

«Unos dos mil tres indios». Reflexiones sobre la pragmática, el principio de economía y la teoría de juegos

Around two thousand and three indians

JESÚS P. ZAMORA BONILLA*

Resumen: El principio de racionalidad, o de economía, puede expresarse como la hipótesis de que los sujetos tienden a llevar a cabo aquellas acciones que maximizan la diferencia entre beneficios y costes. Este principio es ampliamente aplicado en las ciencias sociales, sobre todo, obviamente, en la teoría económica, si bien en el último medio siglo ha sido aplicado también de manera creciente a otras ramas de dichas disciplinas (sociología, ciencia política, antropología, historia, etc.). En este artículo se discute la posibilidad de aplicarlo a la lingüística y la pragmática, comentando varios de los trabajos realizados desde ese punto de vista, e ilustrándolos con su aplicación al caso del uso de expresiones vagas.

Palabras clave: Racionalidad, principio de economía, pragmática, vaguedad, lingüística, teoría de juegos.

Abstract: The rationality principle (or principle of economy) can be expressed as the hypothesis that individuals tend to carry out those actions that maximise the difference between costs and benefits. This principle is widely applied in the social sciences, mainly in economic theory, although during the last half century has also and increasingly been applied to other disciplines, like sociology, political science, anthropology, history, etc. This paper discusses the possibility of applying it to linguistics and pragmatics, by commenting some of the work done in this area from this point of view, and offering as illustration an analysis of vagueness.

Key words: Rationality, principle of economy, pragmatics, linguistics, vagueness, game theory.

1. La actividad lingüística y la teoría de la acción racional

En la mayoría de las llamadas «ciencias sociales» tiene una importancia cada vez mayor el llamado «principio de racionalidad» como fundamento básico para la explicación de las acciones humanas. Según este principio:

a) los hechos sociales deben ser comprendidos como el resultado de las acciones individuales (aunque hay un espacio, más o menos polémico, para tener en cuenta la influencia que el «contexto social» puede ejercer sobre la acción individual); y

Fecha de recepción: 20-05-2010. Fecha de aceptación: 8-07-2010.

* Dirección: Ciudad Universitaria. UNED. Senda del Rey, 7. 28040 MADRID. El profesor Zamora Bonilla es autor de *Ciencia pública – ciencia privada. Reflexiones sobre la producción del saber científico* (Fondo de Cultura Económica, 2005), y *Cuestión de protocolo. Ensayos de metodología de la ciencia* (Tecnos, 2005). Su libro más reciente es *La caverna de Platón y los cuarenta ladrones* (Le pourquois-pas, 2011).

b) las acciones individuales deben ser explicadas mediante el supuesto de que cada persona intenta tomar en cada caso la mejor decisión posible, valorándola de acuerdo con sus propias preferencias y con la información poseída por él.

Es bien sabido que el éxito teórico del principio de racionalidad ha sido, de todas formas, relativamente dispar según en qué ámbitos de las ciencias humanas. Donde más éxito ha obtenido es en la Economía, en parte porque (según la intuición de los economistas clásicos) es en los asuntos relacionados con el dinero donde la gente se preocupa más por tomar decisiones calculadamente, y en parte porque la teoría de la acción racional se presta a un tratamiento matemático muy potente que ha sido extraordinariamente útil para el desarrollo de la teoría económica. En cambio, en el campo de la Antropología se ha hecho hasta el momento un empleo mucho más limitado del principio de racionalidad, tal vez porque esta disciplina se ocupa preferentemente de investigar las tradiciones culturales, ante las cuales los individuos, más que calcular sopesadamente sus decisiones, parecen limitarse a dejarse llevar por las costumbres en las que crecen. Por último, en la Sociología y en la Ciencia Política es donde existe una polémica más viva sobre la validez y aplicabilidad del citado principio, pues ciertas escuelas lo adoptan y utilizan con profusión, mientras que otras son completamente reacias a emplearlo. Pese a todo, puede afirmarse que la tendencia en los últimos treinta o cuarenta años parece estar conduciendo a una modesta aproximación de posturas: por una parte, las investigaciones formales y experimentales van permitiendo desarrollar modelos de acción racional más complejos y más adaptados a las peculiaridades cognitivas de los seres humanos de carne y hueso; por otra parte, los paradigmas sociológicos más opuestos a la consideración del individuo como un sujeto autónomo de decisión (como, sobre todo, el estructuralismo) han ido dejando paso a perspectivas más flexibles y eclécticas.

