

# LA TEORIA SEMANTICA DE LAS LOGICAS DE CONDICIONALES DE STALNAKER Y LEWIS

Mario Gómez Torrente

## Abstract

This paper expounds chronologically the development of the theories of conditionals by Stalnaker and Lewis, paying attention especially to the semantic reasons behind the successive amendments and refinements of both theories concerning the linguistic behavior of subjunctive conditionals. At the end, it is discussed carefully how these different semantic backgrounds may influence the interpretation of indicative conditionals, and Stalnaker's view on these conditionals is advocated with the aid of some new example.

---

Una de las primeras perplejidades con que se encuentra quien estudia lógica la produce la semántica del condicional material. Uno de los ejemplos de algunos manuales para ilustrar esta semántica parece buscado para suscitar esa perplejidad, y no precisamente para iluminarla: 'si llueve, no me mojo' es verdadero, se dice, tanto si no llueve como si no me mojo. Esto suena bastante extraño, y hace sospechar de la adecuación de la semántica del condicional material para tratar todo tipo de casos de construcción condicional; el que no llueva –podría razonarse– no quiere decir que no me haya de mojar si llueve. Hay casos de construcciones en que el conflicto con la intuición es todavía más agudo si intentamos verlos como condicionales materiales: aunque Santo Tomás no fue un pecador, no asentaríamos a quien sostuviera que si hubiera sido un pecador habría sido (de todos modos) canonizado.

Si bien para algunos usos o propósitos la semántica veritativo-funcional del condicional material es apropiada para tratar las construcciones condicionales, esto no ocurre siempre; es cierto que si dos más dos son cinco hay un número finito de primos; pero no es cierto (con toda probabilidad) que si Santo Tomás hubiera sido un pecador Judas Iscariote habría sido canonizado. Algunos casos de condicionales y de inferencias con condicionales requieren, al parecer, un tratamiento sensible a otras propiedades semánticas de los enunciados componentes que el valor de verdad; la semántica de las construcciones condicionales del lenguaje natural no es en general veritativo-funcional.

La simple aceptación de la posibilidad de que existan distintos tipos de enun-

ciados condicionales, que respondan a estructuras semánticas diferentes, es muy reciente. Es por ello que, históricamente, el problema se planteó de un modo muy crud o, como el problema de hallar la «auténtica» naturaleza del condicional; hubo una discusión apasionada en la antigüedad entre los partidarios de la implicación material (el más destacado de los cuales fue Filón de Megara) y los de la implicación estricta (a cuya cabeza estuvo Diodoro Cronos); estos últimos sostenían que la verdad de un condicional suponía la existencia de algún tipo de conexión necesaria entre la verdad del antecedente y la del consecuente. A principios de este siglo, C. I. Lewis construyó un cálculo de implicación estricta con el que pretendía capturar el significado intuitivo de las construcciones condicionales en general, con ayuda de conceptos modales;  $A \rightarrow B$  (leído 'A implica estrictamente B') es verdadero en el cálculo de Lewis si y sólo si  $\Box (A \supset B)$  es verdadero, esto es, si y sólo si es necesariamente verdadero el condicional material  $A \supset B$ . Por razones que veremos más adelante, el análisis primitivo de C. I. Lewis es insatisfactorio como tratamiento de las construcciones condicionales en general.

A finales de los años treinta, se suscitó un interés especial por los llamados 'condicionales contrafácticos', condicionales con antecedente falso (normalmente, también lo es su consecuente) y generalmente en tiempo verbal subjuntivo: 'si yo fuera rico, no trabajaría', 'si Heidegger no hubiese escrito nada, la filosofía habría muerto'. Los filósofos positivistas observaron que los términos disposicionales que abundan en los textos científicos y que se utilizaban en la exposición de sus propias tesis filosóficas podían parafrasearse a través de construcciones contrafácticas: una sustancia es soluble en un líquido si y sólo si se disolvería si fuese vertida en el líquido; un enunciado es verificable si y sólo si si se llevasen a cabo ciertos procedimientos de contrastación empírica, estos procedimientos determinarían su verdad o falsedad. Tratar estos condicionales como materiales supondría aceptar que toda sustancia no vertida en un líquido es soluble en él, y que todo enunciado jamás sometido a contrastación empírica es verificable. Así, ciertas afirmaciones de la ciencia y de la filosofía (científica) parecían tener una lógica no veritativo-funcional, y esto constituía un problema para el proyecto positivista de formalización de la ciencia. Por otro lado, la alusión de las paráfrasis a posibles no actualizados no era menos problemática desde un punto de vista empirista. Los artículos clásicos de Chisholm y Goodman sobre condicionales contrafácticos son intentos de conseguir paráfrasis de estos enunciados aceptables desde un punto de vista filosófico empirista y que sólo involucren condicionales materiales. Las observaciones de Quine y, nuevamente, de C. I. Lewis, fueron también de importancia en estos desarrollos.

