

DOI: <https://doi.org/10.22201/iifs.18704905e.2026.1784>

DESMANTELANDO EL SESGO LINGUACÉNTRICO: LA COMUNICACIÓN PRIMATE ES MULTIMODAL

CAROLINA SCOTTO

Universidad Nacional de Córdoba

Instituto de Humanidades

UNC-CONICET- Facultad de Filosofía y Humanidades

Argentina

carolina.scotto@unc.edu.ar

<https://orcid.org/0000-0002-2013-8539>

RESUMEN: Denomino *sesgo linguacéntrico* a una variedad del sesgo antropocéntrico que consiste en abordar los estudios comparados de la comunicación animal tomando como modelo una caracterización formal e idealizada del lenguaje y la comunicación humana. Dicho modelo amplió injustificadamente la brecha entre ambos. Me propongo identificar sus efectos distorsivos con respecto a su compromiso con una caracterización unimodal de las señales comunicativas. Después de argumentar en favor de la adopción de un enfoque pragmático-interactivo de las formas más básicas de la comunicación humana y de la comunicación intencional en los primates no humanos, propongo que el primer paso para desmantelar otros componentes del *sesgo linguacéntrico* es investigar la comunicación primate bajo un paradigma multimodal.

PALABRAS CLAVE: comunicación animal, primates, multimodalidad, lenguaje humano, sesgos

SUMMARY: I call *linguacentric bias* a variety of anthropocentric bias that consists of approaching comparative studies of animal communication based on a formal and idealized characterization of language and human communication. This model has unjustifiably widened the gap between the two. I propose to identify its distorting effects in terms of its commitment to a unimodal characterization of communicative signals. After arguing for a pragmatic-interactive approach to the most basic forms of human communication and intentional communication in nonhuman primates, I propose that the first step in dismantling other components of the *linguacentric bias* is to investigate primate communication within a multimodal paradigm.

KEYWORDS: animal communication, primates, multimodality, human language, biases

Tanto la comunicación humana como la comunicación animal dependen del empleo de variados tipos de señales en distintas *modalidades*, *i.e.*, producidas y receptadas por medio de diferentes canales sensoriales. Este fenómeno suscita una variedad de cuestiones interesantes, por ejemplo, cómo deberían ser conceptualizadas dichas señales, en qué difiere un acto comunicativo unimodal de uno realizado mediante señales en diferentes modalidades, cuál es el papel de los emisores

y los receptores en asegurar la eficacia de la comunicación mediante el uso de señales en distintas modalidades, qué capacidades son requeridas para su producción y su recepción y, en suma, qué carácter le confieren al sistema comunicativo basado en ellas. Estas y otras cuestiones relacionadas aún no han recibido suficiente atención en la literatura filosófica.

Esta omisión puede deberse, al menos en parte, a los efectos distorsivos de un modelo sedimentado acerca de las características definitorias del lenguaje y, sobre ese núcleo, de la comunicación humana, en tanto operó como un estándar explícito o bien por defecto, para el examen comparado de la comunicación animal. De ese modo se edificó lo que llamaré *sesgo lingüacéntrico*. Asentado sobre un conjunto de rasgos que se consideraron definitorios o característicos del lenguaje —entre los cuales el carácter unimodal, predominantemente vocal de las señales lingüísticas y, sobre esas bases de la comunicación humana—, dicho sesgo distorsionó el abordaje comparado de los sistemas comunicativos en otras especies, especialmente en los primates. Me interesa examinar cómo ha impactado en la caracterización de los rasgos y funciones de las señales, lo que a su vez influyó en el examen de otros rasgos de los sistemas comunicativos. Aunque voy a rastrear el cuerpo de supuestos que le dan su contenido específico sobre todo en la literatura filosófica, sus efectos distorsivos se extienden a los estudios científicos sobre la comunicación animal. Finalmente, desde un ángulo más positivo, me ocuparé de mostrar en qué aspectos un cambio de enfoque puede hacer lugar a una perspectiva comparada más fructífera.

En la primera sección caracterizaré el *sesgo lingüacéntrico* como el fruto de la adopción de un enfoque formal e idealizado acerca del lenguaje y de la comunicación lingüística. La combinación de ambos en un modelo unificado alienta una caracterización abstracta o *amodal* de la estructura lingüística y una caracterización *unimodal* del código lingüístico, cuyos usos comunicativos dependen de requerimientos cognitivos sofisticados. Presentaré, luego, los rasgos generales de una perspectiva diferente, pragmático-interactiva, que hace posible sortear el sesgo unimodal. En la segunda sección, me referiré a los modelos de la comunicación animal discutidos en la literatura reciente y a los rasgos, orígenes y funciones de las señales comunicativas. En la tercera sección, introduciré los conceptos de modalidad y multimodalidad empleados en la literatura sobre comunicación animal, identificaré los tipos de señales multimodales y analizaré los principales fundamentos que justifican el interés de estudiar

la comunicación animal y, en particular, la comunicación primate,¹ adoptando un enfoque multimodal. En la cuarta sección, me referiré a algunas evidencias que ponen de manifiesto los efectos distorsivos del sesgo unimodal en los estudios sobre la comunicación primate y, finalmente, presentaré algunos resultados de las investigaciones empíricas que adoptan una perspectiva multimodal, destacando en qué aspectos sugieren una visión enriquecida de la comunicación primate. Por último, resumiré las principales conclusiones alcanzadas.

1. *El sesgo linguacéntrico en la comunicación animal*

Desde *La historia de los animales* de Aristóteles,² la comunicación animal se ha estudiado sobre todo *en contraste* con el lenguaje humano. Este tipo “negativo” de perspectiva comparada, especialmente notorio en la gran tradición filosófica, desestimuló el estudio de las peculiaridades de los sistemas de comunicación de las distintas especies. Antes bien, el interés por ellos estuvo frecuentemente dominado por la búsqueda de sus *diferencias* o elementos de contraste con el lenguaje humano, y con otras capacidades relacionadas. Esta situación es todavía generalizada en la reciente filosofía de la mente animal. Nótese, además, que no es claro en qué ámbito disciplinar deberían encuadrar las reflexiones filosóficas sobre la comunicación animal.

Considerar que somos criaturas únicas, especiales o incluso excepcionales por poseer una u otra característica, podría ser el efecto de un sesgo: el *sesgo antropocéntrico*. El antropocentrismo puede definirse como la perspectiva objetable según la cual ciertas características cognitivas u otras, consideradas distintivas o únicas de los seres humanos, son empleadas como estándar, explícita o implícitamente, para estudiar y evaluar características similares en distintas especies animales (*cfr.* Scotto 2024, para un examen de este sesgo en la literatura comparada). Como su misma formulación lo pone de manifiesto, dicho sesgo impacta en los exámenes comparados, sean estos

¹ Para facilitar la exposición, usaré “animales” para referir a animales no humanos, “primates” para referir a primates no humanos y “comunicación primate” como equivalente a comunicación entre primates no humanos. Sólo cuando sea necesario, distinguiré a los primates humanos de los no humanos y a la comunicación primate humana de la no humana.

² Es interesante que Aristóteles haya vinculado estrechamente la capacidad humana para pronunciar vocales y consonantes con el lenguaje, pues mientras los animales sólo “hacen ruidos o sonidos”, “el lenguaje es la articulación de sonidos vocales y consonantes por medio de la lengua”. Así el lenguaje es posible por la capacidad para producir sonidos *articulados* (Libro IV, parte 9).

episódicos y meramente ilustrativos o sistemáticos y con pretensiones teóricas.

Ahora bien, dado que es razonable intentar una explicación evolutiva de los orígenes del lenguaje humano, bajo el supuesto que la evolución opera mediante la modificación de características ya existentes, es también razonable investigar sus posibles orígenes en especies ancestrales. Como acertadamente señalan Byrne *et al.* (2017) “en el caso de la comunicación de los simios, no es necesario disculparse por un enfoque explícito que la compara, a veces directamente, con aspectos del lenguaje humano [...] Eso es lo que hay que hacer” (p. 755; la traducción es mía). Sin embargo, orientar la comparación sobre la base de un modelo idealizado del lenguaje humano estaría poniendo “patas para arriba el enfoque evolutivo normal” (Owren, Rendall y Ryan 2010). Pero, además, estaría asumiendo otro supuesto objetable: la existencia de una “escala de progreso” que lleva de capacidades o especies menos sofisticadas (*lower*) a más plenamente evolucionadas (*higher*). Así, la llamada “diferencia antropológica” aplicada al lenguaje, *v. gr.*, la manifestación que se considera más plenamente realizada del conjunto de rasgos o capacidades que constituyen el lenguaje, no debería tomarse como estándar para apoyar los estudios comparados. Expresado de otra manera, los estudios comparados sólo están epistemológicamente justificados, y no antropocéntricamente distorsionados, si pueden informarnos acerca de las semejanzas y diferencias entre distintas expresiones del fenómeno que nos interesa. Sólo de esa manera harían posible reconstruir su probable trayectoria evolutiva.

El *sesgo lingüocéntrico* es coextensivo con el *sesgo antropocéntrico* en tanto diferencia de una manera distorsionada sólo a los seres humanos de toda otra criatura. Sin embargo, es más específico que aquél, puesto que depende de un cierto tipo objetable de enfoques acerca del lenguaje. En efecto, más allá de importantes diferencias entre ellos, una familia de enfoques y teorías convergen en marginalizar y, consecuentemente, en explicar insatisfactoriamente algunos rasgos que desempeñan un papel crítico y diferenciado, tanto en la configuración del lenguaje como de la comunicación humana. Así, la concepción que emerge de ambos fenómenos, empleada luego como estándar en los estudios comparados, ensancha injustificadamente la brecha entre ambos y distorsiona doblemente los estudios sobre la comunicación animal. Ahora bien, dado que me interesa explorar un componente del *sesgo lingüocéntrico*, la caracterización *unimodal* de las señales lingüísticas y, consecuentemente, de la comunicación hu-

mana, es importante identificar primero cómo dicha caracterización unimodal es un elemento constitutivo de los enfoques mencionados.

1.1. El lenguaje como un sistema formal

El *mainstream* en la tradición analítica conceptualizó al *lenguaje* como un fenómeno distintivo e incomparable, que subyace y pre-existe a las lenguas históricas y a sus funciones, y que requiere ser explicado en sus rasgos fundamentales con independencia de estos. Bajo este presupuesto general, distintas teorías edificaron un *modelo idealizado* del lenguaje. Empleo el concepto de “modelo idealizado” recogiendo dos acepciones que están presentes en muchas teorías: una esencialista, *i.e.*, “cuáles son sus rasgos constitutivos” y otra normativa, *i.e.*, “cómo debería ser” el fenómeno que queremos estudiar (*cfr.* Keiser 2023). Según dicho modelo, ciertos rasgos son constitutivos del lenguaje, mientras que otros tienen una importancia secundaria o incluso marginal y, además, están contaminados por ingredientes extra-lingüísticos. Aunque el objetivo de edificar una teoría adecuada de cualquier fenómeno complejo requiere la previa elaboración de *modelos*, adoptando, por ejemplo, alguna versión bien fundada de aquella distinción, el riesgo inherente a esa empresa es la identificación errónea de los rasgos que integran el “núcleo” y de aquellos que son desplazados hacia la “periferia”. En tal caso, un modelo ideal podría convertirse en un “arma de doble filo”: no deja entender adecuadamente los rasgos que se identifican como nucleares y opaca o minimiza aquellos que no son objeto de su enfoque. Es decir, el modelo en cuestión podría representar no simplificada sino inadecuadamente el fenómeno que intenta explicar.³

Según el mencionado modelo, el lenguaje presupone, y en algunas versiones incluso consiste en ciertas capacidades cognitivas, y estas se manifiestan en las propiedades formales que lo estructuran. De ese modo llega a ser definido como un sistema sintáctica y semánticamente composicional, o un sistema recursivo jerárquicamente estructurado, compuesto por elementos simbólicos discretos. La propuesta de Chomsky, en sus diferentes versiones, ejemplifica una variedad puramente sintáctica de un enfoque cognitivo y formal de este tipo (Hauser, Chomsky y Fitch 2002). En la literatura filosófica se han

³ Por exceder ampliamente mis objetivos, no presento todos los componentes del modelo al que me refiero, ni tampoco las diferentes enmiendas posibles al mismo: si debiera ser complementado, invertido, o más radicalmente, abandonado. Es posible que el fenómeno lingüístico deba ser conceptualizado como un sistema heterogéneo, abierto e inestable, en distintos niveles de análisis, difícil de capturar en términos de una distinción entre rasgos nucleares y rasgos periféricos.

elaborado teorías de diferente alcance que presuponen, encuadran o son compatibles, en algún grado, con esta perspectiva general, a las que se agregaron las propiedades semánticas de las estructuras fundamentales en las que consiste o sobre las que se edifica el lenguaje humano y que le permiten desempeñar sus funciones primarias. Los ingredientes del fenómeno lingüístico que quedan fuera de este núcleo formal, son empujados hacia la periferia: refieren a su uso y a los factores pragmáticos que influyen en él y sólo explican su funcionamiento.

Según este tipo de enfoques, el sistema lingüístico está conformado por unidades discretas, los fonemas, que, al combinarse entre sí en sílabas y morfemas, hacen lugar a sus unidades léxicas fundamentales. De ese modo, las señales verbales logran el carácter homogéneo y categórico que les permite desempeñar apropiadamente sus funciones semánticas. De acuerdo con esta caracterización, el código lingüístico, conformado por tales unidades, que luego se combinan entre sí mediante reglas, junto a otras propiedades derivadas, puede desempeñar su función primaria distintiva y eficientemente: transmitir “información”, y una variedad de otras funciones que la presuponen.