En este contexto, ¿cuál es la situación de las ciencias del lenguaje? Pienso que es razonable afirmar que el uso del principio de racionalidad ha sido en ella muy limitado, más, incluso, que en las ciencias sociales más opuestas a dicho principio.¹ A esto ha contribuido, sobre todo, la consideración de la lengua como un sistema abstracto que se «impone» a los individuos, ya sea meramente como un conjunto de tradiciones (como puede ser en el caso del estructuralismo lingüístico), ya sea como un sistema de reglas universal, al que luego se adaptan las peculiaridades de los diferentes idiomas (como en el caso del generativismo). Esto significa que, según las concepciones habituales de la gramática, el hablante no tiene prácticamente libertad de opciones a la hora de construir sus mensajes, ni el oyente de interpretarlos, pues las normas lingüísticas lo hacen casi todo por ellos; de hecho, cualquier desviación de la norma corre el riesgo de terminar en la producción de mensajes antigramaticales. Como se sabe, el moderno interés por la pragmática ha matizado mucho esta postura, ya que las *intenciones* de los hablantes y las peculiaridades de los *contextos* en los que se realiza la comunicación se han mostrado como factores explicativos de gran relevancia a la hora de dar cuenta de numerosos fenómenos lingüísticos. Pero, en general, los principales desarrollos de la pragmática han tendido también a interpretar estos fenómenos desarro-

1 Para más detalles sobre la relación entre la teoría de la elección racional y la pragmática lingüística, puede consultarse el número monográfico sobre «Language and Games» que coordiné para la revista *Journal of Economic Methodology* (del Corral y Zamora Bonilla, 2008).

llando teorías acerca de las *normas* que regulan la comunicación, dejando relativamente poco espacio para el, digamos, «cálculo racional» por parte de los agentes que intervienen en ella.² El hecho de que las preferencias lingüísticas, especialmente en el lenguaje oral, suelen llevarse a cabo de manera casi automática por parte de los hablantes, y ser interpretadas aún con más rapidez por los oyentes, también ha podido crear la impresión de que de ningún modo podía haber presente allí un verdadero proceso de decisión, y menos aún un proceso de deliberación.

A pesar de todo, algunos desarrollos recientes de la lingüística, y de la pragmática en particular, hacen pensar que la situación puede estar cambiando, y que la aplicación de las técnicas analíticas asociadas con la teoría de la acción racional también llegará a ser útil para ofrecer explicaciones de bastantes fenómenos lingüísticos. Esto se corresponde, por otro lado, con el interés cada vez mayor de los especialistas en la teoría económica de los juegos por tener en cuenta explícitamente el medio natural de comunicación entre los agentes que dicha teoría estudia, y con el interés de un nutrido grupo de filósofos por entender la idea de racionalidad utilizando tanto las intuiciones procedentes de la reflexión sobre el lenguaje, como las procedentes de la reflexión sobre las decisiones. También hay que mencionar la ya vieja teoría de que muchos fenómenos lingüísticos (y en especial la evolución histórica de los lenguajes) están gobernados por algún tipo de «principio de economía», una teoría que las mencionadas herramientas analíticas, junto con las de la teoría de la computación, están enriqueciendo considerablemente.

No pretendo ni mucho menos criticar la tesis de que el lenguaje es una actividad fundamentalmente gobernada por normas, aunque sería interesante discutir la naturaleza de dichas normas (por ejemplo, en qué medida influyen en ella los factores genéticos y neurofisiológicos, frente a los sociales o culturales; o en qué medida son esas normas el producto necesario de algún tipo de dinámica, o un resultado determinado convencionalmente entre otros resultados igualmente posibles). Mis objetivos, mucho más modestos, son los siguientes:

- a) intentar mostrar que la interpretación de un fenómeno de interacción entre individuos (y en particular, de los fenómenos lingüísticos) como resultado de la operación de normas no es incompatible con la consideración de los intereses y las decisiones de esos individuos; y
- b) ofrecer un panorama muy sucinto de las herramientas conceptuales que la moderna teoría económica puede aportar para el estudio de los fenómenos de interacción, y de algunas contribuciones al estudio del lenguaje que han hecho uso de esas herramientas.

2. Optimización, comportamiento estratégico, y conducta guiada por normas

¿Qué hace que un determinado modo de conducta pueda considerarse «racional»? Según la concepción «estándar» de la racionalidad (o, al menos, de lo que los filósofos suelen denominar «racionalidad instrumental», o «racionalidad de medios»), una decisión tomada por un individuo será racional cuando sea la mejor decisión que podría haber tomado, teniendo en cuenta tanto *sus* intereses o preferencias, como *su* información (es decir, sus creencias) sobre las probabilidades de obtener unos u otros resultados con cada posible opción alternativa entre las que podía elegir. Otra forma de decir esto mismo es que la decisión de un

2 Así, en la teoría «clásica» de los actos de habla (cf. Austin (1982) y Searle (1986)).

individuo será racional cuando, una vez tomada, este individuo no pueda encontrar *razones* que le lleven a arrepentirse. Un conocido teorema matemático afirma que, si un individuo tiene preferencias «consistentes» ante un conjunto de opciones alternativas cuyos resultados son inseguros,³ entonces esas preferencias podrán ser representadas mediante una función numérica (denominada «función de utilidad»), y la conducta racional será equivalente a aquella en la que se elige aquella opción que proporciona el valor más alto de la esperanza matemática de dicha función (es decir, se maximiza la «utilidad esperada»). Dicho más formalmente, si $U(x)$ es la utilidad que el individuo recibe si obtiene el resultado x , y $p(x/a)$ es la probabilidad con la que cree que se obtendrá ese resultado si él lleva a cabo la acción a , entonces la utilidad esperada asociada a esa acción será:

$$(1) \quad UE(a) = p(x_1/c)U(x_1) + \dots + p(x_n/c)U(x_n) = \sum_{(1 \leq i \leq n)} p(x_i/c)U(x_i)$$