La aparición de la semántica de mundos posibles (y el progresivo relajamiento de las tesis analíticas sobre lo que es filosóficamente aceptable) estimuló un nuevo modo de enfrentarse a la semántica de las construcciones condicionales a partir de conceptos modales. La primera de las teorías de condicionales en el marco de la semántica de mundos posibles fue la de Robert Stalnaker, expuesta por él en 1968 y desarrollada en sus aspectos técnicos por Stalnaker y Richmond Thomason en 1970. La semántica de David Lewis para contrafácticos, expuesta por primera vez en 1971, se apoya en ideas similares a las de Stalnaker. A partir de estos artículos pioneros, varias «lógicas de condicionales» inspiradas en la semántica de mundos

posibles han sido propuestas en las dos últimas décadas. Este tipo de semánticas, cuyo núcleo es lo que se ha dado en llamar 'teoría de Stalnaker-Lewis' han clarificado en un grado considerable la naturaleza semántica de los distintos tipos de construcciones condicionales, y hoy son instrumentos usuales en su discusión<sup>1</sup>. La comprensión proporcionada por estos estudios ha facilitado también el desarrollo de nuevas aproximaciones a distintos problemas filosóficos; entre ellas se hallan el análisis de la noción de causalidad por Lewis, diversos tratamientos de algunas paradojas de la teoría de la decisión, discusiones sobre la noción de probabilidad objetiva, etc.

En las páginas que siguen se exponen en primer lugar las características principales de las teorías originales de Stalnaker y Lewis, y a continuación algunas de las divergencias entre ambas teorías, en concreto las diferencias en cuanto a la validez del llamado 'principio de tercio excluso condicional' y en cuanto a la tesis conocida como 'suposición del límite'. En un último apartado, se habla con algún detalle de las diferencias de Stalnaker y Lewis en la consideración de los condicionales indicativos (los condicionales con el verbo del antecedente en indicativo); creo que este apartado puede ser útil para quien desee una introducción clara o una noticia breve de este aspecto de la teoría de condicionales, sobre el que la bibliografía secundaria es todavía escasísima en cualquier lengua<sup>2</sup>.

## La teoría de Stalnaker

En su artículo de 1968 «A Theory of conditionals», Stalnaker partió de una observación hecha por F. P. Ramsey. Ramsey había sugerido<sup>3</sup> que solemos realizar la evaluación del valor de verdad de un enunciado condicional haciendo el experimento mental siguiente: añadimos hipotéticamente a nuestras creencias el antecedente del enunciado; para ello modificamos nuestro acopio de creencias anteriores, retirando momentáneamente las que contradecirían el antecedente añadido como nueva creencia; tras hacer las mínimas revisiones necesarias de este tipo, examinamos si el consecuente se sigue de nuestro acopio de creencias modificado. Si es así,

<sup>1</sup> Son muy importantes también ciertos desarrollos que usan la teoría de la probabilidad para clarificar la semántica de los condicionales, y que tienen su origen sobre todo en la obra de Ernest Adams. Tendremos ocasión de mencionar algunas de las interrelaciones entre estos desarrollos y las semánticas para condicionales de Stalnaker y Lewis al comentar la obra de éstos.

<sup>2</sup> En lo que sigue se usarán los símbolos lógicos como nombres de sí mismos; una expresión en que aparezcan variables metalingüísticas para enunciados junto con símbolos lógicos se entenderá como un sinónimo de 'el resultado de concatenar las expresiones nombradas en el orden que se nombran en la expresión'. Así, no se usan —por comodidad— *corners* ni artificios similares; por ejemplo, ' $A \supset B$ ' será un sinónimo de 'el resultado de concatenar  $A$  con  $\supset$  con  $B$ '; se entenderá que las variables metalingüísticas están cuantificadas universalmente (en el metalenguaje), salvo cuando se explicita lo contrario. Se emplearán ciertos símbolos como abreviaturas para las supuestas conectivas de los diferentes tipos de construcción condicional; así,  $A \supset B$  será un condicional material,  $A \Box \rightarrow B$  será un condicional contrafáctico,  $A \rightarrow B$  será un condicional indicativo, y  $A > B$  será un condicional de Stalnaker; esta práctica es común (cfr. por ejemplo F. Jackson, *Conditionals*; Blackwell, Oxford, 1987, p. 3).

