

## Debatiendo sobre atención conjunta con una fenomenóloga

Anderson Pinzón-Duarte

La gente pasa y pasa siempre tan igual  
el ritmo de la vida me parece mal.

MARCO ANTONIO SOLÍS, *Si no te hubieras ido*

### Tipos de atención conjunta y el origen del desacuerdo

LA «ATENCIÓN CONJUNTA» ES el nombre que recibe una variedad heterogénea de episodios de interacción social/coordinación. El interés que estos episodios genera a filósofos y filósofas afines a la filosofía analítica y a la fenomenología motiva discusiones entre estos enfoques sobre cuál es la mejor manera de explicar la intersubjetividad y, en particular, cuál es la mejor manera de describir el entendimiento que soporta y explica el éxito de dichos episodios de interacción social/colaboración (cf. Gallagher, 2011; Eilan, 2005; Krueger, 2013; De Jaegher, 2006; Tomasello 2003, 2014). La articulación de estas discusiones se enfrenta a una dificultad: los ejemplos usados provienen de los estudios sobre la interacción social/coordinación realizados en psicología del desarrollo, psicología comparada o en etnometodología (en particular, en sociología), y no es claro si cada una de estas disciplinas habla acerca de lo mismo.

Sin el ánimo de ser exhaustivo y excluyente, es posible establecer diferencias relevantes entre estas disciplinas relativas a este punto. En psicología comparada, por ejemplo, se han estudiado episodios de interacción social/coordinación *humano-primate* o *primate-primate* (cf. Call and Tomasello, 2005; Soproni *et al.*, 2002; Gómez, 2005;

Leavens, 2011), mientras que la psicología del desarrollo humano ha estudiado episodios de interacción social/coordinación *adulto-niño* o *niño-niño* (entre los seis y dieciocho meses de edad) (cf. Tomasello *et al.*, 1998, Hobson y Hobson, 2008). Si bien estos episodios se parecen en cuanto su éxito depende principalmente del ejercicio de habilidades perceptuales, se distinguen porque la presencia de la conducta de señalar con el dedo índice ‘*pointing*’ —la cual es manifiesta en el caso humano y ausente en el caso primate— determina si son episodios de la misma clase (cf. Tomasello, 2003; Leavens, 2011; Eilan, 2005). Por su parte, al otro extremo del espectro, en etnometodología se estudia la interacción social humana mediada por el lenguaje articulado, el conocimiento ya sea implícito o explícito de reglas, y la participación en grupos o instituciones (cf. Kendon, 1990). Esta clase de episodios suele estar clasificada como un tipo radicalmente distinto de interacción social/coordinación en virtud del hecho de que su éxito requiere de alguna clase de entendimiento de reglas (lingüísticas, sociales o institucionales) y, en general, requiere del entendimiento de estructuras institucionales que, difícilmente, luce como algo reducible a un entendimiento perceptual.

Estos tres grupos de episodios de interacción social/coordinación se ven diferentes entre sí, pues aparentemente cada uno instancia tipos de interacción social/coordinación fundamentalmente diferentes entre sí. Se dice que —para coordinar exitosamente— a los primates les basta entender que los objetos y sus congéneres sirven como medios para cumplir la tarea que tienen en frente. Esto significa que ellos no necesitan entender a su congénere como un agente intencional o no necesitan entender que ellos y sus congéneres están realizando la misma tarea (cf. Dunham and Dunham, 1995). En el caso de los niños, por ejemplo, se dice que los episodios de interacción social exitosa dependen, en su mayoría, de que ellos entiendan al otro sujeto (bien sea otro niño o un adulto) como un agente intencional y, al mismo tiempo, de que entiendan que ambos están realizando la misma tarea (cf. Eilan, 2005; Tomasello *et al.*, 1998)<sup>1</sup>. En el caso de los

---

<sup>1</sup> Este último entendimiento es compartido. Y debido al aparente hecho de que los niños (entre seis y dieciocho meses) involucrados en estos episodios carecen

episodios de interacción social estudiados en etnometodología, se dice que su éxito radica en que cada sujeto debe considerar al otro como un agente intencional y darle sentido a la idea de que ambos están compartiendo entendimiento que es —principalmente— acerca de las reglas gramaticales o de las convenciones sociales/institucionales que constituyen la tarea misma. Por lo anterior, es difícil pensar que no hay habilidades cognitivas-conceptuales involucradas constitutivamente en el éxito de los siguientes episodios de interacción social/coordinación: una conversación entre dos personas (*cf.* Kendon, 1990), la organización de una cita con otra persona para el día siguiente (*cf.* De Jaegher, 2008) o, incluso, un episodio de improvisación musical de jazz (*cf.* Krueger, 2014).

Estas consideraciones, respecto al tipo de *entendimiento que explica y soporta el éxito de los episodios de interacción social*, son un buen fundamento para dudar de que quienes hablan de atención conjunta y estudian el fenómeno de la interacción social/coordinación estén realmente hablando acerca de lo mismo. Si el tipo de entendimiento que soporta y explica cada tipo de episodio es diferente, entonces las condiciones de éxito de cada uno de ellos serán distintas y, por lo tanto, cada tipo de episodio de interacción social será distinto.