De acuerdo con el teorema citado, un individuo se comportará racionalmente si y sólo si su conducta es coherente con la maximización de su función de utilidad esperada. Naturalmente, cuando cada acción conduce a un resultado *seguro*, se puede prescindir de las probabilidades, y afirmar que una acción es racional si y sólo si es la que conduce al resultado cuya utilidad es más alta de entre todos los posibles. Es importante notar, también, que el hecho de que las preferencias deban representarse mediante una función *numérica* (la función de utilidad) no es una *limitación* de la teoría, pues no se trata de que la teoría sólo sea aplicable a aquellos casos en los que está dada dicha representación matemática: lo que el teorema demuestra es, más bien, que, *si* la acción es racional, *entonces* las preferencias se podrán representar de ese modo. De hecho, en todos aquellos casos en los que las probabilidades no intervienen de forma específica, la explicación de una acción mediante la teoría de la racionalidad no requiere una presentación matemática en absoluto: basta con indicar por qué unas opciones son mejores que otras para mostrar por qué la elección racional es la que opta por la mejor alternativa posible. Esto, sin duda, es lo que hacen implícitamente muchos argumentos empleados en la literatura sobre pragmática, al indicar por qué ciertas formas de expresión son más «informativas», o «relevantes», o «cortes», o por qué en ciertos la violación de las normas de cooperación lingüística es más útil para el hablante que lo contrario. Aquí estaríamos hablando, entonces, de una especie de uso «intuitivo» de la teoría de la elección racional; lo que podría lograrse con una aplicación más explícita de la teoría a todos esos casos sería, sobre todo, que con ello nos obligáramos a reflexionar con más detalle sobre las formas de expresión alternativas entre las que el hablante podría haber escogido, y sobre las razones específicas por las que cada una de estas expresiones es «peor» que la efectivamente elegida.

Otro aspecto interesante de esta teoría es que *no* exige (contra lo que muchos críticos del enfoque «racional» suponen) que el *proceso* de toma de decisiones de los individuos *consista* de hecho en el cálculo (¿«mental»?) de los valores de la función UE. Lo único que afirma la teoría es que, si la conducta de un individuo *no* es coherente con ella, entonces

3 Las «preferencias consistentes» son definidas mediante una lista de requisitos, todos ellos más o menos razonables, como, p. ej., que las preferencias sean transitivas (si prefiero a a b , y b a c , debo preferir a a c), o que no me hagan asumir riesgos innecesarios (si a y b sólo se diferencian en que, con la primera opción, el resultado mejor es más probable que con la segunda, y el resultado peor es menos probable, entonces preferiré a a b).

podría haber elegido alguna opción mejor de acuerdo con sus propias valoraciones, sean éstas las que sean, y con independencia de cuáles hayan sido los mecanismos «psicológicos» que le hayan llevado a tomar la decisión en cuestión. Esto se puede captar de forma muy ilustrativa al saber que la teoría de la «elección» racional puede ser aplicada con gran éxito a muchos procesos de *optimización* llevados a cabo por «agentes» no humanos, e incluso no sentientes en absoluto (animales, plantas, autómatas, etc.). *Lo único necesario es que seamos capaces de describir el conjunto de opciones del agente, los resultados a los que éstas puedan conducir (si acaso, según ciertas probabilidades), y una medida razonable del grado de «utilidad» o «satisfacción» que el agente obtendría con cada resultado posible.*⁴

Ahora bien, lo más habitual en las situaciones a las que se enfrentan los seres humanos es que los resultados obtenidos por ellos no dependan únicamente de sus propias decisiones (y del «estado de la naturaleza»), sino también de las acciones llevadas a cabo por otras personas. A una situación como esta, en la que las opciones de cada individuo están exhaustivamente descritas, así como la valoración que da cada uno al resultado que se obtiene con cualquier combinación posible de una opción por cada individuo (es decir, lo que yo obtengo si tú haces A y yo X, lo que obtengo si tú haces B y yo X, lo que obtengo si tú haces A y yo Z, etc., y lo mismo para lo que obtienes tú en cada caso); a una situación así, decía, le llaman los economistas un «juego», y la parte de la teoría económica que se dedica a estudiar las propiedades de estas situaciones de mutua interdependencia se la conoce como «teoría de juegos». Encontrarse en una situación así exige que, para elegir racionalmente mi opción, yo tenga que prever la decisión que vas a tomar tú (imaginándonos una situación en la que sólo intervienen dos individuos), pero la cosa se complica, naturalmente, porque tú mismo te basarás, para tomar tu decisión, en la previsión que tú hagas acerca de la decisión que tomaré yo. En una situación como esta, los individuos se ven obligados a tomar una decisión *estratégica*, es decir, en la que se tienen en cuenta las deliberaciones que los demás pueden hacer sobre nuestras decisiones antes de tomar las suyas. La cuestión obvia es que esto parece conducir a un círculo vicioso: puesto que, para actuar racionalmente, yo necesito saber qué vas a hacer tú antes de decidir qué voy a hacer yo, y tú necesitas saber qué voy a hacer yo antes de decidir lo que vas a hacer tú, parece que la acción *racional* no podría comenzar nunca. El principio económico que resuelve este conflicto es el denominado «*equilibrio de Nash*»; los individuos que se enfrentan a una situación de interdependencia están en «equilibrio» cuando cada uno de ellos ha elegido una opción que cumple la propiedad siguiente: dadas las opciones elegidas por los demás, a ninguno de ellos le interesa elegir una distinta de la que ha él ha elegido. Es decir, una combinación de acciones constituye un equilibrio en el juego correspondiente cuando cada «jugador» está haciendo lo mejor que puede, teniendo en cuenta lo que están haciendo los demás.