Devitt y Sterelny (1999), por ejemplo, caracterizaron al lenguaje como un sistema simbólico para la comunicación de los pensamientos que se distingue por un conjunto de “rasgos salientes” que lo convierten en algo “tan especial”. Esos rasgos serían los siguientes: las señales lingüísticas son independientes del estímulo que las causa, tienen un carácter abstracto, son arbitrarias, deben aprenderse, son independientes del medio a través del cual se manifiestan, *i.e.*, son *amodales*, son sistemáticas y versátiles. El conjunto de estos rasgos haría posible identificar la propiedad fundamental del lenguaje: el significado, así como explicarla en términos representacionales. Ello, a su vez, permite dar cuenta del vínculo fundamental que existe entre pensamiento y lenguaje. Mucha literatura filosófica ha puesto el foco en estos rasgos y en su estatus definicional, así como en la caracterización general del sistema lingüístico que contribuyen a delinear.

En la lingüística, una propuesta muy influyente en los estudios comparados sobre la comunicación animal y sobre la evolución del lenguaje fue el modelo de los “rasgos de diseño” del lenguaje de Charles Hockett. En su versión más conocida (1960), el primer rasgo definitorio del lenguaje es “el canal vocal-auditivo” de las señales. Hockett concibe a este rasgo en términos formales: los patrones sonoros en los que consiste el habla están formados por unidades

combinatorias, fonemas y sílabas, que a su vez forman las palabras y las frases, que son los vehículos de información semántica.⁴ Ahora bien, Hockett advirtió que la modalidad vocal determina otros rasgos, aunque ellos sólo emergen del examen de su uso comunicativo: las señales vocales además de transitorias, permiten a los receptores identificar al emisor y a los emisores corregirlas, así como también favorecen la alternancia en la interacción.⁵ Rasgos de diseño como la arbitrariedad de las señales, la semantividad, la transmisión cultural, la aprendibilidad, entre otros, son aquellos en los que Hockett coincide con otros enfoques formales.

Centrados en el sistema lingüístico y no en su uso, este tipo de enfoques también han sido predominantes en los estudios sobre la evolución del lenguaje (*cfr.* Moore y Geurts 2025), llevando a conclusiones más o menos escépticas acerca de la utilidad de los estudios comparados sobre la comunicación animal. Una selección representativa de ellos ocuparía demasiadas páginas, pero interesa destacar que, para la mayor parte de los estudiosos, el carácter vocal de las señales lingüísticas sólo interesa como un medio apropiado para canalizar la estructura fonológica sobre la que se asientan las señales. Cuando los estudios sobre lenguas de señas llevaron a reconocer que el canal vocal no es obligatorio, la visión unimodal se mantuvo y se buscó identificar los mismos rasgos estructurales en la modalidad manual.

El estudio de los sistemas comunicativos animales basado en esta clase de enfoques estuvo sesgado por el intento de identificar alguna manifestación, siquiera mínima, de las propiedades formales consideradas fundamentales del lenguaje, comenzando por su modalidad predominante. Abunda la literatura filosófica y científica que responde a este paradigma. Con variados matices, una conclusión previsible fue que el lenguaje humano es único porque está edificado sobre fundamentos que no se encuentran en ningún sistema comunicativo animal, o bien que no se encuentran todos ellos y/o en ese grado y/o con esa sofisticada complejidad. Para ese tipo de perspectiva comparada, el acento o el interés está puesto *ab initio*, sea inadvertidamente o no,

⁴ Una vez conocidos los estudios sobre las lenguas de señas, Hockett admitió que la modalidad vocal no podía mantenerse como un rasgo definitorio.

⁵ El modelo de Hockett incorpora componentes relativos a las funciones comunicativas del habla, con lo que resulta, a la postre, relativamente híbrido. Además, sostiene que sólo algunos pocos rasgos son únicos al lenguaje. Por estas razones, y por haber desagregado cada rasgo facilitando su operacionalización para la investigación empírica, resultó más versátil que otros para la comparación con los sistemas comunicativos animales.

en la identificación de las discontinuidades entre “el lenguaje humano y la comunicación animal”. Adviértase que así formulada, la cuestión sugiere una manera de conceptualizar las diferencias que, desde otra perspectiva, podría resultar injustificada. En efecto, si fuera posible identificar rasgos comunes en la comunicación primate no humana y humana desempeñando un papel crítico en su configuración, ellos permitirían reconstruir las posibles trayectorias evolutivas que los emparentan. Así, las singularidades de cada sistema comunicativo no son incompatibles con la identificación de rasgos básicos comunes que podrían explicar sus parentescos evolutivos.

1.2. La primacía de la pragmática: el modelo idealizado de Grice

De acuerdo con otro tipo de enfoques, los factores que podrían explicar los rasgos distintivos del lenguaje y sus orígenes evolutivos, serían de naturaleza pragmática. La teoría de H.P. Grice (1957) proporcionó el modelo estándar de la comunicación lingüística, *i.e.*, uno que pretende ser aplicado a todas sus variadas manifestaciones. Según su propuesta, la comunicación es posible en virtud de la implementación de capacidades cognitivo-sociales plenamente desarrolladas, es decir, no sólo competencias intencionales sino “ostensivo-inferenciales” (Scott-Phillips 2015). Estas consisten en “un proceso en dos partes”: (i) el hablante no sólo tiene *la intención de comunicar x —lo que quiere comunicar—*, sino que (ii) tiene también *la intención que su audiencia reconozca i*), haciendo algo que vehicule esa intención (Moore 2017b). A su vez, que la audiencia pueda reconocer efectivamente esa intención comunicativa puede explicar, en parte, que comprenda lo que el hablante quiso comunicarle. Según Grice, además, para que la comunicación sea posible, cada interlocutor debería tener un control intencional pleno de sus participaciones, tanto para atribuir creencias y hacer inferencias como capacidades meta-representacionales para tener y reconocer intenciones comunicativas.

Distintas objeciones fueron dirigidas a rasgos centrales del modelo griceano. Menciono sólo las más relevantes para mis propósitos: i) es posible explicar muchas formas de comunicación humana con una heurística eficiente y capacidades socio-cognitivas implícitas, es decir, sin la intervención de una “teoría de la mente” plenamente desarrollada; ii) la caracterización individualista y unidireccional de la acción lingüística, *i.e.*, centrada en el rol ocupado alternativamente por uno u otro interlocutor, desatiende los rasgos que emergen de la dinámica de la interacción; iii) el modelo es “lingualista” en el sentido que las expresiones verbales se consideran el insumo prin-

cipal de las inferencias que permiten a los receptores reconocer las intenciones comunicativas de los hablantes, lo que opaca el rol crítico y complementario que desempeñan las señales no lingüísticas (cfr. Planer 2023a, 2023b; Scotto 2025).⁶ Es claro que este modelo no puede aplicarse sin muchas enmiendas a un examen comparado de la comunicación animal.

Sin embargo, el modelo griceano alentó la idea de que los posibles precursores de la comunicación humana debían buscarse en sus rasgos pragmáticos e inspiró el enfoque conocido como la *Primacía de la Pragmática* (*Pragmatics-First Approach*). Según este enfoque, la posible trayectoria evolutiva del lenguaje y la comunicación humana requiere identificar sus posibles precursores en la comunicación animal (Moore 2017b).⁷ Su hipótesis fundamental es que el origen y la evolución de las propiedades semánticas y sintácticas del lenguaje deberían explicarse sobre la base de la evolución de los actos comunicativos y las capacidades cognitivas que los hacen posibles, y no al revés (Moore 2017b). En otras palabras, asigna una importancia mayor a los precursores pragmáticos antes que a los sintácticos y/o semánticos, concebidos ahora como la “forma superficial de estos sistemas” (cfr. Rendall, Owren y Ryan 2009; Tomasello 2008).⁸ Esta misma perspectiva, no obstante, dio origen a una nueva versión de la tesis de la singularidad humana, basada esta vez en las competencias cognitivas únicamente humanas que harían posible la comunicación bajo el modelo griceano, reservando el “modelo del código” para los sistemas comunicativos animales (cfr. Scott-Phillips 2015) (véase la sección 2.1).

Los enfoques del tipo *Pragmatics-First* se extendieron más allá de su aplicación al problema de la evolución del lenguaje, para explorar los rasgos que podrían explicar la comunicación en algunas especies animales (cfr. Moore 2016, 2017b; Fröhlich *et al.* 2016; Pika *et al.* 2018; Bar-On 2021, 2025; Moore y Geurts 2025), y también la comunicación humana misma en sus variedades no ideales, es decir, tanto

⁶ El carácter cooperativo-racional de la comunicación humana, que también ha sido objetado, convierte al modelo griceano en un modelo ideal en el sentido normativo antes mencionado. No me referiré a este rasgo, pero también aleja la comunicación griceana de la comunicación animal.

⁷ Ya no se trata sólo de un enfoque sino de un nuevo campo interdisciplinario de estudios: la Pragmática Evolutiva (cfr. Moore y Geurts 2025).

⁸ En la lingüística, los así llamados “enfoques basados en el uso” proponen explicar la emergencia de las estructuras lingüísticas, la “gramática”, a partir de patrones estadísticos basados en el uso de las señales en contextos sociales. Tomasello (2008) ofrece una versión de los orígenes evolutivos del lenguaje apoyado en este tipo de enfoques.

desde una perspectiva ontogenética (Rossano y Liebal 2014; Moore 2017a) como en relación con las interacciones comunicativas adultas más frecuentes, las conversaciones informales cara-a-cara (Levinson 2006, 2019). En cada uno de esos contextos, este enfoque incorpora naturalmente el examen de las funciones comunicativas de las señales no verbales.⁹ De ese modo, después de sortear objeciones en diferentes frentes, el modelo griceano abrió las compuertas para la elaboración de propuestas pragmáticas no ideales.

Así pues, en relación con los enfoques formales e idealizados acerca del lenguaje y la comunicación, las señales lingüísticas se conceptualizan y estudian o bien bajo un enfoque *amodal*, cuando el lenguaje se examina como un sistema formal, o bien bajo un enfoque *unimodal*, cuando se examina como un código comunicativo particular. En el primer caso, se busca identificar las propiedades abstractas o estructurales de las señales, sean estas de naturaleza cognitiva o solo formal, mientras se consideran irrelevantes sus propiedades físicas y perceptibles. En el segundo caso, se busca identificar los rasgos formales propios del canal involucrado, *i.e.*, la forma fonológica que subyace al medio físico predominantemente vocal de las señales verbales, bajo el supuesto de que dichos rasgos hacen posibles las restantes características formales del código lingüístico (y lo propio respecto a las señales manuales en las lenguas de señas). De ese modo, las capacidades perceptuales, motoras u otras involucradas en la producción y la recepción de las señales, así como las propiedades físicas o sensoriales de las señales mismas, es decir, su naturaleza corporeizada y sus instanciaciones o *tokens* cuando son empleadas comunicativamente, no son relevantes para dar cuenta de la naturaleza del lenguaje ni de la comunicación humana. Por cierto, el variado conjunto de señales en otras modalidades que acompañan a la producción del habla cae enteramente fuera de su foco de interés.

Finalmente, este tipo de enfoques dio sustento a las diferentes hipótesis unimodales sobre los orígenes evolutivos del lenguaje. Aquellas que concitaron más atención, *gesture-first* y *vocal-first* (*cfr.* Tomasello 2008) estuvieron guiadas por el potencial atribuido a cada modalidad en contraste con otras, así como a sus características en especies emparentadas, para satisfacer los requisitos más cognitivos del lenguaje humano. Sin embargo, la supervivencia en los humanos de formas híbridas de comunicación, no sólo pone de manifiesto que las señales verbales no reemplazaron a las señales no verbales, sino

⁹ *cfr.* Rossano (2013a) sobre la mirada y Scotto (2022) sobre el rostro.

que el lenguaje pudo haber emergido de ciertas combinaciones entre distintas modalidades (*cfr.* Leavens 2007; Levinson y Holler 2014).

1.3. La primacía de la pragmática: el modelo pragmático-interactivo

Una versión alternativa a la griceana proviene del modelo que llamaré “pragmático-interactivo”. Levinson (2006, 2019, 2022) desarrolló la sugerente hipótesis según la cual el lenguaje emergió en un “nicho ecológico nuclear”, aquel en el que tiene lugar la forma más básica y universal de comunicación humana: la conversación cara-a-cara.¹⁰ La conjetura es que el lenguaje habría evolucionado como un conjunto de herramientas especializadas para favorecer ese tipo de interacciones cercanas, diádicas o en pequeños grupos. Este mismo enfoque también es propuesto para explicar cómo los niños aprenden el lenguaje y cómo es posible la forma de comunicación que empleamos con mayor frecuencia y de la manera más eficiente. Según Levinson, este tipo de comunicación se asienta sobre una “maquinaria interactiva” que consiste en la activación de competencias cognitivo-sociales muy básicas (*cfr.* Scotto 2025), combinada con la implementación de un variado tipo de señales en una dinámica basada en acciones comunicativas contingentemente recíprocas. Dicha “maquinaria” está estructurada mediante mecanismos de secuenciación por turnos, y de coordinación y reparación de fallos, cuya activación está dirigida hacia la comunicación en curso.¹¹ Sus rasgos estructurales y sus funciones también han sido identificados en la comunicación primate (*cfr.* Rossano 2013b; Rossano y Liebal 2014; Fröhlich *et al.* 2016; Fröhlich 2017; Pika *et al.* 2018), como veremos luego. La conjetura sería que el carácter intencional de esta forma de comunicación explica la necesidad de esta dinámica interactiva y que esta, a su vez, podría haber favorecido la emergencia de capacidades de cognición social más sofisticadas. En cualquier caso, la hipótesis acerca de los orígenes “pragmático-interactivos” del lenguaje humano es parcialmente compatible con otras, bajo el supuesto de que la comunicación humana es un sistema “estratificado evolutivamente, compuesto por

¹⁰ El foco puesto en la conversación cara-a-cara hay que rastrearlo en la Sociolingüística y en la tradición del Análisis Conversacional, pero sus diferentes componentes han sido investigados empíricamente por especialistas de diferentes disciplinas (Rossano 2013a y 2013b; Levinson y Holler 2014; Vigliocco, Perniss y Vinson 2014; Levinson 2022, entre otros). Entre los filósofos, Planer (2023b) incorpora las herramientas del análisis conversacional para explicar la transición de la comunicación de los grandes simios hacia la comunicación humana.

