

Recibido: 01.03.2021 • Aceptado: 22.04.2021

Palabras clave: Lenguaje, filosofía, teoría de juegos, comunicación, convención.



Problemas de comunicación: lenguaje, filosofía y teoría de juegos

ANDREA ONOFRI
andrea.onofri@uaslp.mx
FACULTAD DE CIENCIAS SOCIALES Y HUMANIDADES, UASLP

Considera la siguiente situación. Tú y yo queremos inventar un lenguaje, quizás para pasar el tiempo, quizás porque necesitamos comunicarnos por medio de un código secreto. La tarea será más o menos ardua, dependiendo del tipo de lenguaje que queremos inventar, pero tenemos una herramienta a disposición: el lenguaje que ya hablamos (el español, en nuestro caso). Podemos entonces comunicarnos en español para definir las características de nuestro nuevo lenguaje, como sus reglas gramaticales y su léxico. Por ejemplo:

Yo: Llamemos al perro *cat*, ¿te parece?

Tú: No me gusta, mejor *dog*

Hay que tomar en cuenta que este tipo de solución requiere que tú y yo tengamos de antemano un lenguaje común, que usaremos para ponernos de acuerdo sobre nuestro nuevo lenguaje. Por ejemplo, la conversación de arriba sólo puede llevarse a cabo si tú conoces el significado de la palabra *perro* en español; de no ser así, no podrías entender lo que te estoy proponiendo. Como escribió el filósofo británico Bertrand Russell (1921): “We can hardly suppose a parliament of hitherto speechless elders meeting together and agreeing to call a cow a cow and a wolf a wolf” (p. 113).

Para crear un lenguaje por medio de un acuerdo explícito, como en el ejemplo de arriba, debemos entonces tener un lenguaje anterior. Pero ¿cómo se formó este lenguaje anterior? Si fue por acuerdo explícito, tendremos otra vez que explicar cómo fue posible dicho acuerdo. Este es un caso que en filosofía llamamos regreso al infinito: no todo lenguaje puede haberse originado por un acuerdo explícito, ya que esto implicaría una serie infinita de acuerdos de este tipo. Algunos lenguajes deben entonces haberse formado de otra manera. De hecho, esto aplica a muchos lenguajes, incluso el nuestro; el español no nació en una serie de reuniones entre hablantes que quisieran inventar un nuevo idioma.

Pero, si no es por medio de un acuerdo explícito, ¿de qué otra manera podría formarse un lenguaje? Como Aristóteles (*Categoriae et Liber de Interpretatione*, 1949) y

otros autores han señalado, el lenguaje parece ser una convención humana. Por ejemplo, la decisión de utilizar una palabra en lugar de otra es convencional. En español decimos *perro* y en inglés decimos *dog*, pero ninguna de estas expresiones es mejor que la otra y ambas cumplen su función, siempre y cuando sean usadas de manera uniforme por los hablantes del idioma correspondiente (si alguien usara la palabra *dog* para referirse a los gatos, esto afectaría la posibilidad de comunicarnos exitosamente). Encontramos las mismas características en otras convenciones humanas, como las que rigen el sentido de la circulación vial. En México conducimos por el lado derecho de la carretera y en Inglaterra lo hacen por el lado izquierdo, pero las dos reglas son igualmente buenas; lo único que importa es que sean usadas de manera uniforme en el país en cuestión (si alguien condujera por el lado izquierdo en México sería peligroso para él y para los demás conductores).

Parece entonces natural clasificar al lenguaje como una convención. Sin embargo, también parece necesario que haya un acuerdo explícito que permita introducir una convención en una comunidad. Por ejemplo, para introducir la convención de conducir por el lado derecho, parece necesario que los legisladores manifiesten su consenso explícito al respecto. Pero ya hemos visto que no todo lenguaje puede haberse formado por medio de un acuerdo de este tipo. Enfrentamos entonces un dilema (un predicamento común en filosofía): ¿cómo

explicar el carácter de convención del lenguaje sin caer en un regreso al infinito?

El filósofo estadounidense David Lewis intentó ofrecer una nueva solución a este antiguo problema en su libro *Convention* (1969). La obra ha tenido un gran impacto en filosofía del lenguaje y en otras disciplinas (cabe mencionar que Lewis tenía apenas 28 años cuando el libro fue publicado). Lewis trata de demostrar cómo una convención puede formarse sin que haya un acuerdo explícito. En la primera parte de la obra, se ofrece un análisis general de lo que es una convención, para luego enfocarse en el caso específico de las convenciones lingüísticas. Si una comunidad pudiera establecer una convención lingüística sin acuerdo explícito, tendríamos una solución a nuestro problema inicial, pero ¿cómo es posible establecer una convención de esta forma?