Representar una situación de interacción entre dos individuos como un juego es relativamente fácil: ponemos en una tabla las opciones o estrategias entre las que puede elegir

4 Por supuesto, es muy conveniente disponer también de alguna *explicación* del mecanismo o los mecanismos (conscientes o no) por los que las opciones que conducen por término medio a resultados peores son generalmente descartadas; pero la actitud general de los defensores de la teoría de la elección racional es que la búsqueda de ese mecanismo es una tarea secundaria, que únicamente será relevante *después* de que hayamos mostrado que el fenómeno que estamos estudiando se puede explicar adecuadamente mediante alguna hipótesis de optimización. La exposición clásica de este argumento es la de Friedman (1967; ed. orig. 1953).

uno de ellos (A) en las filas, y las del otro (B) en las columnas (no es necesario que las opciones de los dos sean idénticas, ni que coincidan en número); posteriormente, en cada casilla de la tabla así construída, ponemos la valoración (o «utilidad») que cada individuo da al resultado que se obtendría si cada uno de ellos eligiera la opción correspondiente a la fila y a la columna de esa casilla (distinguiendo claramente las valoraciones de A y de B; p. ej., las de A abajo a la izquierda, y las de B arriba a la derecha). Finalmente, buscamos el equilibrio, o los equilibrios, del juego de la manera siguiente: para cada columna, señalamos cuál es la casilla más valorada por el jugador B (es decir, qué preferiría él hacer si A eligiera la acción correspondiente a esa columna), y para cada fila, señalamos la casilla más valorada por A. Aquellas casillas que estén marcadas tanto para A como para B, serán los equilibrios de Nash. (Cf. tabla 1, donde el equilibrio es la casilla correspondiente a la situación en la que A elige I y B elige Z; por ejemplo, ¿por qué no es la primera casilla un equilibrio: porque, aunque es cierto que si B eligiera X, A elegiría I, ocurre que, si A eligiera I, B no elegiría X, sino Z).

| Jugadores B A | X | Y | Z |
|---------------|------------|------------|----------------|
| I | 3 4 (*) | 4 0 | (*) 5 4 (*) |
| II | (*) 4 3 | 0 4 (*) | 2 1 |

Tabla 1

Representar una situación de interacción mediante un juego requiere, por lo tanto, lo siguiente: *describir las opciones de cada individuo, asignar una medida razonable de utilidad a cada individuo en cada resultado posible del juego, y buscar los equilibrios de Nash (o «soluciones» del juego).*

Otra forma de representar y «resolver» un juego es mediante la denominada «forma normal» (la anterior sería la «forma estratégica»), que es especialmente útil cuando, como en el caso de la comunicación humana, los individuos toman sus decisiones uno después del otro. En este caso, se debe dibujar un árbol en el que, de izquierda a derecha, cada nodo representa la decisión de un jugador, y cada línea que conecta con un nodo con los que hay inmediatamente a su derecha representa cada una de las opciones entre las que el jugador puede elegir en ese nodo (cf. fig. 1). Cada uno de los últimos nodos de la figura representa la utilidad que obtendría cada individuo si se siguiera exactamente la serie de líneas (decisiones) que conectan el nodo original del juego con ese nodo final en concreto. Este tipo de juegos se resuelven habitualmente mediante el procedimiento llamado «inducción inversa»: en cada *penúltimo* nodo vemos qué decisión será la más racional para el individuo al que le toque jugar en ese nodo, y tachamos todas las demás líneas que salen de ese penúltimo nodo, con lo que se convierte en un nodo «último», y repetimos así el proceso hasta llegar al primer nodo; la serie de líneas que quede sin tachar, será la solución del juego. En la figura 1 hemos supuesto, para simplificar, que cada jugador obtiene siempre la misma satisfacción que el otro (es lo que se llama un juego de cooperación pura); cuando no sea así, hay que

tener cuidado de considerar, en cada nodo, la utilidad que recibe el individuo al que le toca jugar ahí, y no la del otro jugador. La solución del juego representado en la figura será, obviamente, la situación en la que A elige «arriba» en el nodo 1, y B elige también «arriba» en el nodo 2 (en el nodo 3 elegiría «abajo», pero a ese nodo no se llega, debido a la decisión que toma A en el primero).