Si estos tipos de interacción social son diferentes entre sí, surgen las siguientes preguntas: ¿cuál de éstos es el episodio de interacción social/coordinación más apropiado para entender el concepto de atención conjunta? o ¿cuál de estos episodios es el caso paradigmático de atención conjunta?<sup>2</sup>

En este punto la voz de Michael Tomasello (2003, 2014) es muy relevante. Tomasello (2014) estaría de acuerdo con la idea de que no todos estos ejemplos son episodios de atención conjunta. Su razón es que no todos estos episodios se tienen que explicar a partir del entendimiento (o estructura intencional) propio de los episodios de

---

de habilidades lingüísticas o conceptuales, entonces dicho entendimiento debe ser especificado independientemente del ejercicio de habilidades conceptuales (Eilan, 2005; Hobson and Hobson, 2008).

<sup>2</sup> Por ahora pienso que no es claro cuál es el tipo de episodio de interacción/coordinación paradigmático de la atención conjunta y, por eso mismo, no es claro siquiera si la atención conjunta existe como un fenómeno genuino.

atención conjunta. En sus términos, los episodios de atención conjunta son episodios de interacción social/coordinación basada en *conductas comunicativas*, tal como lo es «señalar con el dedo índice». Esto significa que el entendimiento detrás de la atención conjunta debe hacer posible la comunicación.

Tomasello (2003, pp. 94-133)<sup>3</sup> se suma a la explicación de la comunicación inspirada fuertemente en Grice (1957), según la cual la comunicación es, en parte, resultado de la articulación de actos comunicativos individuales, y cada acto comunicativo se explica en términos de *intenciones comunicativas* individuales que producen dichos actos. Así, comunicarse con alguien implica atribuirle intenciones comunicativas. Por esta razón, el entendimiento que soporta y explica el éxito de esta clase de episodios debe permitir la atribución de esta clase de estados mentales. Así, cuando una niña preverbal participa en un episodio de atención conjunta

requiere que la niña entienda los diferentes roles que el hablante y el oyente desempeñan en la actividad de atención conjunta, así como la intención *comunicativa específica del adulto* dentro de esa actividad— y luego ella debe poder expresar a otras personas la misma intención comunicativa que se expresaba previamente hacia ella.<sup>4</sup> (Tomasello, 2003, p. 95, énfasis agregado)

La intención comunicativa atribuida tiene la siguiente estructura: «tú intentas que (yo comparta atención a X)» (Tomasello, 2003, p. 102). Y esto implica un tipo de estructura intencional compleja. Escribe Tomasello:

Este análisis deja bastante claro que la comprensión de una intención comunicativa es un caso especial de la comprensión de una intención; es entender la intención de otra persona *hacia mi estado de atención*. Comprender esto es claramente más complejo que en-

---

3 Todas las citas de este artículo que originalmente están en inglés son traducción propia.

4 Incluso, más adelante afirma: «Para ver estos ruidos y movimientos de las manos como algo con un significado comunicativo que puede aprenderse y usarse, el niño debe comprender que está motivado por un tipo especial de intención, a saber, una intención comunicativa». (Tomasello 2003, p. 96)

tender la intención de otra persona, *simpliciter*. Para entender que la intención de otra persona es patear una pelota, simplemente debo determinar su objetivo con respecto a la pelota. Pero *para entender lo que otra persona pretende cuando hace el sonido «¡Bola!» en dirección mía, debo determinar su objetivo con respecto a mis estados intencionales/atencionales hacia una tercera entidad [la bola]*. (Tomasello, 2003, pp. 102-3, énfasis agregado)

En estos términos, Tomasello piensa que el tipo de entendimiento que soporta y explica la atención conjunta debe ser la base de la habilidad de atribuir a otro sujeto estados intencionales complejos (estados intencionales *respecto a otros estados intencionales*: ¡los míos!). Esto implica inmediatamente que el entendimiento detrás de los episodios de atención conjunta está estructurado jerárquica y recursivamente. Si un par de sujetos participa en un episodio de atención conjunta, ambos entienden que el otro tiene estados intencionales *respecto a* sus estados intencionales, pero, a su vez, también entienden que estos estados intencionales de segundo orden están presentes; y, en virtud de esto, entienden que estos estados mentales de segundo orden son respecto de los demás, lo que implica la presencia de estados mentales de tercer orden, y así sucesivamente. Tomasello (2014) manifiesta recientemente la misma posición. En las conclusiones de su libro escribe:

[¿Q]ué lleva al individuo a la interacción [social] que le permite participar en la atención conjunta, de una manera que otros simios y niños más pequeños no pueden? [...] para nosotros esto significa que algo como *la lectura de mente recursiva o la inferencia* [...] debe ser parte de la historia de la intencionalidad compartida. Desde el punto de vista del individuo, la intencionalidad compartida simplemente se experimenta como un compartir, pero su estructura subyacente, que refleja su evolución, es que cada participante durante una interacción social *puede potencialmente tomar la perspectiva de otros tomando su perspectiva, y así sucesivamente al menos unos pocos niveles*. (2014, p. 152, énfasis agregado)

Esta explicación es central en el debate sobre la naturaleza de la atención conjunta. Sin embargo, no me detendré a argumentar en

contra de esta posición<sup>5</sup>. Basado en Gallagher (2011), me limitaré a señalar que el problema central de esta explicación es que la compleja estructura cognitiva implicada excluye la idea según la cual el entendimiento que soporta y explica los episodios de atención conjunta es *fundamentalmente* perceptual. Si no fuese así, no podría ser parte de la explicación de cómo las niñas y los niños prevebales entran al mundo social (cf. Gómez, 2005; Hobson and Hobson, 2008; Leavens, 2011; Scaife and Bruner, 1975).