¹¹ Planer (2023b) conjetura que estos mecanismos meta-comunicativos prefiguran la evolución de las capacidades meta-cognitivas humanas.

capas de habilidades de diferentes tipos y diferentes antigüedades” (Levinson y Holler 2014).

Según esta perspectiva pragmática, estudiar al lenguaje aislando el habla de sus contextos de uso, crea “ecologías comunicativamente exóticas” (Perniss 2018), en primer lugar, porque no toma en cuenta que el mismo canal vocal es ya multimodal, *i.e.*, para formar sonidos vocales debemos mover los labios y la boca produciendo movimientos visuales además de sonidos, aun cuando sólo sirvan como claves y no como señales (*cfr.* sección 2); y, sobre todo, porque su uso siempre está acompañado por el empleo simultáneo de señales en diferentes modalidades. Esta distorsión se incrementa por el hecho de que la comunicación oral ha sido estudiada tomando como modelo su expresión escrita, *i.e.*, asumiendo injustificadamente que el pasaje de una a otra modalidad sólo deja fuera a las señales no verbales y a otros componentes que se consideraron irrelevantes para el éxito de las acciones comunicativas. Este es otro efecto que deriva de adoptar una perspectiva idealizada del fenómeno que se quiere explicar.

Linell (2005) lo denominó *sesgo del lenguaje escrito* e identificó sus ingredientes en la lingüística, pero todos ellos impregnan gran parte de la filosofía del lenguaje. Este sesgo refiere, en términos generales, a la conceptualización del habla bajo el modelo de la escritura. No puedo analizar aquí sus posibles fuentes y sus efectos, pero basta señalar que este sesgo está montado sobre el *sesgo lingüocéntrico* que, primero, aísla el habla del sistema multimodal en el que tiene lugar su empleo primario. El *sesgo del lenguaje escrito* convierte a las interacciones verbales fluidas, “incompletas” y situadas, en segmentos independientes y plenamente articulados, tanto sintáctica como semánticamente. Este ingrediente del *sesgo lingüocéntrico* apuntaló la visión de que las unidades de análisis del lenguaje poseen un carácter discreto y segmental,¹² estructurado y secuencial. Como señala Perniss (2018), se trata de características que están mejor adaptadas para su uso escrito antes que vocal, y monológico antes que interactivo. Este sesgo es especialmente interesante para nuestros propósitos, porque pone de relieve que la modalidad de las señales no es irrelevante para comprender el carácter de las acciones comunicativas realizadas mediante ellas.

¹² Segmentos son las unidades discretas o los fonemas, vocales y consonantes, que forman las palabras y las secuencias de estas en el habla. Se distinguen de los rasgos supra-segmentales o prosódicos del habla: el acento, el tono, la duración, el volumen, el ritmo, que afectan a unidades mayores que los fonemas, a una parte de una palabra o a una palabra en una emisión verbal.

Por el contrario, cuando el foco es puesto en los contextos primarios de uso *i.e.*, las interacciones cara-a-cara, se advierte fácilmente que el habla integra un sistema dinámico y complejo, del que forman parte los rasgos prosódicos del habla, así como un variado repertorio de otras vocalizaciones, gestos, expresiones faciales, posturas y movimientos corporales. Ese conjunto de señales co-ocurren invariablemente en la acción comunicativa, cumpliendo diferentes funciones, en diferentes combinaciones y de manera integrada. En esos contextos de cercanía, los interlocutores comparten un gran volumen del *common ground*, por lo que no necesitan recurrir a inferencias o conjeturas ni articular esa información para hacer posible o facilitar la interacción comunicativa.

Con estos fundamentos se ha propuesto caracterizar al lenguaje mismo como un *sistema multimodal*, en el cual el uso de una u otra modalidad sólo es más prevalente en distintas circunstancias y/o condiciones, y ninguna de ellas cuenta como un rasgo definitorio o un “modo dado naturalmente por defecto” del lenguaje humano (Levinson y Holler 2014; Vigliocco, Perniss y Vinson 2014; Perniss 2018). En otras palabras, la comunicación lingüística es por defecto multimodal. Introduciré la noción de modalidad en la siguiente sección, pero es suficiente por ahora relacionarla con la más conocida tipología de las señales: vocales, gestuales, faciales, corporales. Me referiré ahora a los distintos tipos de comunicación animal identificados en la literatura y a las señales comunicativas, sus orígenes, rasgos y funciones. Ambas cuestiones serán relevantes para explorar la comunicación multimodal.

2. La comunicación animal, las señales comunicativas

2.1. Comunicación biológica, informacional e intencional

La comunicación animal ha sido mayormente conceptualizada de acuerdo con dos tipos diferentes: *biológica* e *informacional*. La primera consiste en los cambios en el estado de un organismo, la producción de señales, que cambian el estado de otro organismo, la recepción de esas señales, ambos explicables en términos adaptativos (Maynard-Smith and Harper 2003; Rendall, Owren y Ryan 2009; Andrews 2020). Esta es una forma de comunicación que no supone requerimientos cognitivos por parte de los comunicadores. Por mucho tiempo se presupuso que esta era la única variedad atribuible a los animales, determinada biológicamente. Por lo tanto, ni el origen y la producción de las señales, ni los mensajes y los “motivos” por

los cuales tales comportamientos tenían lugar, ni sus efectos en los destinatarios podían compararse con nuestro uso del lenguaje.

La *comunicación informacional*, en cambio, es conceptualizada como un proceso de transmisión de información, desde un señalizador hacia un receptor a través de un canal. Según este tipo de comunicación, también conocido en la literatura como “el modelo del código” (Scott-Phillips 2015), una señal codifica y transmite información en tanto su producción está correlacionada con la probable presencia de ciertos fenómenos o eventos observables, de manera similar al “significado natural” de Grice. La comunicación es exitosa cuando el receptor es capaz de decodificar la información transmitida por la señal, lo que se manifiesta en el carácter predecible de su respuesta. Es decir, para que esta correlación cuente como un proceso comunicativo, debe ocurrir cierto tipo de “interpretación” por parte del receptor. Por lo tanto, la señal informativa no sería un “constructo explicativo último” como en la comunicación biológica, es decir, una correlación entre eventos explicable en términos evolutivos, sino uno “próximo”, en tanto podría formar parte de procesos psicológicos en el receptor (Palazzolo 2025).¹³

La comunicación humana ha sido tradicionalmente concebida como transmisión de información codificada, lo que explica la centralidad del concepto de *significado* tanto para caracterizar el código lingüístico, como sus usos comunicativos más básicos. Bajo ese modelo, se ha intentado elaborar un concepto de información, significado o referencia mejor adaptado para su aplicación a la comunicación animal, al menos en algunas especies. Tal es el concepto de “referencia funcional”, acuñado para abarcar la información que proporciona una señal sobre la presencia de un objeto o evento, sin descansar en una intención comunicativa del usuario (Seyfarth, Cheney y Marler 1980). Los diseños experimentales sobre la comunicación primate fueron moldeados por mucho tiempo bajo los supuestos de este modelo: se procuraba aislar el componente informativo de señales discretas “específicas al contexto” de su producción que pudieran asociarse fácilmente a una respuesta inmediata por parte del receptor, aún en ausencia del estímulo que las causa. Este modelo ha sido cuestionado, entre otros fundamentos, porque intenta transferir una estructura

¹³ Se ha señalado que el concepto de “transmisión de información” es demasiado abarcativo porque puede aplicarse tanto a intérpretes sofisticados como a las bacterias e incluso a fenómenos en los que no intervienen seres vivos, como las máquinas, lo que impide dar cuenta de las diferencias muy relevantes entre ellos (*cfr.* Andrews 2020; Palazzolo 2025).

elaborada para dar cuenta de ciertos usos destacados del lenguaje humano, a la variedad de roles e intereses que orientan las conductas comunicativas en los animales (Rendall, Owren y Ryan 2009; Wheeler y Fischer 2012). Un efecto de este interés sesgado consistió en haber dejado fuera de examen, por mucho tiempo, un enorme y variado conjunto de comportamientos comunicativos que no respondían a ese patrón (*cfr.* Cartmill 2023).¹⁴

Finalmente, la *comunicación intencional*, aplicada a la comunicación animal, se caracteriza por dos rasgos: en primer lugar, las señales se producen de manera voluntaria o flexible y, por lo tanto, se trata de acciones que una criatura podría elegir no realizar y, en segundo lugar, las señales no se producen de manera automática o habitual, sino con dirección a algún fin, más específicamente, con la intención de que ellas o su uso persistente impacten en el destinatario, no sólo en virtud de su contenido sino en virtud de haber sido dirigidas a él por parte del emisor (Tomasello 2008; Andrews 2020; Palazzolo 2025). Ninguno de los dos rasgos requiere asumir que las intenciones y los fines operen de maneras conscientes, y cuando sí lo hacen, no tienen que adoptar el formato articulado de estructuras representacionales explícitas, capacidades para atribuir creencias y para realizar inferencias sobre actitudes e intenciones. Sin embargo, la comunicación intencional descansa en la posesión de ciertas capacidades cognitivas y, por lo tanto, resulta más interesante para explorar los fundamentos pragmáticos de la comunicación primate y los orígenes evolutivos de la comunicación lingüística y del lenguaje mismo. Sin embargo, el *sesgo linguacéntrico*, moldeado sobre la influyente versión ofrecida por el modelo griceano, desalentó por mucho tiempo la hipótesis de que pudiera atribuirse a algunas especies animales.

El concepto “intencional” ha sido operacionalizado mediante un conjunto de criterios para su aplicación en las investigaciones empíricas. Así, un comportamiento comunicativo es intencional si el señalizador: (1) produce una señal selectivamente, teniendo en cuenta la composición de la audiencia, o incluso, dirigiéndola a cierto individuo, (2) asegurándose que el receptor pueda percibirla o bien manipulando su atención mediante la producción de señales en diferentes modalidades, y (3) persiste en el uso o reelabora la señal en

¹⁴ En el marco de este modelo se estudió la producción de los gritos de alarma en distintas especies primates. Pero este tipo de comportamiento no puede explicarse adecuadamente con el modelo informacional, porque es sensible a la composición de la audiencia y rara vez ocurre en soledad, lo que, a su vez, podría explicar, entre otros rasgos, su estructura, frecuencia y duración (Fedurek y Slocombe 2011).

caso que no haya sido adecuadamente recibida, es decir, si no produce un cambio en la conducta del receptor (Liebal *et al.* 2014).¹⁵ Los rasgos señalados pueden ser más eficientemente satisfechos mediante el empleo de dos o más señales comunicativas, simultáneas o sucesivas, desempeñando la misma o diferentes funciones cada una. En contraste con el modelo informacional, el carácter interactivo, es decir, dinámico, flexible y recíproco del comportamiento de cada participante es constitutivo de la naturaleza intencional del fenómeno comunicativo.

Ahora bien, como señala Andrews (2020), la segunda característica de la comunicación intencional —que esté dirigida a impactar en un receptor— puede satisfacerse de muy diferentes maneras. Como vimos, según la propuesta de Grice, la comunicación intencional descansa en capacidades socio-cognitivas sofisticadas, *i.e.*, una “teoría de la mente” plenamente desarrollada, tanto en el emisor como en el receptor. Diferentes propuestas han intentado identificar variedades menos sofisticadas de esas capacidades. La variedad “mínimamente griceana” de Moore inauguró un conjunto de otras formulaciones neo-griceanas. Entre ellas se pueden mencionar la “intermedia” de Bar-On (2021; 2025) y la “denettiana” de Planer (2023a). Es posible que cada una pueda explicar diferentes rasgos de la comunicación primate.

La comunicación intencional se produce en contextos sociales en los que el señalizador espera una respuesta del receptor y puede persistir hasta obtenerla. Es decir, se requiere que el emisor haga algo para que el receptor también haga algo como efecto de la acción del primero. Así, diferentes comportamientos deberían estar presentes en este tipo de comunicación, en una “estructura intencional de dos niveles” (Tomasello 2008, p. 32): mientras algunas señales “llaman la atención hacia sí mismas”, son seguidas por otras que requieren percepción directa y rastreo de intenciones, tales como pedir u ofrecer algo (p. ej., en contextos de juego, alimentación, acicalamiento), compartir información (p. ej., alertar sobre un peligro), expresar actitudes y estados emocionales (p. ej., manifestar enojo, compasión, impaciencia).

Como dijimos, el mismo tipo distintivo de comunicación interactiva o “conversacional” ha sido identificado en especies altamente

¹⁵ Se han propuesto criterios adicionales, pero no todos ellos excluyen variedades no intencionales, ni están siempre presentes en comportamientos comunicativos que sí lo son (Liebal *et al.* 2014). Para un análisis detallado de estos y otros criterios, *cfr.* Townsend *et al.* 2017.

sociales como los primates: entre las madres y sus crías, entre pares del mismo o diferente sexo, entre parejas y en pequeños grupos. Esta variedad de comunicación intencional, hibridada a partir de los recursos de la comunicación biológica, es realizada mediante señales especializadas que tienen diferente origen, como se verá en la sección 2.2. Su empleo combinado en diferentes modalidades explica en gran medida la eficacia de estas interacciones, como se verá en la sección 3.