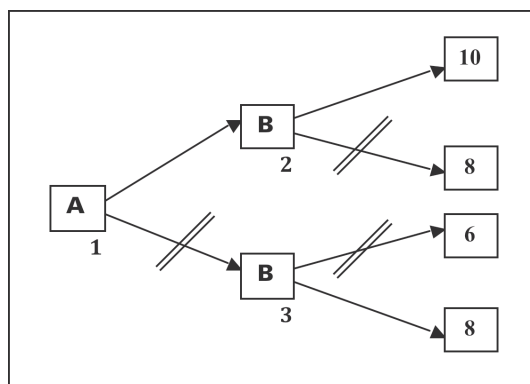


Figura 1

Los principales problemas que podemos encontrar a la hora de «resolver» un juego (una vez que hemos conseguido describirlo) son los siguientes: en primer lugar, *puede que no encontremos ningún equilibrio*; en este caso, un teorema del propio John Nash garantiza que, si ampliamos el conjunto de decisiones que cada jugador puede tomar, considerando también que entre sus opciones está el decidir unas veces una opción y otras veces otras opciones, con una determinada probabilidad cada una de ellas, entonces siempre hay algún equilibrio, aunque puede corresponder con un caso en el que cada jugador no elige una opción directamente, sino una cierta probabilidad de tomar cada decisión. En segundo lugar, *puede que encontremos más de un equilibrio*, en cuyo caso tendremos que buscar argumentos adicionales para decidir por qué uno de los equilibrios es más «razonable» que los otros, o para decidir que todos ellos son igual de probables; de hecho, la existencia de diversas lenguas o de diversos códigos lingüísticos en diferentes grupos o épocas puede interpretarse como la selección de varias soluciones diferentes de un mismo juego. Finalmente, *puede que la experiencia nos diga que la gente no suele comportarse de la manera como la solución de nuestro juego indica*. En este caso tenemos dos estrategias posibles: intentar redefinir el juego (añadiendo opciones, o modificando las valoraciones), pues posiblemente no hayamos sabido captar verdaderamente la situación a la que los agentes *creen* que se están enfrentando; o bien buscar razones que «expliquen» la desviación en esos casos. De todas formas, en general podemos afirmar que, el que una situación sea un equilibrio, es una condición *necesaria* (aunque no suficiente) para que la situación sea *estable*; esto quiere decir que, si reiteradamente observamos que se da una situación que claramente es de «desequilibrio» (es decir, en la que alguien no está haciendo lo mejor que puede, lo que más le interesa, cuando sabe lo que están haciendo los demás), entonces seguramente podremos encontrar algún tipo

de fuerzas que «empujan» a los individuos fuera de su opción más «racional»; por supuesto, identificar claramente esas fuerzas es ya un logro analítico importante.

Finalmente, un par de indicaciones sobre la posible coherencia entre los modelos de decisión racional que acabamos de describir, y la idea de que los seres humanos actúan a menudo basándose en *normas*, más que en el «cálculo» de sus intereses. En primer lugar, algunas veces podemos interpretar las opciones entre las que puede optar cada individuo como «seguir la norma X», «seguir la norma Y», etc. Esto, que será ciertamente válido en ciertos contextos, deja sin explicar muy claramente la posibilidad de que, en otros momentos, el individuo decida «violiar» la regla (pues, de ser así, tendríamos que haber establecido al principio esta posibilidad al describir sus opciones). De todas formas, es habitual suponer que los «jugadores» no son totalmente racionales, en el sentido de que prefieren (o, por sus limitaciones cognitivas, no tienen más remedio que) seguir pautas fijas, y entonces los equilibrios se alcanzan gracias a una dinámica de selección más o menos darwiniana, en la que las pautas o normas más exitosas son «imitadas» cada vez por más individuos. La llamada «teoría de juegos evolutivos» permite, entonces, ofrecer una explicación del surgimiento de ciertos sistemas de normas en vez de otros (v., p. ej., Maynard Smith (1982)).

En segundo lugar, en la medida en que una determinada conducta basada en normas sea habitual, la hipótesis de que los individuos son racionales nos lleva a concluir que, si en ese caso uno de ellos estuviera obteniendo un resultado claramente peor para él que el que obtendría modificando su conducta de cierta manera, entonces ese individuo dejaría de seguir esa norma. Dicho de otra manera, *las normas que terminan imponiéndose en una situación de interacción reiterada entre los mismos individuos* (aunque no necesariamente los mismos jugando con los mismos cada vez), *serán normas que permitan que se dé un equilibrio de Nash*. Esto es lo que sucede, generalmente, en los llamados «juegos de coordinación», donde lo importante para cada individuo no es realizar una acción u otra, sino realizar exactamente aquella que está coordinada con las acciones de los demás. En estos juegos suelen existir varios equilibrios (pensar, p. ej., en el «juego» que consiste en elegir un lado de la carretera para circular habitualmente por él con el coche). De hecho, uno de los primeros análisis de estas situaciones en términos de teoría de juegos fue el libro de David Lewis, *Conventions* (1969), que luego ha sido, como veremos, muy importante en la aplicación de esta perspectiva a la explicación de las normas lingüísticas, pues, especialmente la asignación de un significado determinado a un signo lingüístico es (en general) un asunto puramente convencional, en el que lo importante para los hablantes no es cuál es el significado «razonable» del signo, sino dar a cada signo la misma interpretación que le dan los demás, *sea ésta la que sea*.