La solución de Lewis está inspirada en la teoría de juegos, una rama de la teoría de las decisiones. Un juego es una interacción entre dos o más agentes donde el éxito final de la interacción depende de las decisiones de ambos jugadores. Cada jugador tratará de predecir la decisión del otro para poder tomar la mejor decisión. Hay diferentes tipos de juegos. En algunos de ellos, los jugadores tienen intereses totalmente opuestos: entre más favorable sea el resultado para uno de los jugadores, menos lo será para el otro, y viceversa (véase Ross 2019 para una introducción accesible a este tema). Un ejemplo de este tipo de juegos es el de tres en raya, donde cualquier jugada que favorece a un jugador desfavorece al otro. Un juego donde los intereses de los jugadores son parcialmente opuestos es el famoso dilema del prisionero. Imagina que la policía ha arrestado a dos personas, sabiendo que cometieron un robo a mano armada juntos. No hay evidencia suficiente para condenarlos por este crimen, pero sí para condenarlos a dos años de cárcel por el robo del carro que usaron. El inspector hace entonces una oferta a cada prisionero: si tú confiesas que ambos realizaron el robo a mano armada y tu compañero no confiesa, él será condenado a 10 años y tú serás libre. Si ambos confiesan, ambos serán condenados a 5 años. Si ninguno de los dos confiesa, serán ambos condenados a dos años por robar el carro. Ningún jugador sabe qué decisión tomó su compañero.

Los resultados de las posibles decisiones de los dos jugadores pueden representarse así:

		Jugador II	
		Confesar	No confesar
Jugador I	Confesar	-5, -5	0, -10
	No confesar	-10, 0	-2, -2

Las dos opciones del Jugador I (confesar y no confesar) están representadas a la izquierda de la tabla, las del Jugador II arriba de la tabla. Cada celda representa el resultado de una combinación de acciones de los dos jugadores; por ejemplo, la celda que se encuentra abajo a la izquierda representa lo que va a pasar si el Jugador I no confiesa y el Jugador II confiesa. Los números negativos indican los años de cárcel que el jugador recibirá, así que -10 representa el peor resultado posible (10 años de cárcel) y 0 el mejor resultado posible. Los dos números en cada celda representan los años de cárcel que recibirán el Jugador I y el Jugador II, respectivamente; por ejemplo, la celda de abajo a la izquierda indica que, si deciden actuar de esta forma, el Jugador I recibirá 10 años de cárcel y el Jugador II ninguno.

Observa ahora la celda de arriba a la izquierda. Dada la decisión del Jugador I (confesar), el Jugador II no podría obtener un mejor resultado si cambia su decisión, o sea, decidiendo no confesar, ya que esto resultaría en 10 años de cárcel en lugar de 5. Lo mismo aplica a la decisión del Jugador I, dada la decisión del Jugador II. Este tipo de combinación es llamado equilibrio de Nash: dadas las estrategias de los otros jugadores, ningún jugador podría mejorar su resultado cambiando su estrategia. En el dilema del prisionero, sólo hay un equilibrio de Nash; al examinar las otras tres combinaciones posibles, verás que ninguna de ellas cumple con nuestra definición, (esto tiene consecuencias importantes, tanto en filosofía como en otras ramas, pero no podremos examinarlas aquí). El nombre equilibrio de Nash hace referencia al Premio Nobel John Nash, cuyo trabajo pionero representó una aportación fundamental a la teoría de juegos (Nash 1950, 1951).

En otros juegos, los agentes involucrados tienen intereses comunes, de ahí que, si uno de ellos toma la mejor decisión, esto beneficiará a todos. Lewis llama a estos juegos problemas de coordinación. Los problemas de coordinación son especialmente importantes para Lewis, quien ve la comunicación humana como un problema de este tipo. De hecho, Lewis mismo observa (1969, p. 3) que su propuesta está basada en el estudio de los juegos de coordinación realizado por Thomas Schelling (Schelling 1960), otro ganador del Premio Nobel.