Al integrar a emisores y receptores en una estructura interactiva, esta forma de comunicación contrasta con la visión individualista y estática que ofrecen tanto el modelo informacional como el intencionalista griceano: una secuencia lineal que comienza activamente en el emisor y concluye pasivamente en el receptor. Aunque la producción de una señal puede no recibir una respuesta adecuada o incluso no recibir ninguna, las señales evolucionan, se desarrollan o aprenden conjuntamente con las respuestas a ellas, y así también se modifican. Incluso los repertorios innatos son el producto de “una relación bi-direccional”, bajo la influencia de una presión selectiva aplicada sobre la interacción misma (Rossano y Kaufhold 2021). Así, en esta forma de comunicación interactiva, diádica o entre pocos individuos, muchas veces no es sencillo determinar quién es el emisor y quien el receptor. Dado que tales comportamientos se producen de manera dinámica e incluso sincrónicamente, en muchos casos tendrían que conceptualizarse como “interacciones correguladas” (Shanker y King 2002;¹⁶ Fröhlich *et al.* 2016). Como veremos enseguida, su carácter multimodal es tanto un ingrediente fundamental como un indicador de su complejo carácter interactivo.

2.2. Señales y claves, repertorios y usos

En la literatura sobre comunicación animal se ha consolidado la distinción entre *señales* y *claves*. La definición canónica de *señal* es cualquier rasgo o comportamiento “que altera los comportamientos de otros organismos, que evolucionó a causa de ese efecto, y que es efectiva porque la respuesta del receptor también evolucionó” (Maynard-Smith y Harper 2003, p. 3; la traducción es mía). Una morfología o un comportamiento es una *señal* si satisface esas condiciones: tiene un carácter *adaptativo* para los señalizadores porque

¹⁶ Shanker y King (2002) elaboran un modelo alternativo al informacional, lo que se refleja en la importancia que asignan a los aspectos dinámicos e interactivos de la comunicación en primates.

evolucionó en virtud de su capacidad para *causar regularmente ciertos efectos en otros organismos*. Es decir, las señales tienen una historia evolutiva, son específicas de cada especie y deberían ser conceptualizadas como adaptaciones biológicas.

Una señal es adaptativa porque influye de un modo también adaptativo sobre la respuesta del receptor. Es decir, no habría evolucionado si no hubiera tenido el efecto, bajo ciertas condiciones y con alguna probabilidad, de producir la respuesta apropiada, tanto es así que esta sirve de base para identificar su función. En síntesis, no es posible explicar este tipo de comunicación sin tomar en cuenta los roles activos y de mutuo ajuste de emisores y receptores.¹⁷ Las *claves*, por el contrario, serían rasgos o comportamientos que pueden proporcionar información o servir como guía a un receptor apropiado, pero que no tienen por función hacerlo, es decir, lo hacen sólo como un efecto concomitante o incidental.¹⁸ Otras distinciones podrían ser necesarias, por ejemplo, entre claves comportamentales y rastros de comportamientos que proporcionan información útil a otras criaturas.

Ahora bien, algunas claves podrían dar origen al surgimiento de señales a través de procesos de “ritualización filogenética” por medio de los cuales un comportamiento innato y característico es exagerado y estereotipado, se hace más conspicuo y redundante y es “precedido por componentes de alerta” (Tinberger 1952). Un ejemplo de ellos serían algunos “movimientos de intención”: señalan una acción que el emisor quiere realizar mediante un fragmento o abreviatura de ella (Tomasello 2008). Otras hipótesis se han propuesto para dar cuenta de la emergencia de las señales animales. Por ejemplo, la hipótesis del “reclutamiento” propone que parte del repertorio gestual en los grandes simios hereda sus rasgos de algunos esquemas corporales y de partes y estados del cuerpo (Graham, Rossano y Moore 2025). Estos serían “reclutados” en virtud de que su producción activa asociaciones con la realización de ciertas acciones que es aprovechada

¹⁷ Tomasello (2008) propone una distinción diferente: entre *exteriorizaciones*, “acciones físicas y comportamentales creadas y controladas por procesos evolutivos” y *señales*, elegidas y producidas por los individuos y que forman parte de la comunicación intencional.

¹⁸ Aunque la distinción entre *señal* (o *signo*) y *clave* es muy conocida, todavía es frecuente en la literatura comparada referir a todo rasgo o comportamiento animal que participa de un fenómeno comunicativo como una *clave*, reservando implícitamente el concepto de *señal* para sistemas como el lenguaje humano. Lo propio ocurre con las señales no verbales en la comunicación humana, a las que se hace referencia frecuentemente como *claves*. Esta manera de conceptualizar la distinción, más allá de la terminología elegida, es también un efecto del *sesgo lingüocéntrico*.

con fines comunicativos (p. ej., la cría que levanta un brazo, dirigido hacia la madre, indica con qué parte del cuerpo ella debería interactuar y cómo). Dentro de esta familia, un rasgo o comportamiento que es el fruto de un proceso adaptativo, puede convertirse en una señal social a través de mecanismos socio-cognitivos que “amplifican” las características que permiten detectarlos mejor, cuando son producidos voluntariamente y con fines comunicativos, como es el caso de muchas señales faciales (Glazer 2019). Por ejemplo, los rasgos dinámicos del rostro (la dirección de la mirada, los movimientos de las cejas, los labios, la cabeza), cuando pueden someterse al control voluntario, desempeñan funciones comunicativas específicas que podrían emerger de un mecanismo de ese tipo.

Investigaciones sobre gestos y vocalizaciones en distintas especies animales revelan que estas señales podrían tener “significados” específicos a una población e incluso a individuos (Rossano y Kaufhold 2021). Por ejemplo, se han descrito “dialectos” en las señales vocales en diferentes grupos de chimpancés salvajes (Crockford *et al.* 2004; Fröhlich y van Schaik 2020). En estos casos, se trataría de señales “inventadas” mediante procesos de “ritualización ontogenética”. Por ejemplo, ciertas secuencias repetidas de comportamientos sociales, como “pedir”, se abrevian y convierten en señales comunicativas como “el gesto de alcanzar”. Algunas señales podrían explicarse, en cambio, por medio de “la transmisión y el aprendizaje social” a través de la imitación (Tomasello *et al.* 1994), sobre todo en contextos de interacciones sociales repetidas entre los mismos individuos y en especies capaces de comunicación intencional. La hipótesis de la “negociación social”, finalmente, sugiere que los individuos convierten ciertos comportamientos sociales habituales en comunicativos como el fruto de las interacciones entre ellos, que se adaptan y refinan sobre la marcha, hasta lograr un acuerdo recíproco sobre su uso que, luego, se extiende a otros miembros del grupo (Pika y Fröhlich 2019). Esta hipótesis también explicaría su variabilidad en distintas comunidades.

Así, por ejemplo, gestos como los “movimientos de intención”, a los que nos referimos antes, y los “llamados de atención”, también podrían tener un origen aprendido en los chimpancés (Tomasello 2008). Entre los primeros, por ejemplo, “tocar la espalda” de la cría hacia la madre, para que ella se agache y pueda treparse. Entre los segundos, por ejemplo, “hacer golpes en el piso”, “empujar” o “arrojar objetos”, tienen la función de orientar al receptor hacia el emisor para que su acción comunicativa ulterior pueda ser percibida. También se han descrito diferencias para producir las mismas

señales más o menos correctamente y en los contextos apropiados, en distintas especies primates, entre las crías y los individuos adultos (Ghazanfar 2013). En síntesis, estos hallazgos justificarían distinguir entre señales adquiridas como adaptaciones y señales adquiridas a través de otros mecanismos dependientes de la interacción social: ritualizadas, reclutadas, amplificadas, imitadas, negociadas. Fröhlich y van Schaik (2020) sugieren, incluso, que algunas de estas señales podrían estar compuestas por elementos con orígenes mixtos: innatos y adquiridos.

Por lo tanto, el origen y la transmisión cultural de las señales lingüísticas que para Hockett era un “rasgo de diseño” del lenguaje considerado únicamente humano, podría no tener ese carácter. Se ha señalado también que, contrariamente a cierto consenso en torno a una distinción entre las señales gestuales y las vocales, basado en su carácter más flexible o más rígido o innato, respectivamente, ambas tienen una importante base innata y también ambas son susceptibles de usos más o menos flexibles e intencionales (*cf.* Seyfarth y Cheney 2018; Fröhlich *et al.* 2019; Schel *et al.* 2013). Así, las diferencias atribuidas por mucho tiempo a ambos tipos de señales podrían explicarse, al menos en parte, con base en los supuestos más demandantes que apoyaron las diferentes estrategias de investigación empleadas para estudiar la modalidad vocal a diferencia de la gestual (*cf.* Townsend *et al.* 2017).

Finalmente, Bar-On (2021) ha propuesto una interesante distinción entre el *repertorio* y el *uso* de las señales: mientras el primero (o una parte de él) puede estar limitado a un origen adaptativo y ser fijo y común en cierta especie, su uso puede ser flexible en distintos contextos e incluso variable en diferentes comunidades de esa misma especie. Además, los componentes fijos de ese repertorio pueden combinarse de diferentes maneras en circunstancias distintas, incrementando el número de mensajes que es posible comunicar intencionalmente con un repertorio más o menos estable, como veremos luego. Vale la pena agregar que el repertorio de las señales primates es muy variado: diferentes tipos de gritos, silbidos y otras vocalizaciones, bailes, zapeos, señalamientos, contacto corporal, movimientos de los brazos, del torso y la cabeza, miradas, gestos manuales, posturas corporales, expresiones y gestos faciales, etc. Como veremos en la sección 4.2, esos repertorios varían en distintas poblaciones y grupos de edad y también su uso según los contextos sociales y las condiciones ambientales. Así, la distinción entre repertorio y uso permite apreciar cuánto gana la comprensión del fenómeno comunicativo cuando se incorpora una perspectiva funcional. Dado que uno de los componentes críticos

de la estructura pragmático-interactiva de la comunicación en los primates es el uso de señales en diferentes modalidades, consideremos ahora el papel de las modalidades en la señalización animal.

3. Señales y modalidades en la comunicación animal

3.1. Las señales multimodales

Distintas cuestiones están involucradas en los conceptos de *modalidad* y *multimodalidad* en los estudios sobre comunicación animal. *Modalidad* se aplica a las señales mismas, a sus propiedades físicas y morfológicas perceptibles y al canal sensorial y/o a los recursos corporales empleados por los comunicadores para producir y reconocer dichas señales. En este último sentido, modalidad y canal sensorial suelen emplearse de manera intercambiable.¹⁹

Las *señales* comunicativas se clasifican, entre otros criterios, según el sistema sensorial principal mediante el cual son *producidas* o bien según el cual son *percibidas*, pero también incorporando ambos. Así, por ejemplo, una misma señal puede ser o vocal o auditiva o vocal-auditiva, respectivamente.²⁰ Siguiendo esta conceptualización, también se emplean de manera indistinta los términos *multimodal* y *multisensorial* (Partan y Marler 2005). Mientras los psicólogos comparados suelen emplear el primer criterio, que pone el foco en el señalizador (Fröhlich y van Schaik 2018; Fröhlich *et al.* 2019; Hebets 2011; Liebal *et al.* 2022), los ecólogos y biólogos las clasifican desde el punto de vista de su recepción (Partan y Marler 2005; Higham y Hebets, 2013). Por diferentes razones, es importante tomar en cuenta ambos criterios y, además, la señalización misma: la forma o estructura y la función de las señales (Hebets 2011).

Según su estructura o morfología distintiva, las señales se clasifican en vocalizaciones y otros sonidos, gestos, señales faciales, movimientos corporales, etc.²¹ Esta taxonomía no está basada en el tipo de

¹⁹ Un uso más restrictivo del concepto equipara los rasgos modales de un estímulo con aquellos que son específicos de un canal sensorial, como el sabor dulce respecto al gusto. Según este criterio, rasgos como el tamaño y la textura de un objeto serían *amodales* (cfr. Spence 2011). Pero esta no es la acepción usual de ambos conceptos en la literatura sobre comunicación multimodal.

²⁰ Cuando se hace referencia tanto a la producción como a la recepción de una señal, p. ej., vocal-auditiva, se entiende que la señal es unimodal por estar involucrado un único canal sensorial tanto para producirla como para recibirla. Así mismo, se trata de un único tipo de señal, por ejemplo, una vocalización o una señal vocal.

²¹ En el caso humano, esta tipología se diversifica aún más: señas, dibujos, signos escritos ejemplifican modalidades diferentes, aunque todas son visuales desde el

soporte material que vehiculiza las señales ni en el canal sensorial con el que son producidas o percibidas: en los primates, los gestos no pueden realizarse sólo mediante movimientos corporales, sino que también pueden ser táctiles, faciales o incluso vocales (Liebal *et al.* 2014). Esta manera de conceptualizar las modalidades, que pone el foco en las señales mismas, permite distinguir, por ejemplo, entre dos señales en las que están involucrados los mismos canales sensoriales, pero difieren no sólo por sus características morfológicas sino también por sus funciones. Así, por ejemplo, una señal facial y una gestual son ambas visuales, pero pueden satisfacer diferentes funciones.