El reconocimiento del carácter perceptual del entendimiento detrás de la atención conjunta es esencial para adelantar una discusión informativa respecto a qué es y cómo explican los episodios de esta clase de fenómenos. Sin embargo, no todas las maneras de apelar a la percepción son adecuadas para dar cuenta de la atención conjunta. Esto, porque no todas son suficientes dar cuenta de la *cooperación* y *apertura* involucradas en esta clase de episodios (cf. Gallagher, 2011).

La razón para entusiasmarse con unas teorías de la experiencia perceptual y no con otras se debe a que aún no hay suficiente claridad respecto a cuáles son los aspectos relevantes de la interacción social/ coordinación que implican atención conjunta, pues aún no hay una discusión filosófica que clarifique qué cuenta como un ejemplo paradigmático de atención conjunta, ni que nos diga cuáles son las condiciones suficientes y necesarias o cuál es la definición de la atención conjunta con la que todas las partes estemos de acuerdo.

Tengo la impresión de que esta falta de claridad respecto a qué cuenta como un caso paradigmático de atención conjunta afecta el texto de María Clara Garavito «Atención conjunta en ciencias cognitivas y fenomenología». En términos generales, su tesis —si la interpreto bien— consiste en la afirmación de que los episodios de atención conjunta correctamente descritos son, más bien, un caso particular de un fenómeno cognitivo más general llamado *constitución de sentido*. Y, entendida así, la atención conjunta debería recibir el mismo tipo de explicación.

---

5 Para una revisión detallada de las razones para dudar de la explicación de Tomasello, véase Gallagher (2011). Además, Eilan (2005); Hobson, P. and Hobson, J. (2008); y Leavens (2011) presentan objeciones interesantes.

La constitución de sentido es un fenómeno dinámico que requiere participación activa de los sujetos y que explica por qué razón el *mundo material luce en la experiencia perceptual* tal como lo hace. Esta constitución da cuenta de cómo es el tipo de *entendimiento*, ya sea perceptual o vivencial, que soporta y explica diferentes tipos de interacción social con el mundo y con los otros. Y la explicación —si entiendo correctamente— pasa por el tamiz de la descripción de cómo se aprende a coordinar/sincronizar el cuerpo con otras entidades del ambiente, a saber, con objetos y otros sujetos. Esta coordinación da origen al *cuerpo vivido*<sup>6</sup> y, simultáneamente, al *mundo percibido* o al *mundo entendido*. Mediante esta descripción se hará clara la manera en la que el mundo material está dado (esto es, constituido) en la experiencia perceptual y en el entendimiento. Ella escribe:

El cuerpo vivido, a diferencia del cuerpo orgánico, se constituye en la sedimentación de habitualidades por las que un mundo aparece: que el teclado de mi computador aparezca de cierta forma [...] depende de la organización motora, no solamente de las manos, sino del cuerpo entero. (Garavito, 2021, p. 270)

Su objetivo es presentar una re-descripción de los episodios de atención conjunta en la que sea entendible cómo la coordinación/sincronización entre los cuerpos conduce a una constitución de sentido (esto es, una constitución o *emergencia de mundo percibido/entendido*) que permite dar cuenta del éxito en la interacción social/coordinación de los episodios de atención conjunta. Su ejemplo paradigmático de coordinación entre cuerpos es la producción improvisada de una pieza musical por parte de un cuarteto de jazz. De acuerdo con ella:

a pesar de las diferencias con los casos típicos de atención conjunta, desde una perspectiva en tercera persona uno puede encontrar en la situación de los jazzistas una triangulación: *unas subjetividades se coordinan en torno a una tarea* [...] Si se contrasta el caso de los músicos con un caso típico de intencionalidad compartida, como el

---

6 El cuerpo vivido es la manera en la que cada sujeto vive o, dicho de otro modo, experiencia su propio cuerpo. Se distingue del cuerpo orgánico en tanto que este último es objeto de estudio de la medicina o de la biología, y no necesariamente corresponde con la manera en que cada sujeto vive su cuerpo.

de una díada resolviendo un rompecabezas, se encuentra que una coordinación cinético/cinestésica, *afectiva y perceptual* fundamenta la experiencia; y el rompecabezas resuelto [y la pieza musical improvisada] podría verse como el producto que *emerge* de tal coordinación. (Garavito, 2021, p. 264, énfasis agregado)

El problema es que la descripción que ofrece Garavito de cuál es el mundo de sentido constituido durante la coordinación entre cuerpos es insatisfactoria a la hora de explicar el éxito de un episodio de atención conjunta. Su descripción no permite distinguir si dicha constitución de sentido (si esta manera de darse el mundo material en la percepción) da cuenta del éxito de episodios de actividades individuales o actividades conjuntas. Parte del problema, considero, es su explicación del éxito de la producción improvisada de una pieza musical de jazz. Creo que este episodio de interacción social/coordinación no es un caso de atención conjunta y, por lo tanto, lo que uno diga para explicar estos episodios, desorienta la explicación de la atención conjunta.