Un ejemplo de un problema de coordinación es decidir de cuál lado de la carretera conducir, ya que a los conductores les interesa evitar una colisión y esto dependerá de la decisión de cada uno de ellos. La colisión se evitará en cuanto, digamos, dos conductores manejen por el mismo lado, que puede ser el derecho o izquierdo de la carretera. Podemos entonces representar la situación así:

		Jugador II	
		Conducir por el lado derecho	Conducir por el lado izquierdo
Jugador I	Conducir por el lado derecho	1, 1	0, 0
	Conducir por el lado izquierdo	0, 0	1, 1

Donde 1 representa un resultado positivo (ningún riesgo de colisión) y 0 representa un resultado negativo (riesgo de colisión). Observa cómo en este juego hay dos equilibrios de Nash: si los demás conducen por el lado derecho me conviene conducir por el lado derecho, pero esto cambia si los demás conducen por el lado izquierdo. Afortunadamente, en países como México e Inglaterra es sencillo predecir por cuál lado conducirán los demás, gracias a la existencia de convenciones consolidadas, comúnmente conocidas e

incluso oficializadas por las leyes del país en cuestión. Esto permite que los conductores solucionen su problema de coordinación de una manera segura y eficiente.

¿Debe una convención de este tipo ser introducida por medio de un acuerdo explícito? No necesariamente, observa Lewis (1969, cap. 1). Imagina un pueblo aislado y con pocos habitantes, donde todavía no existe una convención sobre el sentido de la circulación vial. Una persona A decide conducir por el lado derecho; toma su decisión de manera aleatoria, ya que no sabe por cuál lado conducirán sus conciudadanos. La persona B observa la conducta de A y decide hacer lo mismo, ya que por lo menos esto evitará una colisión en caso de encontrar a A. Por las mismas razones, la persona C decide conducir por el lado derecho después de haber observado a A y B. Lo mismo pasa con los demás integrantes de la población (D, E, etcétera), así que poco a poco se establece en el pueblo la convención de conducir por el lado derecho. Puede haber muchas pruebas y errores (y colisiones) en el proceso; puede ser que haya alternativas más seguras y eficientes para establecer nuestra convención, por ejemplo, un acuerdo explícito entre los habitantes. Aun así, el ejemplo indica que, bajo condiciones favorables, una convención puede establecerse sin acuerdo explícito.

¿Podría ocurrir algo similar con las convenciones lingüísticas? Empecemos con algo más sencillo, una forma de comunicación que Lewis llama señalar (*signaling*). Retomemos uno de los ejemplos de Lewis (1969, p. 129), imagina que se ha instalado un semáforo automático en el cruce entre dos calles de nuestro pueblito. El semáforo muestra luz roja a los que vienen de una calle y luz verde a los que vienen de la otra; la luz cambia después de un tiempo determinado. Ahora supongamos que no hay ninguna convención vigente que indique cómo interpretar los colores rojo y verde, ni siquiera por parte de los que instalaron el semáforo; simplemente se instaló para que los conductores lo usaran como referencia y desarrollaran su propia regla, reduciendo así el riesgo de accidentes en el cruce. Al no tener una convención previa, los conductores enfrentan un problema de coordinación: “¿debería detenerme o pasar si el semáforo está en rojo?”. De manera aleatoria, un conductor A decide detenerse cuando el semáforo está en rojo para él; otro conductor B decide entonces aprovechar la ocasión y pasar el cruce; otras personas (C, D, E...) observan