Aunque la comunicación multimodal es ubicua en los animales, el uso comunicativo más o menos prevalente de unas u otras modalidades varía con las especies, dependiendo de sus características biológicas y de los entornos o hábitats en los que se desenvuelven, pero también de sus estilos de vida. Aunque los primates recurren a los llamados sentidos químicos, el olfato y el gusto, para satisfacer diferentes funciones, sólo el olfato y los llamados sentidos físicos, la visión, la audición y el tacto, son empleados con específicas funciones comunicativas (Liebal *et al.* 2014). Dado que los primates comparten con los humanos una morfología similar y recurren a las mismas modalidades, en algunos contextos sociales también similares, la comparación interespecífica se ve claramente facilitada (Andrews 2020).

Las señales multimodales se definen como “las combinaciones simultáneas de señales de dos o más modalidades (señales gestuales, faciales, vocales y olfativas), y/o cualesquiera señales que requieran integración sensorial por parte del receptor” (Liebal *et al.* 2014, p. 115; la traducción es mía). Según esta caracterización inclusiva, una acción comunicativa es multimodal si es el fruto de la producción simultánea de dos o más señales en diferentes modalidades, por lo que su recepción adecuada requiere integrarlas, no sólo a nivel perceptivo, sino incluso a nivel cognitivo y neural. El concepto refiere tanto a señales donde la relación entre los componentes puede ser “fija” u obligatoria como a aquellas donde los componentes se combinan de manera “libre” o “flexible (Partan y Marler 2005), y son producidas simultáneamente o con un mínimo hiato temporal entre ellas.

punto de vista de su percepción. Las señales también se clasifican según la parte del cuerpo o el órgano con los que una criatura señala, *v. gr.* los articuladores: los ojos, el rostro, las manos, el cuerpo, en oculares, faciales, manuales, corporales o posturales, respectivamente.

En la literatura más reciente es frecuente distinguir entre *señales multimodales*, para referir a aquellas que consisten en el acoplamiento obligatorio de dos o más componentes de diferentes modalidades sensoriales (p. ej., movimientos de la boca que producen también sonidos) y *combinaciones de señales multimodales* o también *señales multicomponentes*, que son los acoplamientos flexibles de dos o más señales en diferentes modalidades (p. ej. un gesto y una vocalización). (Fröhlich y van Schaik 2018; Fröhlich *et al.* 2019).²² Dado que las primeras incluyen componentes fijos en los que al menos uno no es producido voluntariamente, participan de una forma de comunicación que puede requerir menor complejidad cognitiva que aquella donde ambas señales se producen y combinan intencionalmente para producir más eficazmente un acto comunicativo. Como veremos luego, sólo algunas de estas *señales multicomponentes* tienen un carácter composicional, *v. gr.*, integran señales que ya poseen significado por sí solas que se combinan para modificar o generar nuevos significados.

El interés de este tipo de señales se puede apreciar en términos de las ventajas que podría proporcionar a una criatura recurrir a la producción simultánea o sucesiva de señales en diferentes modalidades para realizar un mismo acto comunicativo. Su eficacia depende, a su vez, de la capacidad de los receptores para detectarlas y para establecer las adecuadas correspondencias entre ellas. En efecto, el empleo de más de una modalidad sensorial puede estar orientado a la realización de un único acto comunicativo, por ejemplo, dos señales en diferentes modalidades pueden transmitir un mismo mensaje, reforzando su percepción por parte del destinatario. En tal caso, ambos componentes tendrían una función *redundante* (*cfr.* Partan y Marler 1999; 2005). Este tipo de uso de las señales multimodales responde a lo que se ha sido denominado “la táctica de la repetición” (Leavens 2007). Este podría ser el caso donde el empleo de más de una modalidad permite identificar mejor una señal o no confundir su contenido en un contexto ruidoso. También suelen emplearse cuando la producción de una señal no logra satisfacer su función, sea porque no genera ninguna respuesta o porque no genera una respuesta apropiada. En esos casos, el emisor puede remediar el fallo repitiendo la misma señal o bien emitiendo una nueva en otra modalidad que “per-

²² Una señal puede incluir más de dos componentes en distintas modalidades, y cada componente puede estar dirigido a diferentes destinatarios. Por otra parte, las señales pueden ser compuestas, pero en una misma modalidad (en uno de los sentidos antes señalados), por ejemplo, señales faciales y posturales, ambas en la modalidad visual. Lo propio ocurre con las demás modalidades, e incluso con más de dos modalidades. Por razones de simplicidad, me referiré sólo a señales bimodales.

sista” en el objetivo de la anterior, no sólo reiterando sino también “refinando” o “complementando” la información transmitida en la primera señal (Fröhlich y van Schaik 2018). La evolución de este tipo de señales son una muestra de la importancia que tiene para el emisor su apropiada recepción. En ese sentido, la multimodalidad “proporciona una base física para la redundancia” (Leavens 2007).

La producción de señales multimodales también puede explicarse por la satisfacción de funciones comunicativas diferentes. Así, mientras una señal en una modalidad transmite el contenido, otra señal, en una modalidad diferente, tiene la función de llamar la atención del receptor hacia la primera, es decir, no opera sobre su contenido sino sobre el acto que lo vehiculiza. En estos casos, la producción no simultánea de una combinación de señales diferentes podría favorecer la eficacia del acto comunicativo, donde cada componente mantiene con el otro una relación *no redundante* entre sí (*cfr.* Partan y Marler 1999; 2005). También podría producirse una combinación en la que las señales tienen funciones complementarias respecto a la acción realizada, en cuyo caso ejemplifican una forma de sintaxis combinatoria no jerárquica (Partan y Marler 2005). Así, por ejemplo, un componente gestual tiene la función de señalar un objeto y otro facial o vocal hace explícita la intención con la que el primero es producido, por ejemplo, la acción de pedir ese objeto. Es claro que este tipo de señales multimodales hace más complejos los actos comunicativos realizados mediante ellas: al desempeñar funciones diferentes, su empleo responde a la “táctica de la elaboración” (*cfr.* Leavens 2007).

Un ejemplo característico de señales no redundantes del primer tipo es la combinación entre una señal que tiene la función de “llamar la atención” del receptor hacia el emisor (p. ej., produciendo sonidos o haciendo gestos) para luego hacer explícita, mediante otra señal, su intención social (p. ej., jugar, obtener comida, acicalamiento u otra). El acto comunicativo tiene, pues, una estructura compleja en dos niveles, donde cada componente satisface una función diferenciada en la secuencia comunicativa completa.²³ Por otra parte, cuando las señales vocales se emplean en contextos de proximidad también pueden desempeñar otras funciones, tales como mantener el contacto activo,

²³ Las señales unimodales admiten un “análisis descomposicional”, pero de otro tipo, p. ej., una señal vocal puede emitirse o bien con un volumen elevado y/o un ritmo agitado o bien con un volumen más bajo y/o un ritmo más lento. Los estudios sobre las diferentes funciones de los gritos de alarma en los primates identificaron, entre otras, las diferentes propiedades prosódicas de las señales acústicas que contribuyen a especificar el tipo de acción comunicativa que intentan realizar (*cfr.* Filippi 2016).

marcar coordinación o regular la toma de turnos entre quienes participan en la interacción, desempeñando una función organizacional o metacomunicativa en la conversación multimodal (*cfr.* Levinson 2016 sobre la abundante evidencia de estas funciones de las vocalizaciones en distintas especies primates).

Por su parte, la diferencia entre señales redundantes y no redundantes puede identificarse según cuál sea la respuesta del receptor: en el primer caso, los efectos de cada componente, considerados independientemente, deberían ser equivalentes, mientras que, cuando son usados en combinación, deberían producir una respuesta mayor, más “acentuada” (*enhanced*) o más refinada (Hebets y Papaj 2005; Hobaiter *et al.* 2017). En el segundo caso, en cambio, los efectos tendrían que ser diferentes puesto que sus componentes tienen funciones diferentes. A su vez, los efectos causados por la combinación entre dos componentes no redundantes pueden ser de cuatro tipos diferentes: *independencia*, ninguno altera el efecto del otro; *dominación*, uno opaca al otro; *modulación*, uno altera el efecto del otro; o *emergencia*, la combinación de ambos componentes suscita un efecto novedoso (Partan y Marler 1999).²⁴ En el caso de las señales redundantes, son más frecuentes las que acentúan el mensaje y en el caso de las no redundantes, la modulación, siendo la menos frecuente la combinación en la que una señal domina sobre la otra (Partan y Marler 2005).²⁵ El uso de señales no redundantes en los dos últimos tipos de combinaciones flexibles, *modulación* y *emergencia*, ejemplifica una *división del trabajo comunicativo* más compleja, análoga a la que desempeñan distintas señales en la conversación humana cara-a-cara.

Una adecuada conceptualización de las señales multimodales requiere identificar sus repertorios en cada especie y deslindar las combinaciones aleatorias de las que no lo son, así como las fijas de las flexibles. Además, es importante tomar en cuenta no sólo el contenido de las señales sino también su eficacia, que depende de factores contextuales, otro indicador de su posible producción intencional. Muchos estudios observacionales sugieren que el uso de señales multimodales con repertorios variados en combinaciones flexibles y con diferentes funciones, está asociado con demandas sociales exigentes

²⁴ Una clasificación más compleja puede verse en Hebets y Papaj 2005.

²⁵ En la comunicación humana, las señales gestuales y faciales frecuentemente modulan las señales verbales, o bien refinando o complementando algunos contenidos semánticos de la emisión verbal, o bien haciendo más nítida o explícita la fuerza ilocucionaria del acto comunicativo, o bien desempeñando roles metacomunicativos sobre la dinámica conversacional.

que, a su vez, se tornan más robustas y complejas a medida que aumenta el tamaño del grupo y la complejidad de las relaciones sociales entre quienes se comunican (Ghazanfar 2013). A ellos se agregan otros factores como la edad, el sexo, el parentesco, el rango social y los vínculos entre los comunicadores, así como los diferentes contextos sociales en los que las señales son empleadas (*v. gr.*, juego, agresión, reconciliación, acicalamiento, entre otros).

Como muchos especialistas señalan, los beneficios que proporcionan las señales multimodales flexibles a emisores y receptores vienen acompañados de costos mayores en comparación con las señales unimodales, porque no sólo exigen más esfuerzo para su producción y su detección, sino que también exponen más a los comunicadores a posibles predadores, contrincantes o competidores. Ello también podría dar cuenta de la mayor o menor proporción relativa del uso de señales unimodales y multimodales, según los contextos. La importancia relativa de estos diferentes factores pone de relieve que una adecuada comprensión de la comunicación primate no puede limitarse a identificar las intenciones comunicativas y las capacidades cognitivas de las que depende el uso de uno u otro tipo de señal unimodal por parte de los emisores.

4. *Hacia un enfoque multimodal*

4.1. Diagnosticando el sesgo unimodal

Es notable que el marco teórico inaugurado por los trabajos de Par-tan y Marler (1999) y Rowe (1999), entre otros, haya debido esperar tanto tiempo para captar el interés de los primatólogos, donde los estudios unimodales han sido más prevalentes que en otras especies (Fröhlich *et al.* 2019; Wilke *et al.* 2017). En efecto, los estudios multimodales sobre la comunicación en diferentes especies ha sido la norma, por ejemplo, en artrópodos, batracios, abejas, aves y mamíferos (*cfr.* Liebal *et al.* 2014, pp. 118–121 y p. 127). No parece arriesgado conjeturar que esta asimetría obedece, entre otros posibles factores, al sesgo unimodal acerca de la comunicación humana transferido a los estudios comparados y a los estudios evolutivos sobre el origen del lenguaje humano.²⁶ Un análisis bibliográfico realizado

²⁶ Otros factores también explican la brecha entre las investigaciones unimodales y las multimodales, como el hecho que tanto el registro como el análisis de las señales multimodales es más complejo y genera obstáculos metodológicos y requisitos de especialización en los grupos de investigación (*cfr.* Liebal, Slocombe y Waller 2022; Cartmill 2023).

por Slocombe, Waller y Liebal (2011) reveló que los estudios publicados sobre la comunicación primate, tanto observacionales como experimentales —un total de 553 publicaciones—, en un periodo que cubre 5 décadas, entre 1960 y 2008,²⁷ estuvieron mayoritariamente dedicados —un 95 %— a las señales comunicativas espontáneas,²⁸ en una u otra modalidad sensorial de las tres que se consideran principales: vocal, gestual y facial (es decir, excluyendo las modalidades olfativo-gustativa y la táctil).²⁹ Sólo un ínfimo porcentaje del total relevado de las investigaciones —menor al 5 %— buscaba integrar más de una modalidad (28 publicaciones).

El conjunto reveló, además, un claro predominio de los estudios dedicados a la modalidad vocal (el 64 %), frente a la modalidad facial (22 %) y a la gestual (sólo el 9 %). Las estrategias de investigación empleadas, en relación con cada modalidad, eran muy diferentes: la comunicación gestual era estudiada sobre todo en grandes simios, con animales cautivos y usando tanto métodos observacionales como experimentales, mientras que las señales faciales, eran estudiadas en monos, también cautivos, y usando métodos observacionales. En cuanto a las vocalizaciones, tanto en simios como monos, cautivos o en escenarios naturales, eran estudiadas sobre todo con métodos experimentales. En otras palabras, las modalidades vocal y facial fueron escasamente estudiadas en grandes simios y muy poco en escenarios naturales. En cuanto a la modalidad vocal en particular, fue muy poco estudiada con métodos observacionales, mientras que escasos estudios experimentales se dedicaron a la comunicación facial. En cuanto a la modalidad gestual, fue apenas estudiada en monos y en escenarios naturales.