Mi objetivo en este breve texto es doble. En primer lugar, deseo presentar un ejemplo de interacción social comunicativa como el ejemplo paradigmático de un episodio de atención conjunta. A partir de este ejemplo, señalaré los que considero son elementos constitutivos del entendimiento que soporta y explica el éxito de estos episodios, y argumentaré por qué creo que los episodios de improvisación de piezas musicales no son episodios de atención conjunta. En segundo, lugar señalaré dónde está mi desacuerdo con María Clara Garavito. Mi punto será que la descripción que ella nos ofrece conduce, entre otras cosas, a eliminar el fenómeno de la atención conjunta. Al final creo que el punto de diferencia está en que ella y yo estamos hablando de fenómenos diferentes.



## Un episodio afianzado de atención conjunta

Empecemos por aquí:

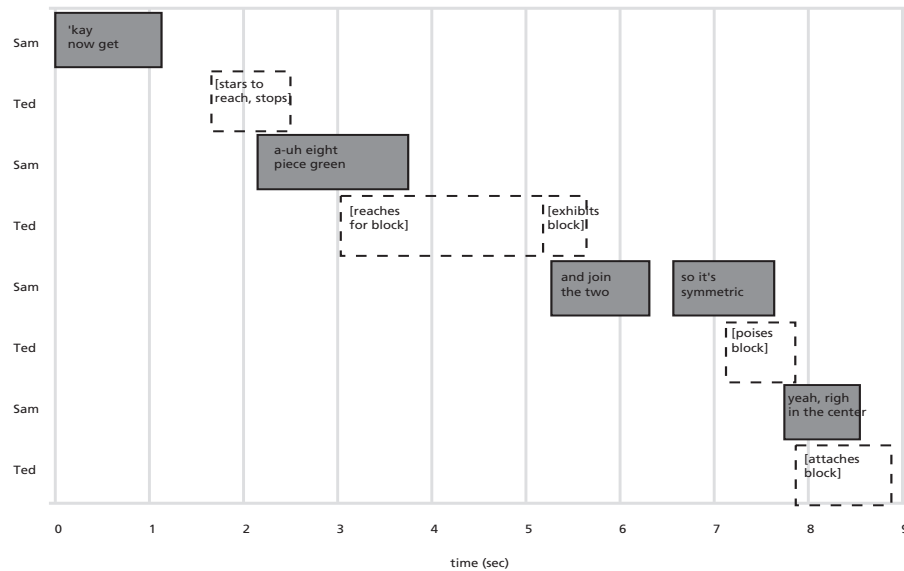


FIGURA 1. Clark and Krych (2004, p. 74)]<sup>7</sup>

La figura 1 registra una instancia paradigmática que orientará mi comentario sobre la atención conjunta. Esta figura muestra sólo un fragmento de 9 segundos, aproximadamente, de un episodio de interacción social más largo, de 10 minutos aproximadamente, entre Sam y Ted (dos estudiantes universitarios de la Universidad de Stanford). Ellos están resolviendo una tarea compartida: ensamblar un modelo de Lego® basado en un conjunto de instrucciones y un grupo de fichas específico.

Para resolver esta tarea (esto es, ensamblar el modelo) es necesario identificar las fichas y los lugares donde deben ir ubicadas. En las condiciones del experimento de Clark y Krych (cf. 2004, pp. 65-66) se establece que Sam conoce las instrucciones para ensamblar el modelo mientras Ted las ignora, pero Ted puede manipular las fichas mientras Sam no tiene acceso a ellas. Por esta razón, ninguno puede resolver la

<sup>7</sup> De acuerdo con las convenciones usadas por Clark y Krych (2004), lo que está encerrado en rectángulos grises son las emisiones de oraciones, el área rodeada de una línea punteada indica los movimientos realizados y la extensión de cada área dice cuánto tiempo tomó cada acción: emisión o movimiento corporal.

tarea individualmente, necesitan del otro; no basta con que Ted sepa cuál es la ficha correcta, es necesario que Sam también lo sepa, y viceversa. La tarea de ensamblar el modelo será un fracaso a menos que cada ficha sea al menos *mutuamente identificada* por ambos (cf. Eilan, 2005).

Lo más notable de este episodio de interacción social/coordinación es la experticia de ambos al hacer que la ficha sea mutuamente identificada,<sup>8</sup> ¡tan solo tardaron tres segundos!<sup>9</sup> ¿Cómo lo hacen? Comencemos por el camino que toman Sam y Ted. Sam inicia con la siguiente instrucción, tarda casi cuatro segundos en emitirla: «Ok, ahora coge... eh, una ficha verde de ocho». Aquí Sam se interrumpe a sí mismo mientras espera la reacción de Ted y éste, por su parte, anticipándose al final de la instrucción —justo en la mitad— inicia el proceso de búsqueda de la ficha, pero se interrumpe a sí mismo hasta que la instrucción esté completamente emitida. Entre el segundo 5 y 6, ambos identifican la ficha correcta.<sup>10</sup>

La mutua identificación de la ficha depende de que haya un acuerdo manifiesto respecto a *cuál* es la ficha<sup>11</sup>. Ambos logran este

8 Ocurre entre el segundo 2 y el segundo 5; y la localización del lugar correcto tardó igualmente tres segundos: ocurre entre el segundo 6 y el segundo 9.