<sup>3</sup> Véase «General propositions and causality», en la colección de artículos de Ramsey *Foundations* (editada por D. H. Mellor), Routledge and Kegan Paul, Londres, 1978, p. 143, nota.

añadimos el enunciado condicional a nuestro acopio de creencias habitual (lo creemos verdadero); si el consecuente no se sigue, creemos que el enunciado condicional es falso.

La idea central de Stalnaker en «A theory of conditionals» consiste en extrapolar esta sugerencia relativa a las condiciones bajo las cuales creemos un enunciado condicional al estudio de las condiciones bajo las cuales un enunciado condicional es verdadero, al estudio de las condiciones de verdad de los condicionales; «el concepto de *mundo posible*», dice Stalnaker, «es precisamente lo que necesitamos para hacer esta transición, puesto que un mundo posible es el análogo ontológico de un acopio de creencias hipotéticas»<sup>4</sup>; Stalnaker propuso, en analogía con el experimento de Ramsey, un conjunto de condicionales de verdad para condicionales basado en el concepto de mundo posible: «consideremos un mundo posible en el que *A* es verdadero, y que por otra parte difiere mínimamente del mundo actual. «Si *A*, entonces *B*» es verdadero (falso) precisamente cuando *B* es verdadero (falso) en ese mundo posible»<sup>5</sup>.

Estas son condiciones de verdad para un condicional en términos del valor de verdad de sus enunciados componentes en ciertos mundos posibles. A primera vista, no parece haber dificultad en precisar estas condiciones de verdad con una semántica de tipo kripkeano; sin embargo hay un elemento que ofrece problemas: la referencia a «el mundo en el que *A* es verdadero que difiere mínimamente del actual». Las condiciones de verdad para enunciados con los operadores modales usuales sólo emplean determinantes cuya semántica está bien fijada por la lógica clásica:  $\Box A$  es verdadero si y sólo si *A* es verdadero en *todo* mundo posible de un cierto tipo;  $\Diamond A$  es verdadero si y sólo si *A* es verdadero en *algún* mundo de un cierto tipo. El estudio de las propiedades de ‘todo’ y ‘algún’ están bien fijadas. Esto no ocurre con ‘el mundo en el que *A* es verdadero difiere mínimamente del actual’, y ello motiva la introducción por Stalnaker de una precisión formal de la idea de mundo mínimamente diferente o máximamente similar a uno dado; los diversos modos de precisar formalmente la noción de similaridad entre mundos constituirán el núcleo de cada semántica de condicionales basada en mundos posibles.

Para tratar con precisión la idea de mínima diferencia, Stalnaker introduce en el aparato habitual de la semántica de mundos posibles (cuyo elemento más característico es una relación de accesibilidad entre índices o mundos posibles) las *funciones de selección*. Una función de selección es una función diádica que toma como argumentos a un enunciado y a un mundo y que tiene como valor el mundo posible más similar al mundo argumento en el que el enunciado es verdadero; si *f* es una función de selección,  $\alpha$  es un mundo y *A* un enunciado,  $f(A, \alpha)$  es el mundo más similar a  $\alpha$  en el que *A* es verdadero. Con ayuda de las funciones de selección, las condiciones de verdad de un condicional propuestas por Stalnaker pueden expresarse así:  $A > B$  es verdadero en el mundo  $\alpha$  si y sólo si *B* es verdadero en  $f(A, \alpha)$ .

<sup>4</sup> «A Theory of conditionals», en W. L. Harper, R. Stalnaker y G. Pearce (eds.): *Ifs. Conditionals, belief, decision, chance and time*, Reidel, Dordrecht, 1981, p. 45.

<sup>5</sup> «A theory of conditionals», p. 45.

<sup>6</sup> Leído ‘Si *A* entonces *B*’;  $>$  es un signo introducido por Stalnaker para abreviar una construcción condicional cualquiera.

Stalnaker también introduce un artificio secundario para tratar los antecedentes imposibles: el mundo imposible,  $\lambda$ , que es el valor de una función de selección  $f$  para  $\langle A, \alpha \rangle$  cuando  $A$  es imposible respecto a  $\alpha$  (cuando  $A$  es falso en todo mundo accesible desde  $\alpha$ ). En  $\lambda$  todo es verdadero, y esto hace verdadero a todo condicional con un antecedente imposible.

Los requisitos formales que parezca plausible exigir que reúna una función de selección serán de extrema importancia a la hora de estudiar la semántica de las construcciones condicionales y determinarán la validez o invalidez de esquemas o patrones de inferencia que involucren condicionales (del mismo modo que las restricciones formales impuestas a la relación de accesibilidad determinan la validez o invalidez de ciertos esquemas modales). Naturalmente, no es posible caracterizar formalmente una función de selección apropiada para fijar el valor de verdad de todo condicional, pero las restricciones de tipo formal que podamos imponer a toda función apropiada serán decisivas para extraer consecuencias acerca del comportamiento semántico de  $>$ <sup>7</sup>.