3. Algunas aplicaciones a la lingüística en general, y a la pragmática en particular

La comunicación entre seres humanos es, claramente, un fenómeno de carácter intencional, en el que los participantes intentan obtener ciertos resultados, de tal manera que parece razonable, en principio, intentar analizarlo bajo el supuesto de que cada individuo realizará aquella acción que le parezca más satisfactoria. El hecho de que estas acciones se lleven a cabo de manera casi automática no significa, por supuesto, que los mecanismos cognitivos subyacentes no estén programados de forma que lleven los individuos a tomar

decisiones relativamente óptimas. Esta sola consideración ya hace lógico, por sí misma, el intentar encontrar sistemas de *optimización* en los procesos de producción e interpretación de mensajes lingüísticos. Por otro lado, la comunicación es un proceso en el que interviene por lo general más de un individuo, y las acciones que cada uno lleva a cabo en ese proceso también afectan, evidentemente, al resultado de todos los demás; por tanto, parece lógico asimismo que se intente representar las situaciones de interacción comunicativa mediante las herramientas formales de los «juegos» que hemos descrito en el apartado anterior.

Ciertamente, como indicaba al principio, no es mucha la literatura que ha utilizado explícitamente estas técnicas en los diversos campos de la lingüística. El ámbito en el que son más claramente de aplicación es el de la *sociolingüística*, rama esta que, como su nombre indica, está en la frontera de la lingüística y las ciencias sociales; pero aquí voy a referirme específicamente a los territorios más centrales de la disciplina, como son la *sintaxis*, la *semántica* y la *pragmática* (permítaseme indicar, de pasada, que aunque esta última rama ha sido considerada tradicionalmente como menos «central», la situación ha cambiado mucho al respecto en el último cuarto de siglo, llegando incluso algunos autores a considerarla como una parte más fundamental para la comprensión de la estructura básica del lenguaje, ya que éste es, por esencia, un fenómeno *intencional*; la aplicación de la teoría de la elección racional y la teoría de juegos a la sintaxis y a la semántica es coherente con esta consideración, pues tiende a mostrar que las características *estructurales* del lenguaje se deben de forma primordial a las necesidades de su *uso*). Otras aplicaciones del pensamiento «económico» a la lingüística, aparte de las mencionadas, se refieren a las cuestiones sobre el *origen* y la *evolución* del lenguaje; v. p. ej., Zipf (1949), Keller (1994), Croft (2000), o Noble (2000).

Por falta de tiempo y espacio, no voy a glosar aquí todos y cada uno de los trabajos en los que las herramientas de análisis económico han sido utilizadas en la lingüística (en muchos casos, de forma más bien metafórica), sino que me limitaré a remitir a los lectores a la bibliografía (que tampoco es exhaustiva), e indicaré aquí las dos líneas de investigación que me parecen más prometedoras. Por una parte, está la denominada «*teoría de la optimalidad*» («*optimality theory*»; v. p. ej., Archangeli y Langendoen (1997) y Kager (1999); Hendriks y de Hoop (2001) ofrecen una aplicación a la semántica). Según esta teoría, las reglas del lenguaje (p. ej., las reglas de transformación de la gramática generativa) no son algoritmos que generen automáticamente un único posible enunciado con sentido semántico bien definido para cada mensaje que se quiera transmitir, sino que se limitan a imponer *restricciones* («*constraints*») a los posibles mensajes; el enunciado que se considerará «gramatical» será, en cada caso, aquel que menos de estas restricciones viole, o el que viole preferentemente las restricciones menos importantes (una idea interesante de este enfoque es que las restricciones serían universales, mientras que cada lenguaje ordenaría el grado de importancia de estas restricciones de forma diferente a los demás). Así, cada posible construcción lingüística va asociada con un cierto *valor* (para un mensaje o contenido dado), y se trata de elegir aquella construcción cuyo valor es más alto.

Desde el lado de la pragmática, se han intentado en particular ofrecer una explicación en términos de teoría de juegos de fenómenos o principios como las implicaturas, las máximas de cooperación, la argumentación o la cortesía (v. p. ej., Lewis (1979), Merin (1994), y Parikh (2000) y (2001)). La idea principal es que el *hablante*, al intentar transmitir un determinado mensaje, elegirá una forma lingüística u otra teniendo en cuenta la decisión que