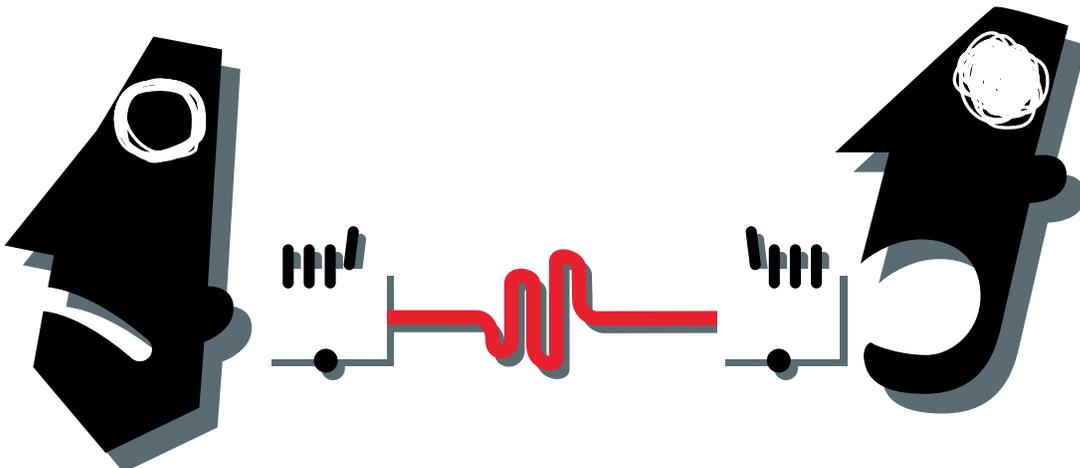
lo sucedido. Por el éxito de esta primera interacción, cada uno de los conductores presentes (A, B, C, D, E...) decide emplear la misma estrategia en futuras ocasiones (rojo → detenerse, verde → pasar), esperando que los demás hagan lo mismo. Al observar esto, otros conductores eligen la misma estrategia para evitar colisiones con los del primer grupo. La convención (rojo → detenerse, verde → pasar) llega así a consolidarse progresivamente en la comunidad.

Ahora parece que los colores rojo y verde tienen un significado para los miembros de la comunidad, es decir, se han convertido en señales (*signals*): rojo quiere decir detenerse y verde quiere decir pasar. Pero los colores no tenían ningún significado cuando el semáforo fue instalado, fue algo que adquirieron. Y es importante notar cómo lo adquirieron: no fue por algún acuerdo explícito, sino por una serie de decisiones de los conductores, que reforzaron poco a poco la convención hasta que finalmente ésta llegó a establecerse. Así, un sistema de señales puede formarse sin necesidad de un acuerdo explícito. Si esto aplica a las señales, ¿no podría pasar lo mismo con otra herramienta comunicativa, o sea el lenguaje humano?

La idea de Lewis es fascinante, pero enfrenta también varios retos. Lewis mismo señaló uno de ellos: nuestro lenguaje es mucho más complejo y flexible que un sistema de señales (Lewis 1969, pp. 160-161). La explicación de arriba podría entonces fracasar si se aplica al lenguaje, incluso suponiendo que funcione para sistemas comunicativos más sencillos, como las señales de tránsito. Sea como sea, el caso de las señales presenta sus propios problemas. Skyrms (1996, p. 84) evidencia uno de ellos. En el ejemplo del semáforo, considera el primer encuentro entre los conductores A y B. El semáforo está en rojo para A y en verde para B; A decide detenerse; B pasa el cruce sin inconvenientes. Imagina ahora un segundo encuentro. El semáforo está

otra vez en rojo para A y en verde para B, pero B podría tener dudas sobre qué decisión tomará A esta vez: "Es posible que, por el color del semáforo, A repita la decisión anterior; en este caso, yo podría pasar sin problema. Pero es también posible que A decida pasar el cruce, esperando que yo me detenga para regresarle el favor; en este caso, debería detenerme para evitar una colisión". Al no tener una respuesta segura, es prudente que B se detenga, aunque el semáforo esté otra vez en verde para él. Es decir, ninguno de los conductores tiene conocimiento suficiente sobre los demás, como para poder prever qué harán; será entonces mejor detenerse, en cualquier caso, haciendo caso omiso del color del semáforo. Bajo estas condiciones, la convención de pasar con el verde no podría formarse en esta comunidad, ni tendría el color verde algún significado para sus integrantes.

A pesar de estas dificultades, muchos autores han retomado las ideas de Lewis, desarrollándolas en nuevas direcciones e intentando solucionar los problemas antes mencionados. Una propuesta interesante e influyente ha sido elaborada por Brian Skyrms en una serie de trabajos (1996; 2010). Skyrms propone aplicar a las ideas de Lewis algunas nociones tomadas de la teoría evolutiva de juegos. En esta rama de la teoría de juegos, el mismo tipo de interacción (o sea, el mismo juego) se repite a lo largo del tiempo. Los jugadores participantes pertenecen a una población y usan una estrategia en el juego; las mejores estrategias son "premiadas" y se difunden en la población, las peores son "penalizadas" y se vuelven menos comunes. La teoría evolutiva de juegos se ha aplicado en diferentes ramas, por ejemplo, en la teoría de la selección natural en biología: si una especie actúa de manera más exitosa que otra, esto puede incrementar su capacidad de reproducirse, así que el comportamiento de la especie



Es doctor en filosofía por la University of St. Andrews, St. Andrews, Reino Unido. Cuenta con estudios posdoctorales en el Instituto de Investigaciones Filosóficas de la UNAM y por la University of Graz, Graz en Austria. Es profesor de tiempo completo y coordinador de la Licenciatura en Filosofía en la Facultad de Ciencias Sociales y Humanidades de la UASLP, en donde desarrolla investigaciones sobre el problema filosófico de la comunicación.



podría en el futuro volverse más común en la población en cuestión.