Acabamos de ver que uno de los requisitos de Stalnaker para una función de selección es que verifique  $f(A, \alpha) = \lambda$  si  $A$  es imposible respecto a  $\alpha$ . Otro de los requisitos impuestos por Stalnaker es que si  $A$  es verdadero en  $\alpha$ ,  $f(A, \alpha)$  sea el mismo  $\alpha$ ; esto parece una experiencia obvia, habida cuenta de que  $\alpha$  debe ser el mundo que más se parece a  $\alpha$ ; así cuando  $A$  es verdadero en el mundo real, la semántica de  $A > B$  es la del condicional material para antecedentes verdaderos. Un último requisito que impone Stalnaker a una función de selección  $f$  es:

(\*) Para todo mundo posible base  $\alpha$  y cualesquiera antecedentes  $B$  y  $B'$ , si  $B$  es verdadero en  $f(B', \alpha)$  y  $B'$  es verdadero en  $f(B, \alpha)$ , entonces  $f(B, \alpha) = f(B', \alpha)$ .

Intuitivamente, (\*) también parece bastante plausible; si  $f(B, \alpha)$  y  $f(B', \alpha)$  fueran diferentes,  $B$  y  $B'$  serían ambos verdaderos en dos mundos diferentes, pero parece que uno de ellos habría de ser más similar a  $\alpha$  que otro, luego la selección de  $f$  para  $\langle B, \alpha \rangle$  y  $\langle B', \alpha \rangle$  debería dar el mismo resultado, a saber, el más próximo a  $\alpha$  de los mundos en que  $B$  y  $B'$  son ambos verdaderos.

Una función de selección proporciona un orden lineal de los mundos accesibles desde uno dado en razón de su similaridad a éste. Tomemos el conjunto de los valores (otros que  $\lambda$ ) de  $f(A, \alpha)$  para todos los enunciados  $A$  de nuestro lenguaje, y llamémosle 'V'; definimos, para  $\beta$  y  $\gamma$  pertenecientes a V:  $\beta$  es tanto o más similar a  $\alpha$  que  $\gamma$  ( $\beta \leq_{\alpha} \gamma$ ) si y sólo si para algún enunciado  $A$  que valga tanto en  $\beta$  como en  $\gamma$ ,  $\beta$  es  $f(A, \alpha)$ <sup>8</sup>. El orden es también un buen orden si suponemos que para todo

<sup>7</sup> Veremos que la imposibilidad de caracterizar unívocamente a una función de selección se debe a la extrema ambigüedad pragmática de un condicional; diversas especificaciones de una función de selección corresponden a distintas desambiguaciones o explicitaciones del contexto en que se profiere un condicional.

<sup>8</sup> Que el orden así definido es lineal se sigue de (\*).  $\leq_{\alpha}$  es obviamente reflexiva;  $\leq_{\alpha}$  es antisimétrica: si  $\beta \leq_{\alpha} \gamma$  y  $\gamma \leq_{\alpha} \beta$ , hay  $B, B'$  verdaderos en  $\beta$  y  $\gamma$  tales que  $f(B, \alpha) = \beta$  y  $f(B', \alpha) = \gamma$ ; entonces por (\*)  $\beta = \gamma$ ;  $\leq_{\alpha}$  es transitiva: observemos primero que de (\*) se sigue que si  $f(A, \alpha) = \beta$  y  $f(B, \alpha) = \gamma$ , entonces  $f(A \vee B, \alpha) = \beta$  o  $f(A \vee B, \alpha) = \gamma$ ; supongamos que  $\beta \leq_{\alpha} \gamma$  y  $\gamma \leq_{\alpha} \delta$ ; entonces hay  $A$  verdadero en  $\beta$  y  $\gamma$  tal que  $f(A, \alpha) = \beta$  y  $B$  verdadero en  $\gamma$  y  $\delta$  tal que  $f(B, \alpha) = \gamma$ ; por 1 a observación y (\*) nuevamente,  $f(A \vee B, \alpha) = \beta$  ( $A$  es verdadero en  $f(A \vee B, \alpha)$ , y  $A \vee B$  es verdadero en  $f(A, \alpha)$ , que es  $\beta$ ); como  $A \vee B$  es verdadero en  $\delta$ ,  $\beta \leq_{\alpha} \delta$ ;  $\leq_{\alpha}$  es conexa: supongamos que no es cierto que  $\beta$