el oyente debe tomar después sobre si interpretar de una manera u otra la forma finalmente emitida por el hablante. P. ej., en la figura 1 podemos interpretar el problema de A como el de decidir si transmitir un mensaje ambiguo o uno no ambiguo, y el de B como el de decidir si el mensaje que se le ha transmitido debe ser interpretado de manera literal o no literal. Simplemente a modo de ilustración precipitada,⁵ imaginemos que A es un explorador del salvaje Oeste que informa de la presencia de indios en las cercanías del fuerte; él sabe que son exactamente 1.962 indios, pero tiene que transmitir su mensaje muy deprisa porque los trae pisándole los talones; puede gritar «¡vienen dos mil indios!» -decisión «arriba» en el nodo 1 de la figura- o «¡vienen mil novecientos sesenta y dos indios!» -decisión «abajo»- (también podría decir, «vienen unos dos mil indios», usando una expresión que indica a los oyentes que está haciendo un uso no literal de los numerales, pero por simplificar ignoraremos esa posibilidad; también podría decir, como en el chiste, «vienen unos dos mil tres indios», teniendo en cuenta que vienen tres delante y unos dos mil detrás; en fin, también podría callarse y seguir corriendo más allá del fuerte...). En los nodos segundo y tercero, B, el comandante del fuerte, debe decidir si lo que ha dicho A debe interpretarse de forma aproximada («arriba») o exacta («abajo»). Jugando nosotros mismos a asignar utilidades más o menos razonables a cada nodo final, podremos buscar el equilibrio del juego (que no necesariamente coincidirá con el de la figura), y ofrecer así una explicación de por qué suelen utilizarse ciertas formas lingüísticas para expresar ciertos mensajes (aunque no el significado literal de las primeras no coincida con el contenido semántico de lo que se quiere transmitir).

La idea ha sido generalizada por los defensores de la llamada «teoría de la optimalidad bi-direccional» (v., p. ej., Blutner (2000) y van Rooy (2003)) de la manera siguiente: cada «acto lingüístico» básico es la elección conjunta de un mensaje o significado (m) y un enunciado o forma de expresión (f); un par $\langle f, m \rangle$ debe ser óptimo en el sentido de satisfacer simultáneamente el «principio del hablante» y el «principio del oyente», a saber: para cualquier otra forma alternativa f' , el par $\langle f', m \rangle$ sería *peor desde el punto de vista del hablante* que el par $\langle f, m \rangle$, y a la vez, para cualquier otro significado o interpretación alternativa, m' , el par $\langle f, m' \rangle$ sería *peor desde el punto de vista del oyente* que el par $\langle f, m \rangle$. Como señala van Rooy, la exigencia de que se cumplan simultáneamente estos dos principios es equivalente a exigir que el par $\langle f, m \rangle$ elegido constituya un equilibrio de Nash.

Para concluir, mencionaré algunas contribuciones que se han hecho al análisis del lenguaje mediante la teoría de juegos desde fuera del terreno académico de la lingüística. Básicamente podemos citar, por un lado, la obra de algunos economistas, que han llegado a este tema preocupados por analizar el proceso de *comunicación* que debe darse entre los jugadores de cualquier juego para que cada uno de ellos pueda conocer el punto de vista desde el que el otro individuo percibe el juego (lo que los teóricos suelen denominar «*cheap talk*»). Aquí la obra principal es Rubinstein (2000); v. tb., Marschak (1965), Kreps y Sobell (1994), Kim y Sobel (1995), o Warneryd (1993). Otro terreno desde el que se ha aplicado la teoría de juegos para intentar comprender ciertos aspectos del lenguaje es la semántica formal (entendida como una rama de la lógica, más que de la lingüística); la principal contribución es la llamada «*game-theoretical semantics*», desarrollada sobre todo por Jaakko Hintikka; v., p. ej., Hintikka (1973) y Saarinen (1979).