Estos registros pusieron en evidencia, en primer lugar, que la comunicación primate fue muy poco estudiada “holísticamente”, antes bien se focalizó en una u otra de las modalidades predominantes, consideradas separadamente. El supuesto que podría explicar esta estrategia es que el interés por la comunicación primate está guiado implícitamente por el modelo unimodal del lenguaje humano. Por otra

²⁷ La búsqueda se limitó a revistas especializadas de habla inglesa utilizando las tres bases de datos más relevantes, y se basó en términos como: comunicación o expresión facial, despliegue facial, despliegue y conducta, gesto, comunicación gestual, vocalización, y mono, macaco, chimpancé y similares.

²⁸ La conversación espontánea difiere de aquella en la que se solicita a los primates que interpreten o utilicen señales artificiales humanas.

²⁹ El relevamiento también identificó la especie primate estudiada, si se trataba de animales en cautiverio o en sus entornos naturales, y si los estudios se focalizaban sobre el productor o sobre el receptor de las señales.

parte, la fragmentación de la investigación en una u otra modalidad y con diferentes aproximaciones metodológicas no sólo obstaculiza la integración de sus resultados, sino que impide detectar la interacción funcional de las señales en diferentes modalidades. Como señalaron sus autores años después, esta publicación estimuló un “creciente consenso teórico” en favor de encarar investigaciones multimodales o, al menos, sobre la necesidad de integrar los hallazgos de los estudios en diferentes modalidades. Sin embargo, en las investigaciones empíricas se mantuvo el predominio del paradigma unimodal.

Dos revisiones bibliográficas más recientes arrojaron resultados que ratificaron la vigencia de ese paradigma. Uno de ellos, sobre distintos sesgos en la investigación empírica sobre comunicación animal, abarcó más de 900 artículos especializados, publicados en el periodo 2005–2015. Los resultados evidenciaron un notable predominio de los estudios en la modalidad vocal en primates, un 58 %, frente a sólo un 7 % dedicados a la comunicación multimodal (Ullrich, Mittelbach y Liebal 2018). La otra revisión fue realizada por el mismo grupo de investigadores que hizo el estudio antes mencionado del año 2011, esta vez sobre 294 artículos sobre comunicación primate, publicados en revistas especializadas entre 2011 y 2020, en una o en varias de las tres modalidades más estudiadas. Aunque esta revisión revela algunos progresos, por ejemplo, un incremento de los estudios sobre vocalizaciones en grandes simios y de estudios experimentales sobre expresiones faciales, entre otros, el sesgo unimodal o el “velo del lenguaje” continuó generando una brecha. Ello también dificultó los avances en el campo de la evolución del lenguaje humano (Liebal, Slocombe y Waller 2022).

Estas investigaciones bibliográficas revelaron distorsiones sistemáticas que orientaron la investigación sobre la comunicación en primates incluso hasta las primeras décadas de este siglo, y ponen de manifiesto la necesidad de un sustantivo cambio de enfoque. Muchos investigadores concuerdan en esta conclusión (Leavens 2007; Pika *et al.* 2014; Genty *et al.* 2014; Wilke *et al.* 2017, entre otros). El *sesgo unimodal* explica, además, que no se haya prestado igual atención a todas las modalidades comunicativas en la proporción con las que son empleadas por cada especie, y que no se haya recurrido para su estudio a las mismas estrategias metodológicas en relación con los mismos contextos, lo que impide que los resultados sobre una modalidad sean comparables con los evidenciados en las demás. Por ejemplo, mientras los estudios sobre la modalidad vocal se orientaron tanto hacia las conductas del emisor como del receptor, en escenarios urgentes, sobre todo en animales en sus entornos naturales y

empleando métodos experimentales, los estudios sobre gestos se focalizaron más bien en animales en cautiverio, en contextos socialmente relajados como el juego, estuvieron mayormente centrados en la conducta del señalizador, empleando métodos observacionales (Liebal *et al.* 2014).

En la siguiente sección, haré un breve repaso de algunas evidencias significativas recogidas por las investigaciones multimodales que ponen de manifiesto rasgos interesantes de la comunicación multimodal primate que el sesgo unimodal impedía advertir.

4.2. La comunicación multimodal primate: algunas evidencias

Los estudios observacionales y experimentales reflejan que diferentes especies primates usan fluida e intencionalmente ricos repertorios gestuales, vocales y faciales, tanto separadamente como en combinación, con sensibilidad al contexto social, y dependiendo tanto del estado atencional de sus destinatarios como del grado de éxito de las interacciones previas y de otros factores sociales. Se han identificado repertorios unimodales en chimpancés (Hobaiter y Byrne 2014), bonobos (Graham, Furichi y Byrne 2017) y otros primates, así como también repertorios de combinaciones multimodales de gestos y gruñidos u otras vocalizaciones en bonobos (Genty *et al.* 2014), en chimpancés (Wilke *et al.* 2017), de vocalizaciones y señales visuales en chimpancés (Mine *et al.* 2024), y de señales vocales y faciales en distintas especies primates (Ghazanfar 2013). En Hobaiter *et al.* (2017) puede encontrarse un amplio registro de los diferentes repertorios gestuales en distintas especies primates. Las evidencias del carácter intencional del uso de estas señales son abundantes: se dirigen a su audiencia, esperan una respuesta, persisten en obtenerla y reelaboran las señales (Hobaiter *et al.* 2017).

Un estudio comparó la frecuencia de uso de señales unimodales y multimodales por parte de chimpancés en escenarios naturales, en relación con diferentes contextos sociales y factores demográficos (sexo, rango y edad), constatándose que las señales multimodales (de las que se documentaron 48 combinaciones libres sobre un total de 62 señales de las modalidades gestual, vocal y facial) eran empleadas con relativa menor frecuencia, aunque en contextos y condiciones similares a las unimodales (descanso, alimentación y acicalamiento). Se constató, además, que las combinaciones más frecuentes, vocal-gestuales, producían una respuesta con mayor probabilidad que sólo una señal vocal, aunque no más que una gestual. Este hecho podría explicarse por la diferente función de ambas, atribuyéndose a las

vocales la función específica de captadoras de la atención (Wilke *et al.* 2017).

Un interesante estudio de Genty *et al.* (2014), acerca de un tipo particular de vocalización empleado por los bonobos, el grito de disputa (*contest hoot*), muestra que este se produce en combinación con diferentes gestos y señales faciales, así como movimientos de balanceo corporal más rudos o más suaves (*rough or soft*). Esa vocalización expresa la intención de desafiar o provocar una interacción social, dirigida por los machos hacia otros machos de estatus social relativamente alto, por lo que se conjetura que permiten al señalizador desplegar su propio estatus social. Ahora bien, el tipo de interacción que suscita, o bien agonística o bien afiliativa, está modulada por los gestos con los que se combina, por ejemplo, un pisotón (*a threatening "stomp"*) o una juguetona "sacudida de muñecas" (*a playful "wrist shake"*), respectivamente. Este es un caso donde las dos señales no tienen una función redundante, sino que un componente *modula* el significado del otro. Este tipo de acción comunicativa es posible en virtud de la combinación flexible y el empleo simultáneo de señales unimodales ya disponibles en el repertorio vocal y gestual de esta especie, y ejemplifica un caso de limitada capacidad composicional de las señales: mientras la vocalización intenta provocar una respuesta, diferentes gestos sirven para explicitar con cuál de dos intenciones antagónicas posibles se realiza esa "invitación".

En los primates, el empleo de señales faciales, gestuales y vocales combinadas entre sí, es muy frecuente en contextos de proximidad (*short-range communication*) (Fröhlich *et al.* 2019). Parece razonable suponer que las señales multimodales simultáneas pueden resultar más eficaces cuando la interacción comunicativa ocurre de manera suficientemente próxima de modo que el destinatario pueda ver, escuchar, oler y/o tocar al emisor, e incluso "puede facilitar el uso de señales graduadas más que discretas" (Partan y Marler 2005, p. 233). Este aspecto es importante para las interacciones comunicativas cara-a-cara entre individuos que mantienen vínculos habituales. Un estudio realizado en chimpancés identificó el papel que desempeña el relamerse de los labios (*lip-smacking*), que consiste en los chasquidos provocados por abrir y cerrar rápidamente la boca y los labios, una señal multimodal intencional con un componente visual y otro vocal, común en muchos primates. Este comportamiento es propio de las interacciones cara-a-cara, frecuente entre las madres y sus crías, pero también es usado para indicar una actitud afiliativa, como la disposición para iniciar y mantener el acicalamiento social, para coordinar y prolongar dicha tarea, para facilitar la reciprocidad

por parte del receptor, así como, más ampliamente, para intensificar los lazos sociales (Ghazanfar 2013; Fedurek *et al.* 2015).

En las madres chimpancés se ha descrito el uso habitual de la combinación de un tipo de vocalización dirigida hacia sus crías, “suave hoo” (*soft hoo*) combinada con un gesto de “brazo extendido” (*arm reach*) o “espalda presente” (*present back*), como invitación a subirse a su espalda. Estos intercambios tienen una estructura similar a la toma de turnos que organiza las conversaciones cara-a-cara en humanos (Fröhlich *et al.* 2016). Rossano (2013b) estudió la estructura secuencial y el *timing* que organiza ese mismo tipo de interacción comunicativa entre las madres bonobo y sus crías y que precede a la actividad de co-locomoción. Se trata de una secuencia de señales gestuales, táctiles o visuales, mediante una organización temporal por turnos que es, en lo fundamental, la misma que se ha identificado con las herramientas del análisis conversacional en las comunicaciones cara-a-cara entre los humanos. Estas formas de comunicación interactiva ponen de relieve que una comprensión adecuada de las señales requiere tomar en cuenta no sólo su forma, ni su producción o su recepción, sino también la estructura y la dinámica en la que desempeñan su función.

Un estudio realizado sobre chimpancés salvajes exploró los patrones de desarrollo ontogenético manifestados en la producción de señales unimodales y multimodales, con expresiones faciales, vocalizaciones y gestos (un repertorio de 66 señales en esas modalidades: 6 faciales, 8 vocales y 52 gestuales, 13 señales multimodales fijas y 316 combinaciones multimodales flexibles), en individuos pequeños y jóvenes (entre 1 y 11 años de edad), en distintos contextos (Doherty, Davila-Ross y Clay 2023). El patrón observado fue el siguiente: es más frecuente el uso de señales unimodales que multimodales, aunque estas se incrementan con la edad, en contextos de juego en los más pequeños y de agresión en los adolescentes, lo que podría explicarse porque una señal ambigua en tales contextos podría tener un mayor costo para los emisores. Se observó también que el uso de gestos es muy frecuente en los infantes, probablemente porque sus interacciones tienen lugar en mayor medida en estrecha proximidad con otros individuos, con sus hermanos y su madre, así como también en contextos de juego. Otro estudio sobre el uso de un variado repertorio de señales vocales en chimpancés en cautiverio reveló que aproximadamente la mitad se produjeron en combinación con otra modalidad, y que, en contraste con las señales unimodales, las multimodales con un componente vocal estaban dirigidas a un individuo

específico, con fines comunicativos inmediatos y en contextos de proximidad espacial, por lo que refuerza la interpretación que el uso de señales vocales también es intencional cuando la interacción es próxima (Tagliatela *et al.* 2015).

Otro estudio observacional en chimpancés salvajes se propuso identificar el repertorio de señales multimodales vocal-visuales, diferenciándolas de las combinaciones azarosas entre ellas (Mine *et al.* 2024). Respecto al componente vocal, se registraron las 7 vocalizaciones o “llamadas” más frecuentes de las 13 previamente descritas en la literatura: gruñido, ulular suave, ladrado jadeante, gruñido jadeante, ulular jadeante, grito y gemido (*grunt*, *soft hoo*, *pant bark*, *pant grunt*, *pant hoot*, *scream*, *whimper*) y una variedad de 31 combinaciones diferentes con variados comportamientos visuales (posturas y movimientos corporales, gestos y expresiones faciales). Ello permitió identificar y registrar el primer repertorio de señales vocal-visuales en chimpancés: 108 señales multimodales. Aunque algunos de esos comportamientos pueden ser claves antes que señales (p. ej., la postura corporal, la orientación de la cabeza, la dirección de la mirada, movimientos de las manos, los brazos, etc.), por lo cual contribuyen a formar señales multimodales fijas, ofrecen a los receptores información relevante para comprender las señales. El estudio también mostró que no todas las vocalizaciones son usadas en la misma proporción con señales o claves visuales.

El otro hallazgo interesante del mismo estudio es que una proporción de 90 % de los comportamientos visuales se combinan con vocalizaciones diferentes. Los autores interpretan que el gruñido (*grunt*), por ejemplo, usado en diferentes contextos, es desambiguado al combinarse con otras señales. Estos podrían ser casos donde la combinación es composicional porque esta sólo la modula o refina. Otro ejemplo de un tipo de combinación lo proporciona el uso de la señal *play face*, una expresión relajada con la boca abierta, muy frecuente en los primates, a la que se atribuye una función metacomunicativa en tanto altera el significado de la acción comunicativa principal a la que precede o acompaña: en contextos de juego, anticipa la intención no agresiva de las sucesivas conductas comunicativas (Amici, Oña y Liebal 2024).

También refleja una estructura comunicativa compleja el uso de vocalizaciones y de gestos visuales o táctiles para captar la atención del destinatario, justamente cuando este no está mirando al emisor (Tomasello 2008). Así, su función no es expresar un significado en particular, sino que el destinatario preste atención a las

próximas señales, que sí lo hacen. Su función sería indicativa de la presencia de un tipo de comunicación intencional “mínimamente griecano” (Moore 2017a). Este comportamiento se ha descrito también en orangutanes y chimpancés que interactúan con humanos cuando estos no los están mirando, por ejemplo, para poder solicitarles alimento. Los chimpancés en cautiverio usan gestos y expresiones faciales combinadas, pero sólo cuando el experimentador los mira (Hostetter, Cantero y Hopkins 2001). Este tipo de comportamientos ha sido identificado en numerosos experimentos, y es considerado por algunos especialistas un indicador de comportamiento ostensivo del tipo requerido para la comunicación intencional. Cartmill y Byrne (2007) constataron que los orangutanes no sólo persisten, sino que reelaboran sus señales, e incluso lo hacen discriminando entre diferentes grados de éxito comunicativo, persistiendo con la misma señal o una similar si el resultado es parcialmente alcanzado, mientras que, si la respuesta es insatisfactoria, adoptan la táctica de la elaboración, recurriendo a diferentes tipos de señales.