conjunto de mundos posibles hay un enunciado de nuestro lenguaje verdadero en todos y sólo esos mundos, esto es, si nuestro lenguaje tiene medios para expresar toda proposición (identificamos, como es usual, una proposición con un conjunto de mundos posibles). Esta suposición no es en general legítima, pero podemos hacer que una función de selección tome como argumentos a proposiciones y mundos, en lugar de enunciados y mundos, y reformular las condiciones de verdad de  $A > B$  así:  $A > B$  es verdadero en  $\alpha$  si y sólo si  $B$  es verdadero en  $f$  (la proposición expresada por  $A, \alpha$ )<sup>9</sup>. Que  $\leq_\alpha$  es un buen orden de los mundos de  $V$  se ve ahora fácilmente; sea  $A$  un conjunto no vacío y no igual a  $(\lambda)$  de mundos en  $V$ ; entonces  $f(A, \alpha)$  es el primer elemento de  $A$ .

En «A theory of conditionals», Stalnaker ofrece un cálculo axiomático para la lógica proposicional suplementada con la conectiva  $>$ <sup>10</sup>. Stalnaker y Thomason publicaron en 1970 una prueba de que este cálculo es completo, es decir, una prueba de que todas las fórmulas válidas de acuerdo con la semántica indicada (incluyendo en particular los requisitos formales impuestos a las funciones de selección) son deducibles en ese cálculo (el cálculo llamado 'C2'). Los operadores modales usuales son definibles en C2 así:  $\Box A = \text{def } \neg A > A$ ;  $\Diamond A = \text{def } \neg (A > \neg A)$ .

Es fácil comprobar que la semántica de  $>$  hace inválidos cientos patrones de inferencia válidos para el condicional material y la implicación estricta<sup>11</sup>. Uno de ellos es el llamado 'silogismo hipotético': de  $A > B$  y  $B > C$  no es posible inferir que  $A > C$ . Una evidencia intuitiva en favor de esta característica de  $>$  es el siguiente argumento inválido con la forma de un silogismo hipotético: si Tales no hubiese sido el primer filósofo, Tales no lo habría sido; así, si Anaxímenes hubiese sido el primer filósofo, Anaximandro habría sido el primer filósofo<sup>12</sup>.

El tipo de inferencia en que se refuerza el antecedente de un condicional (el pase de  $A > B$  a  $A \wedge C > B$ , válido para el condicional material y el condicional estricto) también es inválido según la semántica de  $>$ <sup>13</sup>. La invalidez intuitiva del siguiente argumento apoya esta característica de la semántica de  $>$ : si bebo agua saciaré mi sed; luego si bebo agua y he echado ácido sulfúrico en el agua saciaré mi sed.

También es válida para el condicional material y la implicación estricta la inferencia por contraposición: de «Si  $A$  entonces  $B$ » se concluye «Si no- $B$  entonces

---

$\leq_\alpha \gamma$  y que no es cierto que  $\gamma \leq_\alpha \beta$ ; entonces para todo  $A$  verdadero en  $\beta$  y  $\gamma$ ,  $f(A, \alpha) \neq \beta$  y  $f(A, \alpha) \neq \gamma$ ; sea  $A_0$  tal que  $f(A_0, \alpha) = \beta$  y sea  $A_1$  tal que  $f(A_1, \alpha) = \gamma$ ; entonces  $\beta \neq f(A_0 \vee A_1, \alpha)$  y  $\gamma \neq f(A_0 \vee A_1, \alpha)$ ; sea  $\delta = f(A_0 \vee A_1, \alpha)$ ; o  $A_0$  o  $A_1$  son verdaderos en  $\delta$ ; si  $A_0$  es verdadero en  $\delta$ , como  $A_0 \vee A_1$  es verdadero en  $\beta$  tenemos por (\*) que  $f(A_0, \alpha) = \beta = f(A_0 \vee A_1, \alpha) = \delta$ , que es absurdo; análogamente razonamos si  $A_1$  es verdadero en  $\delta$ .

<sup>9</sup> Stalnaker sólo habla de enunciados como valores de  $f$  en «A theory of conditionals», pero en su obra posterior los valores de  $f$  son proposiciones. Este es un punto no esencial, aunque la prueba de que  $\leq_\alpha$  es un buen orden depende de él. Stalnaker, sin embargo, siempre había dado por bueno que el orden de similitud entre mundos inducido por una función de selección es un buen orden.

<sup>10</sup> Cfr. «A theory of conditionals», p. 48.