5 En Parikh (2000) y (2001) se ofrecen ejemplos mucho más razonados y detallados.

Bibliografía

- Archangeli, D., y D. T. Langendoen (eds.), 1997, *Optimality Theory – an Overview*, Oxford, Blackwell.
- Asher, N., I. Sher, y M. Williams, 2001, «Game theoretical foundations for Gricean constraints», In R. van Rooy & M. Stokhof (eds.), *Proceedings of the Thirteenth Amsterdam Colloquium*, ILLC, Amsterdam.
- Austin, J. L., 1982, *Cómo hacer cosas con palabras*, Barcelona, Paidós (ed. orig. 1960).
- Beaver, D., 2003, «The optimization of discourse». *Linguistics and Philosophy* (en prensa).
- Blutner, R., 1998, «Lexical pragmatics», *Journal of Semantics* 15, 115-162.
- Blutner, R., 2000, «Some aspects of optimality in natural language interpretation», *Journal of Semantics*, 17, 189-216.
- Del Corral, M., y J.P. Zamora Bonilla, 2008, «*Also Sprach der Homo Oeconomicus*», *Journal of Economic Literature*, 15, 241-44. (Introducción al número monográfico «Language and Games»).
- Croft, W., 2000, *Explaining Language Change: An Evolutionary Approach*, Longman Linguistics Library, Harlow.
- Desalles, J. L., 1998, «Altruism, status and the origin of relevance», en J. R. Hurford *et al.* (eds.), *Approaches to the Evolution of Language: Social and cognitive bases*, Cambridge, Cambridge University Press.
- Fox, D., 2000, *Economy and semantic interpretation*, Cambridge (Mass.), The MIT Press.
- Friedman, M., 1967, *Ensayos de economía positiva*, Madrid, Gredos.
- Hasida, K., K. Nogao and T. Miyata, 1995, 'A game-theoretic account of collaboration in communication', en *Proceedings of the First International Conference on Multi-Agent Systems*, San Francisco.
- Henderson, W., T. Dudley-Evans, y R. Backhouse (eds.), 1993, *Economics and Language*, London, Routledge.
- Hendriks, P. y H. de Hoop, 2001, «Optimality Theoretic Semantics», *Linguistics and Philosophy*, 24: 1-32.
- Hintikka, J., 1976, *Lógica, juegos de lenguaje, e información*, Madrid, Tecnos.
- Hurd, P. L., 1995, «Communication in discrete action-response games», *Journal of Theoretical Biology*, 174, 217-222.
- Jaeger, G., 2000, «Some notes on the formal properties of bidirectional optimality theory», en R. Blutner & G. Jaeger (eds.), *Studies in Optimality Theory*, Linguistics in Potsdam, volume 8, pp. 41-63.
- Kager, R., 1999, *Optimality theory*, Cambridge, Cambridge University Press.
- Keller, R., 1994, *On Language Change. The Invisible Hand in Language*, London, Routledge.
- Kim, Y. G., y J. Sobel 1995, «An evolutionary approach to pre-play communication», *Econometrica*, 63, 1181-93.
- Krebs, J. R., y R. Dawkins, 1984, «Animal signals: mind reading and manipulation», en J. R. Krebs y N. B. Davies (eds.), *Behavioural Ecology: An evolutionary approach*, Oxford, Blackwell.
- Kreps, D., y J. Sobel, 1994, «Signalling», en R.J. Aumann & S. Hart (eds.), *Handbook of Game Theory*, Vol. 2, Elsevier.

- Lewis, D., 1969, *Convention*, Cambridge: Harvard University Press.
- Lewis, D., 1979, «Scorekeeping in a language game», *Journal of Philosophical Logic*, 8, 339-359.
- Marschak, J., 1965, «The economics of language», *Behavioral Science*, 10, 135-40.
- Maynard Smith, J., 1982, *Evolution and the Theory of Games*, Cambridge, Cambridge University Press.
- Maynard Smith, J., y D. G. C. Hasrper, 1995, «Animal signals: models and terminology», *Journal of Theoretical Biology*, 177, 305-311.
- Merin, A., 1994, «The algebra of elementary social acts», en S. Tsohatzidis (ed.), *Foundations of Speech Act Theory*, London, Routledge.
- Merin, A., 1999, «Information, relevance, and social decisionmaking», *Logic, Language, and Computation*, Vol. 2, L. Moss, J. Ginzburg, M. de Rijke (eds.), CSLI publications, Stanford, 179-221.
- Moreno Cabrera, J. C., 2002, «El motor de la economía lingüística: de la ley del mínimo esfuerzo al principio de la automatización retroactiva», *Revista Española de Lingüística*, 32.1, 1-32.
- Newmeyer, F., 1991, «Functional explanation in linguistics and the origin of language», *Language and Communication*, 11, 67-69.
- Noble, J., 2000, «Cooperation, competition and the evolution of prelinguistic communication», en C. Knight *et al.* (eds.), *The Emergence of Language*, pp. 40-61, Cambridge University Press.
- Oliphant, M., 1996, «The dilemma of Saussurean communication», *BioSystems*, 37, 31-38.
- Parikh, P., 1991, «Communication and strategic inference», *Linguistics and Philosophy*, 14: 473-513.
- Parikh, P., 1992, «A game-theoretical account of implicature», en Y. Vardi (ed.), *Proceedings of the fourth TARK conference*, Monterey, California.
- Parikh, P., 2000, «Communication, meaning, and interpretation», *Linguistics and Philosophy*, 23: 185-212.
- Parikh, P., 2001, *The use of Language*, CSLI Publications, Stanford, California.
- Prince, A., y P. Smolensky, 1997, «Optimality: From Neural Networks to Universal Grammar», *Science*, 275, 1604-1610.
- Rooy, R. van, 2002, «Utility, informativity, and protocols», en Bonanno *et al.* (eds.), *Proceedings of LOFT 5: Logic and the Foundations of the Theory of Games and Decisions*, Torino.
- Rooy, R. van, 2003, «Signaling games select Horn strategies», *Linguistics and Philosophy* (en prensa); puede consultarse por internet en la página *web* del autor.
- Rubinstein, A., 2000, *Economics and Language*, Cambridge University Press.
- Rundalab, G., y D. B. Kronenfeld, 2003, The inevitability of folk etymology: a case of collective reality and invisible hands, *Journal of Pragmatics*, 35, 119-138.
- Saarinen, E., (ed.), 1979, *Game-Theoretical Semantics*, Dordrecht, D. Reidel.
- Searle, J., 1986, *Actos de habla*, Madrid, Cátedra (ed. Orig., 1969).
- Warneryd, K., 1993, «Cheap talk, coordination and economic stability», *Games an economic behaviour*, 5, 532-46.
- Zipf, G., 1949, *Human behavior and the principle of least effort*, Cambridge, Addison-Wesley.