Un estudio realizado en chimpancés cautivos (*Pan troglodytes*), en un grupo integrado por 110 individuos (Leavens, Russell y Hopkins 2010) examinó cómo integran múltiples modalidades en la comunicación interespecífica, poniendo de manifiesto que tanto *persisten* como *elaboran* la señal comunicativa en una modalidad diferente, cuando la primera falla. En el estudio mencionado se identificaron distintos tipos de señales comunicativas, gestuales y vocales, además de los comportamientos visuales, auditivos y táctiles de “llamar la atención” (*attention-getting behaviour*) en cada modalidad. Se estudió la frecuencia y el grado de elaboración de estas conductas como efecto de la orientación del receptor humano. Los resultados mostraron que los chimpancés producían más comportamientos vocales para llamar la atención cuando un humano no estaba orientado hacia ellos. En caso contrario, producían más gestos visuales e incluso elaboraban la señal en cada modalidad prestando atención a la orientación y la recepción del observador humano. De esta manera, se pudo constatar que “siguen un patrón lógico y eficiente de permutaciones específicas a la modalidad” (p. 38). El estudio también puso de manifiesto que las condiciones empobrecidas de cautividad ante tareas como obtener alimento o escoger el alimento preferido, obligan a los chimpancés a desarrollar soluciones comunicativas eficientes.

Una de las conclusiones más interesantes que pueden extraerse de estos resultados es que los primates parecen entender o, al menos, “tener el sentimiento de entender” (Planer 2023b) la función comunicativa de su conducta, porque persisten, modulan o cambian

su comportamiento comunicativo dependiendo de la respuesta del receptor, es decir, inhiben una señal en una modalidad dependiendo de la ausencia o inaccesibilidad del destinatario y cambian tácticamente su modalidad primaria dependiendo de su orientación física, por ejemplo, de la mirada del experimentador humano o de otro coespecífico (Tomasello *et al.*, 1994; Hostetter *et al.* 2001; Cartmill y Byrne 2007; Leavens 2007).

Así, las investigaciones sobre las señales multimodales en distintas especies primates alientan la adopción de un enfoque pragmático e intencional puesto que incorpora el papel tanto de emisores como receptores y de los contextos sociales y ambientales en los que ocurren las interacciones. En contextos de proximidad, el enfoque pragmático-interactivo puede dar cuenta de las funciones de señales y claves en diferentes modalidades que concurren en la realización de variados tipos de acciones comunicativas, distinguiendo entre aquellas en las que, mediante una división del trabajo comunicativo, unas transmiten el contenido principal y otras lo desambiguan o refinan, o bien, unas están dedicadas a asegurar su eficacia, es decir, tienen una función metacomunicativa, mientras que otras sirven para realizar la acción comunicativa.

Sin embargo, las evidencias disponibles son todavía insuficientes en relación con todos los parámetros relevantes, por ejemplo, cuáles son los repertorios comunicativos de las diferentes especies primates, cuáles son las habilidades cognitivas en diferentes etapas del desarrollo entre animales jóvenes y adultos y en qué contextos se emplean señales de uno u otro tipo con mayor frecuencia, en qué grado se registran variaciones en el uso de señales multimodales entre los individuos según diferentes factores sociales (edad, sexo, relaciones de parentesco, afinidad y rango entre los interlocutores) y en diferentes contextos sociales, cómo influyen las diferentes condiciones comunicativas que imponen los escenarios naturales o la vida en cautiverio, entre otras. En cualquier caso, la imagen que comienza a emerger de las investigaciones multimodales es más compleja que la edificada por el enfoque unimodal y ha comenzado a modificar la visión recibida. En particular, parece claro que el empleo combinado de señales de diferentes tipos, innatas y aprendidas, espontáneas y deliberadas, en diferentes combinaciones, más o menos flexibles, podría ser más frecuente que sus variedades “puras”, así como también variable según la situación comunicativa, las condiciones del contexto y otros factores relevantes, como el tipo de relación social que vincula a los comunicadores. De ese modo, el examen del papel que pueden desempeñar las diferentes modalidades en la comunicación primate,

en todas las dimensiones señaladas, y en particular su uso combinado e intencional, devuelve una imagen más enriquecida, dinámica y compleja que la ofrecida por los enfoques influidos por el sesgo unimodal.

5. *Algunas reflexiones finales*

Emplear uno u otro modelo del lenguaje y la comunicación humana para estudiar la comunicación animal, poniendo la atención en sus rasgos diferenciales, lleva implícito el efecto de asimilar injustificadamente todas las formas de comunicación animal entre sí. Desde una perspectiva evolutiva, sin embargo, asimilar el comportamiento comunicativo de las moscas, las ranas, los patos, los reptiles y los primates no humanos no está más justificado que explorar semejanzas y diferencias entre especies emparentadas y, por lo tanto, también entre los primates humanos y no humanos.

El intento de dar cuenta de los rasgos estructurales de las señales animales y de los usos comunicativos que se han considerado más fundamentales, transferidos de los enfoques elaborados para dar cuenta de las estructuras lingüísticas y de formas idealizadas de comunicación humana, no permiten comprender adecuadamente las peculiaridades de la comunicación primate. En particular, las investigaciones predominantemente unimodales de las señales comunicativas primates, una expresión del *sesgo linguacéntrico*, ofrecieron una imagen inadecuada de sus ricos repertorios, usos flexibles y variadas funciones en la dinámica interactiva. Así, reconocer el carácter multimodal de la comunicación primate abre las compuertas para elaborar nuevos enfoques acerca de otros rasgos importantes que dan su perfil distintivo a los sistemas comunicativos de los primates, es decir, alienta la revisión de otros componentes del *sesgo linguacéntrico*, v. gr. el carácter cultural y aprendido de las señales, sus propiedades combinatorias, diferentes tipos de acciones comunicativas, etcétera.

He puesto el foco en el carácter multimodal de la comunicación intencional en los primates no humanos, porque se trata de un rasgo presente en ciertas formas distintivas de comunicación humana, pero la comunicación biológica también es multimodal en casi todas las especies. En otras palabras, el carácter multimodal de la comunicación en muchas especies, incluido el *Homo Sapiens*, está notoriamente extendido. Sin embargo, los estudios multimodales, frecuentes en el caso de otras especies tuvo que esperar mucho más tiempo en el caso de los primates. Que no se haya advertido su importancia como una

guía para comprender sus rasgos peculiares, requiere una explicación. Es la que he intentado ofrecer aquí.

El foco crítico puesto en el *sesgo lingüacéntrico* contribuye también al campo de estudios sobre la evolución del lenguaje; si el modo por defecto de la comunicación primate es multimodal, habremos dado con uno de los puentes más básicos que podría explicar la evolución de la comunicación lingüística humana, dado que esta también puede ser caracterizada por defecto como una forma de comunicación multimodal. En otras palabras, explicar la singularidad del lenguaje humano no exige renunciar a una explicación evolutiva basada en capacidades comunicativas y cognitivo-sociales de ciertos tipos más básicos, presentes tanto en nuestros parientes más próximos como en las formas de comunicación más extendidas entre nosotros.³⁰

BIBLIOGRAFÍA

- Amici, Federica, Linda Oña y Katja Liebal, 2024, “Compositionality in Primate Gestural Communication and Multicomponent Signal Displays”, *International Journal of Primatology*, vol. 45, pp. 482–500.
- Andrews, Kristin, 2020, *The Animal Mind. An Introduction to Animal Cognition*, Routledge, Londres y Nueva York.
- Aristotle, 1970, *Historia Animalium*, vol. II, Books IV–VI, Cambridge University Press, Cambridge.
- Bar-On, Dorit, 2025, “‘Pragmatics-First’: Animal Communication and the Evolution of Language”, *Review of Philosophy and Psychology*, vol. 16, pp. 1–28.
- Bar-On, Dorit, 2021, “How To Do Things with Nonwords: Pragmatics, Biosemantics, and Origins of Language in Animal Communication”, *Biology & Philosophy*, vol. 36, no. 6, pp. 1–25.
- Byrne, Richard W., Erica Cartmill, Emilie Genty, Kirsty E. Graham, Catherine Hobaiter y Jessica Tanner, 2017, “Great Ape Gestures: Intentional Communication with a Rich Set of Innate Signals”, *Animal Cognition*, vol. 20, pp. 755–769.
- Cartmill, Erica, 2023, “Overcoming Bias in The Comparison of Human Language and Animal Communication”, *Proceedings of the Royal Academy of Sciences*, 230, 47, e2218799120.
- Cartmill, Erica A. y Richard W. Byrne, 2007, “Orangutans Modify Their Gestural Signaling according to Their Audience’s Comprehension”, *Current Biology*, vol. 17, no. 15, pp. 1345–1348.

³⁰ Deseo agradecer los valiosos comentarios y sugerencias realizados por los evaluadores de este trabajo porque me permitieron mejorar la presentación de sus contenidos. La investigación realizada fue posible gracias al apoyo de los subsidios PIP CONICET y CONSOLIDAR SECYT-UNC: “Mentes intencionales: variedades de contenido, expresión y atribución”.

- Crockford, Catherine, Roman M. Wittig, Roger Mundry y Klaus Zuberbühler, 2012, “Wild Chimpanzees Inform Ignorant Group Members of Danger”, *Current biology*, vol. 22, no. 2, pp. 142–146.
- Crockford, Catherine, Ilka Herbinger, Linda Vigilant y Christophe Boesch, 2004, “Wild Wild Chimpanzees Produce Group-Specific Calls: A Case For Vocal Learning?”, *Ethology*, vol. 10, no. 3, pp. 221–243.
- Devitt, Michael y Kim Sterelny, 1999, *Language and Reality: An Introduction to the Philosophy of Language*, Wiley-Blackwell, Cambridge.
- Doherty, Emma, Marina Davila-Ross y Zanna Clay, 2023, “Multimodal Development Communication in Semiwild Chimpanzees”, *Animal Behavior*, vol. 201, pp. 175–190.
- Fedurek, Pawel y Katie E. Slocombe, 2011, “Primate Vocal Communication: A Useful Tool for Understanding Human Speech and Language Evolution?”, *Human Biology*, vol. 83, no. 2, pp. 153–173.
- Fedurek, Pawel, Katie E. Slocombe, Jessica Hartel, y Karl Zuberbühler, 2015, “Chimpanzee Lip-Smacking Facilitates Cooperative Behavior”, *Scientific Reports*, 5, 13460.
- Filippi, Piera, 2016, “Emotional and Interactional Prosody across Animal Communication Systems: A Comparative Approach to the Emergence of Language”, *Frontiers in Psychology*, vol. 7, 1393.
- Fröhlich, Marlen, 2017, “Taking Turns across Channels: Conversation-Analytic Tools in Animal Communication”, *Neuroscience Biobehavioral Reviews*, vol. 80, pp. 201–209.
- Fröhlich, Marlen y Carel P. van Schaik, 2020, “Must All Signals Be Evolved? A Proposal for a New Classification of Communicative Acts”, *Wiley Interdisciplinary Reviews: Cognitive Science*, vol. 11, no. 4, e1527.
- Fröhlich, Marlen y Carel P. van Schaik, 2018, “The Function of Animal Multimodal Communication”, *Animal Cognition*, vol. 21, no. 5, pp. 619–629.
- Fröhlich, Marlen, Christine Sievers, Simon W. Townsend, Thibaud Gruber y Carel P. van Schaik, 2019, “Multimodal Communication and Language Origins: Integrating Gestures and Vocalizations”, *Biological Reviews*, vol. 94, no. 5, pp. 1809–1829.
- Fröhlich, Marlen, Paul Kuchenbuch, Gudrun Müller, Barbara Fruth, Takeshi Furuichi, Roman M. Wittig y Simone Pika, 2016, “Unpeeling The Layers of Language: Bonobos and Chimpanzees Engage in Cooperative Turn-Taking Sequences”, *Scientific Reports*, vol. 6, no. 1, pp. 1–14.
- Genty, Emilie, Zanna Clay, Catherine Hobaiter y Klaus Zuberbühler, 2014, “Multi-Modal Use of Socially Directed Call in Bonobos”, *PlosOne*, vol. 9, no. 1, e84738.
- Ghazanfar, Asif A., 2013, “Multisensory Vocal Communication in Primates and the Evolution of Rhythmic Speech”, *Behavioral Ecology and Sociobiology*, vol. 67, no. 9, pp. 1441–1448.
- Graham, Kirsty E., Federico Rossano y Richard Moore, 2025, “The Origins of Great Ape Gestural Forms”, *Biological Reviews*, vol. 100, no. 1, pp. 190–204.