9 Doy por sentado que la estrategia no fue establecida previamente. Esto es, es una estrategia improvisada.

10 La estrategia comunicativa que usan Sam y Ted no implica comprometerse con la explicación sugerida por Grice (1957) o Tomasello (2003; 2014). John McDowell (1998) ha adelantado un argumento en contra de la explicación tipo Grice de los actos comunicativos. Luego de desarrollar su argumento, McDowell afirma que, por ejemplo, las conductas comunicativas como el «señalamiento con el dedo índice» son expresiones inmediatas de la intención de compartir conocimiento respecto de cuál aspecto del ambiente se ha identificado. Estas intenciones son abiertas en el sentido en que son un asunto de conocimiento común entre los sujetos. La apertura de estas intenciones implica que la acción de un hablante cuenta como la realización de este acto comunicativo solamente si una audiencia lo reconoce como tal; de manera que lo único que requiere una audiencia para reconocer esto es percibir las intenciones del hablante manifestadas en las expresiones corporales del hablante. Si esta propuesta tiene sentido, es posible articular una noción de intención comunicativa sin el aparataje cognitivo que implica la propuesta de Grice.

11 Deben estar de acuerdo en que la ficha que debe ser movida (información que tiene Sam) y la ficha que se va a mover (información que tiene Ted) sean la misma.

acuerdo mediante la *exhibición y confirmación* de que ésa es la ficha. Siguiendo lo que registra la figura 1, podemos decir que —a la vez que la ficha se exhiba durante los segundos 5 y 6— el simple hecho de que simultáneamente Sam prosiga con las instrucciones para ubicar dicha ficha sobre el modelo constituye, en sí mismo, la confirmación de que *ambos* identifican *esa* ficha como la correcta. Esto hace que sea racional continuar con la realización de la tarea.

Debido a que el acuerdo manifiesto respecto de cuál es la ficha correcta se basa, parcialmente, en la habilidad de exhibir la ficha, dicho acuerdo depende de que ambos *identifiquen perceptualmente* la ficha (cf. Evans, 1982, 151-170). El acuerdo consistirá en que haya una fusión entre la identificación perceptual de Sam acerca de la ficha y la identificación perceptual de Ted acerca de la ficha. El éxito de esta tarea descansa precisamente en esta fusión, pues es la base informacional sobre la cual los sujetos monitorean las acciones presentes durante los episodios en los que se resuelve una tarea de este tipo.

### ***El corazón de la atención conjunta: la apertura***

¿Cómo lograr este acuerdo, esta fusión? El acuerdo se logra cuando Ted le muestra la ficha a Sam. Y esto ocurre bajo dos condiciones:

- (I) Ted y Sam identifican perceptualmente la misma ficha, y
- (II) Esta identificación perceptual que ocurre durante el episodio interactivo es abierta.

Una identificación perceptual es abierta cuando —durante el episodio interactivo— la identificación perceptual de la ficha es, al menos, *mutuamente accesible* a Ted y a Sam; es decir, cuando dicho aspecto *de* la interacción social es *público* (cf. Eilan, 2005). Si este aspecto de identificación perceptual del episodio no fuera público, sería privado e inaccesible a los demás y, en ese caso, Ted y Sam no podrían generar el acuerdo manifiesto sobre cuál es la ficha; por lo tanto, no podrían colaborar efectivamente para resolver la tarea: ensamblar el modelo de Lego®. El uso de la expresión «apertura» apunta a señalar el carácter público de la percepción. El hecho de que ésta sea abierta hace parte de la explicación del éxito de este tipo de tareas colaborativas.

La apertura de la identificación está soportada por un tipo de entendimiento que orienta a Sam y Ted para lograr el éxito en su tarea. Ahora, ¿cómo especificar dicho entendimiento? La historia que debemos contar para especificar dicho entendimiento debe tomar en cuenta que el acuerdo que se requiere para el éxito de la tarea está anidado en la articulación de diferentes habilidades cognitivas; éstas son habilidades verbales (como ofrecer instrucciones y seguir instrucciones) y habilidades no-verbales (como discriminar perceptualmente objetos o exhibir/mostrar objetos). El entendimiento que se requiere para dar cuenta de la apertura es el mismo que se requiere para dar cuenta de la articulación de estas habilidades, porque sólo podemos hacer pública la percepción mediante la articulación de dichas habilidades (tal como se muestra en la figura 1). El éxito de la articulación descansa en el *monitoreo* continuo de las condiciones cognitivas y materiales del episodio de interacción social/coordinación. Con esto a la mano, sólo necesitamos explicar cómo el ejercicio de las habilidades verbales y no-verbales conduce a que la percepción sea abierta/pública; o, en términos de Garavito (2021): necesitamos explicar cómo *se constituye el sentido* abierto/público de la percepción a partir del ejercicio y monitoreo de dichas habilidades.