<sup>11</sup> La relación de implicación entre el ángulo de Stalnaker y el condicional estricto y el condicional material es la siguiente:  $A \dashv\vdash B$  implica  $A > B$  y  $A > B$  implica  $A \supset B$ . Las conversas no valen.

<sup>12</sup> Para ejemplos similares, cfr. «A theory of conditionals», p. 48 y R. Stalnaker: *Inquiry*, M. I. T. Press, Cambridge, Mass., 1984, p. 123.

<sup>13</sup> Cfr. «A theory of conditionals», pp. 48-49, e *Inquiry*, p. 123.

no- $A$ ». Pero  $A > B$  puede ser verdadero y  $\neg B > \neg A$  falso<sup>14</sup>. Un ejemplo que apoya este rasgo de la semántica de  $>$  puede ser el siguiente argumento intuitivamente inválido: si no pego a mi hijo, no deja de llorar; así, si deja de llorar, le pego.

Es importante señalar ahora que en «A theory of conditionals», Stalnaker consideraba su teoría del comportamiento semántico de los condicionales como una teoría general que explicaba la estructura conceptual de toda construcción condicional; las diferencias entre diversos tipos de construcciones condicionales quedarían recogidas por el cálculo C2 y su semántica. Además, estas características de la teoría tendrían relevancia epistemológica: «Una interpretación general», dice Stalnaker, «que evita dividir los sentidos y las explicaciones del comportamiento de un concepto en varios contextos se adecúa al patrón familiar de la explicación científica en el hecho de que fenómenos diversos y aparentemente desiguales se contemplan como derivados de una fuente común. Por estas razones, entiendo que es un argumento potente en favor de la teoría semántica el hecho de que trate al condicional como un concepto unívoco»<sup>15</sup>.

## La teoría de Lewis

En su artículo de 1971 «Completeness and decidability of three logics of counterfactual conditionals»<sup>16</sup>, David Lewis expuso por primera vez públicamente una teoría semántica de los condicionales contrafácticos en la que había trabajado desde finales de los sesenta, en contacto con Stalnaker. Aunque similar en espíritu a la teoría de Stalnaker, la teoría de Lewis no es una teoría del comportamiento semántico de todas las construcciones condicionales, como pretende ser la teoría expuesta en «A theory of conditionals», sino únicamente una teoría de los contrafácticos. La razón de esta autorrestricción la comentaremos con un cierto detalle al hablar posteriormente de las diferencias entre condicionales indicativos y condicionales subjuntivos.

En 1973, en su artículo «Counterfactuals and comparative possibility», Lewis expuso informalmente su teoría (el artículo de 1971 ofrecía únicamente una condensada prueba de la completitud y decibilidad de una teoría axiomatizada con una conectiva contrafáctica primitiva); de un modo aún más detenido, detalló todo su trabajo anterior sobre contrafácticos en su libro *Counterfactuals*, también de 1973. Expondremos sucintamente el núcleo de su teoría, en base a los dos trabajos recién citados.

Lewis motiva su análisis en las deficiencias de la implicación estricta para proporcionar una explicación correcta de la naturaleza semántica de los contrafácticos. Supongamos que los condicionales contrafácticos, de la forma  $A \Box \rightarrow B$ , fueran condicionales estrictos, esto es, condicionales materiales «necesitantes», de la forma  $\Box (A \supset B)$ , donde  $\Box$  es un operador modal de necesidad cualquiera; entonces

<sup>14</sup> Cfr. «A theory of conditionals», p. 49, e *Inquiry*, p. 124.

<sup>15</sup> Cfr. «A theory of conditionals», p. 50.

<sup>16</sup> *Theoria*, 37 (1971), pp. 74-85.

$A \Box \rightarrow B$  sería verdadero si y sólo si  $A \supset B$  es verdadero en todos los mundos posibles de un cierto tipo, o equivalentemente, si y sólo si todos los mundos de un cierto tipo en que valga  $A$  son mundos en que vale  $B$ .

Para mostrar que este análisis es incorrecto, Lewis exhibe secuencias de condicionales en las que apela esencialmente al fenómeno de la falacia de reforzar el antecedente, que ya hemos comentado:

(1) Si hubiese bebido agua, habría saciado mi sed

(2) Si hubiese bebido agua y hubiese echado ácido sulfúrico en el agua, habría saciado mi sed

(3) Si hubiese bebido agua y hubiese echado ácido sulfúrico en el agua y después de echar ácido sulfúrico en el agua hubiese destilado el agua, habría saciado mi sed

(4) Si hubiese bebido agua y hubiese echado ácido sulfúrico en el agua y después de echar ácido sulfúrico en el agua hubiese destilado el agua y después de destilar el agua hubiese vuelto a echarle ácido sulfúrico, habría saciado mi sed.