- Graham, Kirsty E., Takeshi Furuichi y Richard W. Byrne, 2017, "The Gestural Repertoire of the Wild Bonobo (*Pan Paniscus*): A Mutually Understood Communication System", *Animal Cognition*, vol. 20, no. 2, pp. 171–177.
- Glazer, Tim, 2019, "The Social Amplification View of Facial Expression", *Biology & Philosophy*, vol. 34, no.33.
- Grice, Herbert P., 1975, "Logic and Conversation", en Peter Cole y Jerry L. Morgan (eds.), *Syntax and Semantics*, vol. 3: Speech Acts, Academic Press, Nueva York, pp. 33–48.
- Grice, Herbert P., 1977, *Significado*, Cuadernos de Crítica, Instituto de Investigaciones Filosóficas-UNAM, México.
- Hauser, Marc D., Noam Chomsky y W. Tecumesh, Fitch, 2002, "The Faculty of Language: What Is It, Who Has It, and How Did It Evolve?", *Science*, vol. 298, no. 5598, pp. 1569–1579.
- Hebets, Eilen A., 2011, "Currents Status and Future Directions of Research in Complex Signaling", *Current Zoology*, vol. 57, no. 2, pp. i–v.
- Hebets, Eilen A. y Daniel R. Papaj, 2005, "Complex Signal Function: Developing a Framework of Testable Hypotheses", *Behavioral Ecology and Sociobiology*, vol. 57, no. 3, pp. 197–214.
- Higham, James P. y Eilen A. Hebets, 2013, "An Introduction to Multimodal Communication", *Behavioral Ecology of Sociobiology*, vol. 67, pp. 1381–1388.
- Hobaiter, Catherine, Richard W. Byrne y Klaus Zuberbühler, 2017, "Wild Chimpanzees' Use of Single and Combined Vocal and Gestural Signals", *Behavioral Ecology and Sociobiology*, vol. 71, 96.
- Hobaiter, Catherine y Richard W. Byrne, 2014, "The Meanings of Chimpanzee Gestures", *Current Biology*, vol. 24, no. 14, pp. 1596–1600.
- Hockett, Charles F., 1960, "The Origin of Speech", *Scientific American*, vol. 203, no. 3, pp. 88–111.
- Hostetter, Autumn B., Monica Cantero y William D. Hopkins, 2001, "Differential Use of Vocal and Gestural Communication by Chimpanzees (*Pan Troglodytes*) in Response to the Attentional Status of A Human (*Homo Sapiens*)", *Journal of Comparative Psychology*, vol. 115, no. 4, pp. 337–343.
- Keiser, Jessica, 2023, *Non-Ideal Foundations of Language*, Routledge, Nueva York.
- Leavens, David, Jamie L. Russell, y William D. Hopkins, 2010, "Multimodal Communication by Captive Chimpanzees (*Pan Troglodytes*)", *Animal Cognition*, vol. 13, no. 1, pp. 33–40.
- Leavens, David, 2007, "Animal Cognition: Multimodal Tactics of Orangutan Communication", *Current Biology*, vol. 17, no. 17, R762.
- Levinson, Stephen C., 2022, "The Evolution of Interactional Engine: Cutness Selection and the Evolution of Interactional Base of Language", *Philosophical Transactions of the Royal Society B*, vol. 377, no. 1859, 20210108.

- Levinson, Stephen C., 2019, “Interactional Foundations of Language: the Interaction Engine Hypothesis”, en Peter Hagoort (ed.), *Human Language: from Genes and Brain to Behavior*, MIT Press, Cambridge, Mass., pp. 189–200.
- Levinson, Stephen C., 2016, “Turn-Taking in Human Communication: Origins and Implications for Language Processing”, *Trends in Cognitive Science*, vol. 20, pp. 6–14.
- Levinson, Stephen C. y Judith Holler, 2014, “The Origin of Human Multi-Modal Communication”, *Philosophical Transactions of the Royal Society B*, vol. 369, no. 1651, 2013030.
- Levinson, Stephen C., 2006, “On the Human ‘Interaction Engine’”, en Nicholas J. Enfield, Stephen C. Levinson (eds.), *Roots of Human Sociality: Culture, Cognition and Interaction*, Berg, Oxford, Reino Unido, pp. 39–69.
- Liebal, Katja, Katie E. Slocombe y Bridget M. Waller, 2022, “The Language Void 10 Years On: Multimodal Primate Communication Research is Still Uncommon”, *Ethology, Ecology & Evolution*, vol. 34, no. 22, pp. 274–287.
- Liebal, Katja, Bridget M. Waller, Anne M. Burrows y Katie E. Slocombe, 2014, *Primate Communication: A Multimodal Approach*, Cambridge University Press, Cambridge.
- Linell, Per, 2005, *The Written Language Bias: Its Nature, Origins, and Transformations*, Routledge, Londres y Nueva York.
- Maynard-Smith, John y David Harper, 2003, *Animal Signals*, Oxford University Press, Nueva York.
- Mine, Joseph G., Claudia Wilke, Chiara Zulberti, Chiara, Melika Behjati, Alexandra Bosshard, Sabine Stoll, Zarin P. Machanda, Andri Manser, Katie E. Slocombe, Simon Townsend y W. Simon, 2024, “Vocal-Visual Combinations in Wild Chimpanzees”, *Behavioral Ecology and Sociobiology*, vol. 78, no. 10, 108.
- Moore, Richard, 2017a, “Gricean Communication and Cognitive Development”, *Philosophical Quarterly*, vol. 67, no. 267, pp. 303–326.
- Moore, Richard, 2017b, “Pragmatics’s First Approaches to the Evolution of Language”, *Psychological Enquiry. An International Journal for the Advancement of Psychological Theory*, vol. 28, no. 2–3, pp. 206–210.
- Moore, Richard, 2016, “Meaning and Ostension in Great Ape Gestural Communication”, *Animal Cognition*, vol. 19, no. 1, pp. 223–231.
- Moore, Richard y Bart Geurts, 2025, “Introduction”, en Bart Geurts y Richard Moore (eds.), *Evolutionary Pragmatics: Communicative Interaction and the Origins of Language*, Oxford University Press, Oxford, pp. xiii–xxii.
- Owren, Michael J., Drew Rendall y Michael J. Ryan, 2010, “Redefining Animal Signaling: Influence vs. Information in Communication”, *Biology & Philosophy*, vol. 25, pp. 755–780.
- Palazzolo, Giulia, 2025, “What is Animal Communication?”, *Ergo an Open Access Journal of Philosophy*.

- Partan, Sarah R. y Peter Marler, 2005, "Issues in the Classification of Multisensory Communication Signals", *American Naturalist*, vol. 166, no. 2, pp. 231–245.
- Partan, Sarah R. y Peter Marler, 1999, "Communication Goes Multimodal", *Science*, vol. 283, no. 5406, pp. 1272–1273.
- Perniss, Pamela, 2018, "Why We Should Study Multimodal Language", *Frontiers in Psychology*, vol. 9, 1109.
- Pika, Simone y Marlen Fröhlich, 2019, "Gestural Acquisition in Great Apes: The Social Negotiation Hypothesis", *Animal Cognition*, vol. 22, no. 4, pp. 551–565.
- Pika, Simone, Ray Wilkinson, Kobin H. Kendrick y Sonja C. Vernes, 2018, "Taking Turns: Bridging the Gap between Human and Animal Communication", *Proceedings of the Royal Society B*, vol. 285, 1880, 20180598.
- Planer, Ronald J., 2023a, "Going Dennettian about Gricean Communication", *Philosophical Psychology*, vol. 38, no. 4, pp. 1271–1294.
- Planer, Ronald J., 2023b, "Conversation and the Evolution of Metacognition", *Evolutionary Linguistic Theory*, vol. 5, no. 1, pp. 53–78.
- Rendall, Drew, Michael J. Owren y Michael J. Ryan, 2009, "What Do Animal Signals Mean?", *Animal Behavior*, vol. 78, no. 2, pp. 233–240.
- Rossano, Federico y Stephan Kaufhold, 2021, "Animal Communication Review", en Allison B. Kaufman, Josep Call y James C. Kaufman (eds.), *The Cambridge Handbook of Animal Cognition*, Oxford University Press, Reino Unido, pp. 5–35.
- Rossano, Federico y Katja Liebal, 2014, "Requests" and "Offers" in Orangutans and Human Infants", en Paul Drew y Elizabeth y Couper Kuhlen (eds.), *Requesting in Social Interaction*, vol. 26, John Benjamins, Amsterdam, pp. 335–364.
- Rossano, Federico, 2013a, "Gaze in Conversation", en Jack Sidnell y Tanya Stivers (eds.), *The Handbook of Conversational Analysis*, Wiley-Blackwell, Reino Unido, pp. 308–329.
- Rossano, Federico, 2013b, "Sequence Organization and Timing of Bonobo Mother-Infant Interactions", *Interaction Studies*, vol. 14, no. 2, pp. 160–189.
- Rowe, Candy, 1999, "Receiver Psychology and the Evolution of Multicomponent Signals", *Animal Behaviour*, vol. 58, pp. 921–931.
- Schel, Anne M., Simon W. Townsend, Zarin Machanda, Klaus Zuberbühler y Katie E. Slocombe, 2013, "Chimpanzee Alarm Call Production Meets Key Criteria for Intentionality", *PLoS ONE*, vol. 8, no. 10, e76674.
- Scotto, Carolina, 2025, "Cognición social de segunda persona en la comunicación cara-a-cara", en Ángeles Eraña y Santiago Echeverri (comps.), *Yo, tú, nosotras (lenguaje, cognición y creencia)*, Bonilla/Instituto de Investigaciones Filosóficas-UNAM, México, pp. 105–141.
- Scotto, Carolina, 2024, "The Anthropocentric Bias in Animal Cognition", *ArtefaCTos*, vol. 13, no. 1, pp. 85–116.

- Scotto, Carolina, 2022, “A Pragmatics-First Approach to Faces”, *Topoi*, vol. 41, pp. 641–657.
- Scott-Phillips, Thom C., 2015, *Non-Human Primate Communication, Pragmatics, and the Origins of Language*, *Current Anthropology*, vol. 56, no. 1, pp. 56–80.
- Seyfarth, Robert y Dorothy Cheney, 2018, “Pragmatic Flexibility in Primate Vocal Production”, *Current Opinions in Behavioral Sciences*, vol. 21, pp. 56–61.
- Seyfarth, Robert y Dorothy Cheney, 2017, “The Origin of Meaning in Animal Signals”, *Animal Behavior*, vol. 124, pp. 339–346.
- Seyfarth, Robert M., Dorothy L. Cheney y Peter Marler, 1980, “Monkey Responses to Three Different Alarm Calls: Evidence of Predator Classification and Semantic Communication”, *Science*, vol. 210, no. 4471, pp. 801–803.
- Shanker, Stuart G. y Barbara J. King, 2002, “The Emergence of a New Paradigm in Ape Language Research: Beyond Interactionism”, *Behavioral and Brain Sciences*, vol. 25, no. 5, pp. 646–651.
- Slocombe, Katie E., Bridget M. Waller y Katja Liebal, 2011, “The Void of Language: the Need for Multimodality in Primate Communication Research”, *Animal Behaviour*, vol. 81, no. 5, pp. 919–924.
- Spence, Charles, 2011, “Crossmodal Correspondences: A Tutorial Review”, *Attention, Perception, and Psychophysics*, vol. 73, pp. 971–995.
- Taglialatela, Jared P., Jamie L. Russell, Sara H. Pope, Tamara Morton, Stephanie Bogart, Lisa A. Reamer, Stephen J. Schapiro y William D. Hopkins, 2015, “Multimodal Communication in Chimpanzees”, *American Journal of Primatology*, vol. 77, no. 11, pp. 1143–1148.
- Tinbergen, Nikolaas, 1952, “‘Derived’ Activities; Their Causation, Biological Significance, Origin, and Emancipation during Evolution”, *The Quarterly Review Biology*, vol. 27, no. 1, pp. 1–32.
- Tomasello, Michael, Joseph Call, Katherine Nagel, Raquel Olguin y Malinda Carpenter, 1994, “The Learning and Use of Gestural Signals by Young Chimpanzees: A Transgenerational Study”, *Primates*, vol. 35, pp. 137–154.
- Tomasello, Michael, 2008, *Los orígenes de la comunicación humana*, Katz Ed. Buenos Aires, 2013.
- Townsend, Simon W. et al., 2017, “Exorcising Grice’s Ghost: An Empirical Approach to Studying Intentional Communication in Animals”, *Biological Review*, vol. 92, no. 3, pp. 1427–1433.
- Townsend, Simon W., Stuart K. Watson y Katie Slocombe, 2020, “Flexibility in Great Ape Vocal Production”, en Lydia M. Hopper, Stephen R. Ross (eds.), *Chimpanzees in Context: A Comparative Perspective on Chimpanzee Behavior, Cognition, Conservation, and Welfare*, University of Chicago Press, Chicago y Londres, pp. 260–280.
- Ullrich, Robert, Moritz Mittelbach y Katja Liebal, 2018, “‘Scala Naturae’: the Impact of Historical Values on Current ‘Evolution of Language’ Discourse”, *Journal of Language Evolution*, vol. 3, no. 1, pp. 1–12.

- Vigliocco, Gabriella, Pamela Perniss y David Vinson, 2014, “Language as a Multimodal Phenomenon: Implications for Language Learning, Processing and Evolution”, *Philosophical Transactions of the Royal Society B*, vol. 369, 20130292.
- Wheeler, Brandon C. y Julia Fischer, 2012, “Functionally Referential Signals: A Promising Paradigm Whose Time Has Passed”, *Evolutionary Anthropology: Issues, News, and Reviews*, vol. 21, no. 5, pp. 195–205.
- Wilke, Claudia, Eithne Kavanagh, Ed Donnellan, Bridget M. Waller, Zarin P. Machanda y Katie E. Slocombe, 2017, “Production of and Responses to Unimodal and Multimodal Signals in Wild Chimpanzees, *Pan Troglodytes Schweinfurthii*”, *Animal Behaviour*, vol. 123, pp. 305–316.

Recibido el 20 de agosto de 2025; aceptado el 12 de enero de 2026.