### **La improvisación musical y la atención conjunta**

Los episodios de atención conjunta, de acuerdo con mi descripción, son episodios de interacción social/coordinación en los que se resuelve una tarea colectiva basada en conductas como señalar, exhibir o mostrar aspectos del ambiente, o incluso usar expresiones verbales como «eso» o «aquí» tal que estas identificaciones sean abiertas.

En estos términos, la interacción social/coordinación que ocurre durante los episodios de atención conjunta depende del uso y entendimiento de estas conductas de señalamiento o exhibición. Durante un episodio de atención conjunta, un sujeto entiende la conducta de señalamiento o exhibición cuando —basado en la percepción— singulariza (selecciona) un aspecto particular del ambiente respecto del cual puede monitorear si la interacción social/coordinación contribuye a la realización de la tarea o no (en el caso la interacción social que registra la Figura 1, si la interacción social contribuye al ensamble del modelo de

Lego® o no). Hay un episodio cabal de atención conjunta cuando dicha singularización es abierta/pública: esto ocurre cuando ambos sujetos monitorean la contribución de la interacción social/coordinación a la realización de la tarea sobre la base de la misma singularización.

Algunos de los ejemplos de interacción social/coordinación —mencionados antes— que motivan la discusión sobre la atención conjunta no encajan en esta caracterización. Los episodios de interacción social/coordinación entre primates, o entre primates y humanos, no cuentan como episodios de esta clase porque el uso de gestos de señalamiento o exhibición no cuentan con la suficiente complejidad para garantizar la *apertura* de la identificación perceptual (cf. Call, J. and Tomasello, M., 2005). En otras palabras, no coordinan a partir de la apertura de la identificación de algunos objetos singulares del ambiente. Así, el tipo de entendimiento que soporta y explica el éxito de estos episodios es distinto al que corresponde con la atención conjunta.

Sucede igual con algunos casos de interacción social/coordinación entre niños y algunos casos de interacción social/coordinación entre adultos. No son casos de atención conjunta las conversaciones que no están basadas en la percepción singular de algún aspecto del ambiente como lo es una conversación telefónica y tampoco lo es el caso de la producción improvisada de una pieza musical de jazz (en adelante: una improvisación musical).

En la *descripción* de Garavito (2021) del éxito de una improvisación musical, no hay nada que cuente como una singularización, o una identificación perceptual, de algún aspecto *singular* del ambiente. De hecho —éste será mi punto en la siguiente sección—, esta descripción no es necesariamente acerca de lo que ocurre durante un episodio grupal. Veamos.

Su idea es que el nivel correcto de descripción de la interacción social/coordinación de un episodio de improvisación musical no está en términos de un entendimiento acerca de los estados mentales de otros sujetos, singularizaciones u objetos compartidos, sino en términos de cómo la percepción *afectiva*<sup>12</sup> del propio cuerpo se *sincroniza*

---

<sup>12</sup> Todas estas percepciones son afectivas puesto que están constituidas por una valoración afectiva, al menos en términos de agrado o desagrado.

‘*Entrainment*’<sup>13</sup> con la percepción de los rasgos del ambiente. Estos rasgos ambientales percibidos afectivamente pueden ser movimientos o sonidos de otros cuerpos, movimientos o sonidos del propio cuerpo, o sonidos rítmicos (provenientes de los instrumentos, por ejemplo). En todo caso, deseo subrayar, ninguno de estos rasgos está dado en la percepción (a la consciencia) como un objeto singular y, por lo tanto, ninguno puede ser objeto de singularización. En sus palabras:

En los músicos, el *entrainment* es la improvisación debida a pequeñas alteraciones en los patrones de la interpretación individual que son inmediatamente seguidas por modificaciones en la interpretación de los otros miembros del grupo, por ejemplo, a través del cambio en el movimiento de las manos. Aquí el *entrainment* lleva a la coordinación de la percepción auditiva de la interpretación de otro con la percepción cinestésica del movimiento de sus manos. (Garavito, 2021, p. 269)

Esta sincronización es dinámica en la medida que fluctúa y se mantiene a través del tiempo que dura la improvisación. El entendimiento (afectivo-perceptual) que soporta y explica el éxito de los episodios de improvisación musical gira en torno a —esto es, está tematizada alrededor de— esta sincronización dinámica. Así, está dado el mundo en la percepción afectiva (en la vivencia experiencial afectiva) durante esta clase de episodios. Sin embargo, la sincronización dinámica *no está* —y *no puede estar*— dada como un objeto singular durante el episodio; la sincronización dinámica es un rasgo del episodio y no un objeto. En este sentido, el entendimiento que soporta y explica el éxito de la improvisación no es acerca de nada singular, sino más bien trata un rasgo bastante complejo que constituye el episodio de improvisación.

---

<sup>13</sup> María Clara Garavito toma este concepto de la teoría de sistemas dinámicos. En esta teoría, dos (o más) procesos se encuentran en *entrainment* cuando la periodicidad de sus movimientos es la misma. El ejemplo paradigmático aquí es la periodicidad de dos péndulos al moverse de izquierda-derecha. Estos están sincronizados cuando el tiempo que demoran en llegar al punto máximo de altura a su izquierda (o derecha) es el mismo para ambos.