En principio la secuencia podría extenderse indefinidamente<sup>17</sup>. Seguramente, todos convendríamos en que (1) es verdadero (bajo ciertos supuestos razonables acerca del contexto en que (1) se profiera). Pero si (1) es verdadero, todos los otros miembros de la lista lo son, si consideramos que un condicional contrafáctico es un condicional estricto. Esto no es deseable, pues (2) y (4) son palmariamente falsos; así, concluye Lewis, los condicionales contrafácticos no pueden ser condicionales estrictos.

Lo que ocurre es, según Lewis, que lo estricto que es un contrafáctico varía; al evaluar el valor de verdad de los contrafácticos de una secuencia como la exhibida, no prestamos atención siempre a los mundos que verifican ciertos requisitos fijados; más bien, lo que ocurre es que a medida que «reforzamos» el antecedente los requisitos de semejanza al mundo real son menos estrictos; inversamente, son más estrictos cuantas menos premisas hagamos formar parte del antecedente. Por ello Lewis habla de los condicionales contrafácticos como «variablemente estrictos»<sup>18</sup>.

El valor de verdad de un contrafáctico  $A \Box \rightarrow B$  no depende del valor de verdad de  $B$  en todos los mundos de un cierto tipo en que vale  $A$ , sino del valor de  $B$  en los mundos más parecidos al actual en que vale  $A$ ; hay múltiples grados de similitud al mundo real y esto hace que los respectos relevantes en la consideración del valor de verdad de diversos contrafácticos sean también muy variados, de modo que ningún operador modal (estricto) puede cumplir la misión de restringir la consideración a los mundos apropiados en cada caso. En concordancia con este razonamiento, la teoría de Lewis se caracteriza por ofrecer (en distintas versiones) una precisión formal de la noción de similitud comparativa a un mundo dado (de los mundos accesibles desde éste). Con ayuda de estas precisiones formales, Lewis ofrece condiciones de verdad para los contrafácticos.

Una de estas precisiones (la favorecida en *Counterfactuals*) reposa en la noción de sistema de esferas. Un sistema de esferas  $S$  es una función que toma un conjunto

<sup>17</sup> Cfr. *Counterfactuals*, Blackwell, Oxford, 1986 (reimpr. revisada); p. 10, y «Counterfactuals and comparative possibility», en D. Lewis: *Philosophical papers*, vol. II, Oxford U. P., Oxford, 1986, p. 4.

<sup>18</sup> Cfr. *Counterfactuals*, p. 13, y «Counterfactuals and comparative possibility», p. 5.

posible  $i$  como argumento y le asigna como valor un conjunto de conjuntos (esferas) de mundos posibles,  $\$i$ ; cada uno de los conjuntos de  $\$i$  representa un grado o ámbito de similitud a  $i$ : si  $S \in \$i$ , los mundos de  $S$  son más similares a  $i$  que los que no están en  $S$ <sup>19</sup>. Como ocurría con las funciones de selección de Stalnaker, resulta plausible intuitivamente imponer algunas condiciones formales a los sistemas de esferas, si éstos han de proporcionar escalas de similitud entre mundos, pero no es posible determinar completamente un sistema de esferas que fije el valor de verdad de todo condicional<sup>20</sup>.

Los requisitos formales que Lewis impone a un sistema de esferas  $\$$  son los siguientes: «(C)  $\$i$  está *centrado sobre  $i$* ; esto es, el conjunto  $\{i\}$  con  $i$  como único miembro pertenece a  $\$i$ ; (1)  $\$$  está *nidificado*, esto es, siempre que  $S$  y  $T$  pertenecen a  $\$i$ , o  $S$  está incluido en  $T$  o bien  $T$  está incluido en  $S$ , (2)  $\$$  está *cerrado bajo uniones*, esto es, siempre que  $S$  es un subconjunto de  $\$i$  y  $US$  es el conjunto de todos los mundos  $j$  tales que  $j$  pertenece a algún miembro de  $S$ ,  $US$  pertenece a  $\$i$ ; (3)  $\$i$  está *cerrado bajo intersecciones (no vacías)*, esto es, siempre que  $S$  es un subconjunto no vacío de  $\$i$  y  $\cap S$  es el conjunto de todos los mundos  $j$  tales que  $j$  pertenece a todo miembro de  $S$ ,  $\cap S$  pertenece a  $\$i$ »<sup>21</sup>.