Skyrms argumenta que esta nueva perspectiva puede ayudar con los problemas de la teoría de Lewis. Volvamos al ejemplo del semáforo. En este caso, lo que no permite que la convención se consolide es la falta de conocimiento recíproco en los jugadores. Pero, observa Skyrms (1996, cap. 5; 2010, cap. 2), la comunicación no es un privilegio de nuestra especie, ya que muchas otras especies son capaces de comunicar con señales de diferentes tipos. ¿Cómo lo hacen? En muchos casos, (por ejemplo, las bacterias) es evidente que los miembros de la especie no tienen conocimiento en absoluto, ya que carecen de las capacidades cognitivas necesarias. Aun así, pueden comunicarse (por ejemplo, las bacterias lo hacen mediante señales químicas). Esto sugiere que la comunicación por medio de señales no requiere conocimiento recíproco por parte de los jugadores, sino un mecanismo de otro tipo.

¿Podría esto ayudar con el problema del semáforo? Imagina que, a la hora de acercarse al cruce, cada conductor tuviera una estrategia predeterminada (por ejemplo, una estrategia codificada en sus características genéticas). Algunos emplean la estrategia: rojo → detenerse, verde → pasar. Pero hay otros que se detienen en cualquier caso (es decir, hacen caso omiso del color del semáforo); finalmente, hay otros que pasan en cualquier caso (otra vez haciendo caso omiso del color). La primera estrategia parece superior a las otras, ya que la segunda obliga a todos los conductores presentes a gastar su tiempo esperando y la tercera conlleva una gran cantidad de accidentes. Bajo las condiciones adecuadas, la primera estrategia se reproduce entonces de manera más exitosa que las demás, por el beneficio otorgado a los jugadores que la emplean. Poco a poco, la estrategia empieza a difundirse en la población, hasta que es adoptada por todos sus miembros. La convención (rojo → detenerse, verde → pasar) se ha consolidado y los colores actúan ahora como señales,

y todo esto sin apelar a conocimientos ni acuerdos explícitos por parte de los jugadores!

¿Es posible aplicar una solución de este tipo en el caso del lenguaje humano, que es considerablemente más complejo que muchos sistemas de señales? Responder a esta pregunta de manera rigurosa requeriría una discusión mucho más extensa, que involucra tanto modelos matemáticos como consideraciones filosóficas (ver Planer y Godfrey-Smith, 2020, para una reseña más detallada); lo que he expuesto es, en el mejor de los casos, una serie de hipótesis interesantes y ejemplos sugerentes. A pesar de sus limitaciones, espero que esta exposición haya despertado tu interés en la reflexión filosófica sobre el lenguaje. Y espero también haber mostrado cómo un antiguo problema filosófico puede ser abordado con herramientas científicas, como la teoría de juegos y la teoría de la selección natural; intentos de este tipo son importantes si queremos que la filosofía sea integrada en una visión más amplia del mundo y no corte sus vínculos con otras disciplinas. 

Agradecimientos

A todos los que asistieron a mis ponencias sobre estos temas en Graz, Ciudad de México, Valencia y (en modalidad virtual) París. Muchas gracias a la doctora Erika Torres por la revisión formal del texto.

Referencias bibliográficas:

- Aristóteles, (1949). *Categoriae et Liber de Interpretatione*, (Minio Paluello, L. Ed.). Oxford: Clarendon Press.
 Nash, J. (1950). Equilibrium Points in n-Person Games. *Proceedings of the National Academy of Science*, pp. 48-49.
 Russell, B. (1921). *The Analysis of Mind*. London: George Allen & Unwin.
 Schelling, T. (1960). *Strategy of Conflict*. Cambridge: Harvard University Press.
 Skyrms, B. (1996). *Evolution of the Social Contract*. Cambridge: Cambridge University Press.