Este hecho inmediatamente implica que los episodios de improvisación musical no sean casos de atención conjunta, por su diferencia en el tipo de entendimiento. Sin embargo, me gustaría decir algo más fuerte respecto a este tipo de descripción de la improvisación musical en conexión con la atención conjunta.

### **Sincronización en soledad**

El entendimiento que soporta y explica la improvisación es acerca de la sincronización, esto es, *el tema* de la percepción al interior del episodio es la sincronización de las percepciones afectivas del propio cuerpo, de los rasgos ambientales y de los movimientos de otros cuerpos. Esto implica que los objetos singulares —*qua* objetos singulares— quedan por fuera del tema de la percepción durante la improvisación; no están dados en el entendimiento ni en la percepción como objetos singulares. Igualmente sucede con los otros sujetos: no están dados al entendimiento ni a la percepción como otros sujetos individuales, ni como otros cuerpos individuales. Esto no significa que estén fuera de la percepción absolutamente, sino que hacen parte de la experiencia perceptual pre-reflexiva; esto es, están en la periferia de la experiencia perceptual, pero no son el tema de la percepción o entendimiento.

En esa coordinación, esa agencialidad y alteridad que se constituyen en la constitución de unidad, son experiencias implícitas, pre-reflexivas [...], es decir, no aparecen como tema. Por ejemplo, cada uno de los músicos se coordina con los otros en la emergencia de una pieza musical sin que para esto medie, en su experienciar, una tematización de los otros y de ellos mismos. (Garavito, 2021, p. 271, énfasis agregado)

Cualquier elemento singular, sean objetos u otros sujetos, queda en el plano prereflexivo, y esto significa que, pese a que hagan parte de las condiciones de posibilidad del episodio de improvisación, no son lo que explica el éxito de este episodio, porque no son lo que orienta la improvisación.

Esto también localiza los gestos de exhibición —como señalar con la mano— en el reino de las experiencias pre-reflexivas. Esta con-

ducta (sea hecha por el sujeto mismo o por otro sujeto) no se da en la percepción o entendimiento como lo que hace un agente intencional con un cuerpo individual, sino se da como un rasgo del ambiente. Cualquier expresión o gesto demostrativo se percibe como un rasgo ambiental indistinguible de cualquier otro con el que se aprende a sincronizar. Cuando uno mismo señala —dice Garavito (2021)— aprende por asociación a hacer que las cosas cambien de lugar, se habitúa a un cambio en los rasgos del ambiente por otros rasgos del ambiente.

En el caso concreto de la tarea, el gesto es un hábito relacionado con cierta dación de un objeto o tarea y que tiene que ver con la coordinación motora y afectiva con el otro —por ejemplo, cada vez que señalo, la ficha es movida a cierta ubicación. De modo parecido, en la coordinación de los músicos, quien que cambia la posición de las manos por un cambio en la interpretación de otro, no lo hace porque le sea dado, ni explícita, ni implícitamente, el gesto del rostro del otro que anticipa el cambio de su interpretación. (Garavito, 2021, p. 272)

En este nivel de descripción, la conducta de señalar es una manera de responder sincronizadamente a rasgos ambientales o de hacer que otros rasgos ambientales cambien sincronizadamente y, en estos términos, es indeterminado si dicha sincronización hace parte o no de una constitución de un mundo de sentido intersubjetivo. En otras palabras, no es claro si dicha sincronización contribuye al tipo de entendimiento que soporta y explica un episodio de interacción social/coordinación con otros, o si más bien contribuye al tipo de entendimiento que soporta y explica un tipo de coordinación/sincronización individual. Mi impresión es que dicha sincronización contribuye a lo segundo.

En los términos de Garavito, el mundo constituido mediante esta sincronización está dado a la percepción o el entendimiento como una relación sincronizada dinámicamente entre lo que ocurre con mi cuerpo y otros rasgos del ambiente, y debido a que ninguno de estos está dado singularmente, el resultado de esta sincronización será una manera de darse el mundo tal que explique y soporte episodios de *coordinación individual* entre el cuerpo del sujeto y algún aspecto del ambiente. Esta sincronización puede ocurrir cuando una niña de



tres meses, por ejemplo, aprende a sincronizar percepciones entre lo que hace con su cuerpo y lo que sucede con un sonajero. Este tipo de sincronización entre las percepciones de los rasgos del ambiente y del propio cuerpo conducirá a una manera de darse el mundo individualmente, y no grupalmente. Los otros desaparecen en la percepción, así como señala el epígrafe de este texto.

Si éste es el nivel correcto de descripción, es consistente con el hecho de que la manera en la que el mundo es dado a la percepción durante el episodio de improvisación corresponda con el entendimiento individual y no grupal (¿solipsismo?). No hay diferencia sustancial en la constitución del entendimiento que explicaría que un sujeto improvise en solitario, apoyado con alguna grabación de otros instrumentos musicales, y el entendimiento que explicaría que improvise en grupo.

En mi explicación, la diferencia entre los casos de sincronización individual y grupal se encuentra, en parte, en que hay cosas que no pueden ocurrir durante el episodio, como *mostrar* aspectos singulares del ambiente. Y esta posibilidad se explica con un tipo de entendimiento apropiado. En contraste con la explicación de Garavito, dado que durante un episodio de improvisación no hay singularización, entonces no hay atención conjunta, y el tipo de entendimiento que explicaría el éxito en la improvisación no puede explicar el éxito en el caso de Ted y Sam, esto es, no puede dar cuenta del éxito de los episodios de atención conjunta. Por esta razón, al cambiar de nivel de descripción cambiamos de tema y, por lo tanto, no podemos sacar consecuencias de un fenómeno a otro. Investigar sobre el éxito durante la improvisación para explicar el éxito durante la atención conjunta es un camino desorientador. Esto no significa que el nivel de descripción de Garavito (2021) para dar cuenta de la improvisación sea incorrecto; por el contrario, lo encuentro sumamente atractivo. Sólo que no es el nivel apropiado para la atención conjunta.

### Bibliografía

- Call, J. and Tomasello, M. (2005). What Chimpanzees Know about Seeing, Revisited: An Explanation of the Third Kind. En N. Eilan, J. Roessler, Ch Hoerl, and T. McCormack (eds.), *Joint Attention: Communication and Other Minds*. New York: Oxford UP.

- Clark, H., and Krych, M. (2004). Speaking while monitoring addressees for understanding. *Journal of Memory and Language*, 50 (1), 62-81.
- De Jaegher, H. (2006). *Social Interaction Rhythm and Participatory Sense-Making: An Embodied, Interactional Approach to Social Understanding, with Some Implications for Autism*, [Tesis doctoral]. Sussex: University of Sussex. <https://hannedejaegher.files.wordpress.com/2010/05/hannedphilboekformaat.pdf>
- Dunham, Ph., and Dunham, F. (1995). Optimal Social Structures and Adaptive Infant Development. En Ch. Moore, and Ph. Dunham (eds.), *Joint Attention: Its origins and Role in Development*. New York: Psychology Press.
- Eilan, N. (2005). Joint Attention, Communication and Mind. En N. Eilan, J. Roessler, Ch Hoerl, and T. McCormack (eds.), *Joint Attention, Communication and Other Minds, Issues in Philosophy and Psychology*. Oxford: Oxford UP.
- Gallagher, S. (2011). Interactive Coordination in Joint Attention. En A. Seemann (ed.), *Joint Attention: New Developments in Psychology Philosophy of Mind, and Social Neuroscience*. Massachusetts: MIT Press.
- Garavito, M. (2021). Atención conjunta en ciencias cognitivas y fenomenología. *Imágenes de la mente, el lenguaje y el conocimiento*. Bogotá: Universidad Nacional de Colombia.
- Gómez, J. (2005). Joint Attention and the Notion of Subject: Insights from Apes, Normal Children, and Children with Autism. En N. Eilan, J. Roessler, Ch Hoerl, and T. McCormack (eds.), *Joint Attention: Communication and Other Minds*. Oxford: Oxford UP.
- Grice, H. (1957). Meaning. *Philosophical Review*, 66, 377-388.
- Hobson, P. and Hobson, J. (2008). Engaging, sharing, knowing: Some lessons from research in autism. En J. Zlatev, T. Racine, Ch. Shina, and I. Itkonen (eds.), *The Shared Mind: Perspectives on Intersubjectivity*. Philadelphia: John Benjamins Publishing Company.
- Leavens, D. (2011). Joint Attention: Twelve Myths, En A. Seemann (ed.), *Joint Attention: New Developments in Psychology, Philosophy of mind and Social Neuroscience*. Massachusetts: MIT Press.

- Kendon, A. (1990). *Conducting Interaction: Patterns of Behavior in Focused Encounters (Studies in Interactional Sociolinguistics, 7)*. Cambridge: Cambridge UP.
- Krueger, J. (2013). Merleau-Ponty on Shared Emotions and the Joint Ownership Thesis. *Continental Philosophy Review*, 46 (4), 509-531.
- Krueger, J. (2014). Affordances and the musically extended mind. *Frontiers in Psychology*, 4, 1-13.
- Soproni, K., Miklósi, A., Topál, J., and Csányi, V. (2002). Dogs' (Canis familiaris) responsiveness to human pointing gestures. *Journal of Comparative Psychology*, 116 (1), 27-34.
- Scaife, M. and Bruner, J. S., (1975). The capacity for joint visual attention in the infant. *Nature*, 253, 265-266.
- Tomasello, M. (2003). *The Cultural Origins of Human Cognition*. London: Harvard UP.
- Tomasello, M. (2014). *A Natural History of Human Thinking*. London: Harvard UP.
- Tomasello, M., Carpenter, M., Nagell, K. (1998). Social Cognition, Joint Attention, and Communicative Competence from 9 to 15 Months of Age. *Monographs of the Society for Research in Child Development*, 63 (4), 1-174.
- McDowell, J. (1998). Meaning, Communication and Knowledge. En J. McDowell (ed.), *Meaning, Knowledge and Reality*. Cambridge: Harvard UP.