La justificación intuitiva del requisito (C) es que seguramente  $i$  es tan similar a sí mismo como cualquier otro mundo y también más similar que ningún otro<sup>22</sup>. (1) es un modo de exigir que todos los mundos accesibles desde  $i$  sean comparables en cuanto a su similitud a  $i$ . Las condiciones (2) y (3) se justifican observando que las uniones e intersecciones de conjuntos de mundos nidificados deben constituir ámbitos de similitud: los mundos que pertenezcan a una unión de esferas serán más similares al real que los que queden fuera de la unión; similarmente con la intersección; como dice Lewis, la nidificación garantiza por sí sola (2) y (3) si asumimos que cada  $\$i$  es un conjunto *finito* de esferas.

Con la ayuda de la noción de sistema de esferas, Lewis da condiciones de verdad para los condicionales contrafácticos del siguiente modo<sup>23</sup>.  $A \Box \rightarrow B$  es verdadero en un mundo  $i$ , relativamente a un sistema de esferas  $\$$ , si y sólo si o bien (1) ningún mundo en que  $A$  es verdadero pertenece a ninguna esfera  $S$  en  $\$i$ , o bien (2) alguna esfera  $S$  en  $\$i$  contiene al menos un mundo en que  $A$  es verdadero y  $A \supset B$  vale en todo mundo de  $S$ . Dicho de forma más intuitiva: un contrafáctico  $A \Box \rightarrow B$  es verdadero en un mundo  $i$  si o bien  $A$  es imposible respecto a  $i$  (como la de Stalnaker la teoría de Lewis hace verdaderos los contrafácticos con antecedente imposible) o

<sup>19</sup> Cfr. *Counterfactuals*, p. 14.

<sup>20</sup> Como en el caso de las funciones de selección, los distintos sistemas de esferas aceptables intuitivamente corresponderían a las diversas explicitaciones del contexto relativas al criterio de similitud al mundo real; en cierto contexto, unos mundos posibles pueden estar más cercanos (en cuanto a similitud) que otros, mientras que en otro contexto puede ocurrir justamente lo contrario.

<sup>21</sup> *Counterfactuals*, p. 14.

<sup>22</sup> Aunque en esto último Lewis no está tan seguro; en «Completeness and decidability of three logics of counterfactual conditionals» y en el capítulo 6 de *Counterfactuals* explora las propiedades de una lógica de contrafácticos enteramente similar a la lógica cuya semántica estamos esbozando, salvo en que no incluye la restricción (C), la llamada 'lógica CO' en aquel artículo; CO más (C) es la lógica llamada 'C1'.

<sup>23</sup> Cfr. *Counterfactuals*, p. 16.

bien  $A$  es posible en algún ámbito de similaridad tal que  $B$  es verdadero en todos los mundos de ese ámbito de similaridad en que  $A$  es verdadero.

Un modo preliminar conveniente de comparar la teoría de Lewis con la de Stalnaker es mostrar la naturaleza formal de la relación de similaridad comparativa entre mundos que genera un sistema de esferas. Partiendo de un sistema  $\$$  de esferas, definimos para cada mundo  $i$  la relación  $\leq_i$  entre los mundos accesibles desde  $i$ <sup>24</sup>:  $j \leq_i k$  si y sólo si toda esfera  $S$  de  $\$$  que contiene a  $k$  contiene también a  $j$ . Entonces se muestra fácilmente que  $\leq_i$  es reflexiva, transitiva, y conexa. Sin embargo,  $\leq_i$  no es antisimétrica (de modo intuitivo: puede haber mundos diferentes que pertenezcan exactamente a los mismos ámbitos de similaridad a  $i$ , mundos diferentes igualmente similares a  $i$ ); así,  $\leq_i$  no es un orden, sino lo que a veces se llama cuasi-orden o preorden. Naturalmente, no es un buen orden, pero ni siquiera se exige que no haya cadenas infinitas descendentes de mundos cada vez más similares a  $i$ , cadenas de la forma  $\langle i_n: n \in \omega \rangle$  con  $i_{n+1} <_i i_n$  (e  $i_m \neq i_n$  para  $m \neq n$ )<sup>25</sup>. En el siguiente apartado comentaremos el significado intuitivo de estas diferencias entre las relaciones de similaridad inducidas por las funciones de selección de Stalnaker y los sistemas de esferas de Lewis.

Lewis probó en su artículo citado en 1971 que las verdades lógicas de un sistema de lógica proposicional con  $\Box \rightarrow$  de acuerdo con la semántica propuesta son exactamente las fórmulas deducibles de un sistema axiomático<sup>26</sup> (mostró además que este conjunto de fórmulas es decidible).

Los requisitos formales impuestos por Lewis a la relación de similaridad intermundana son, como hemos visto, más débiles que los de Stalnaker. De hecho, las relaciones de similaridad de Stalnaker son casos especiales de relaciones de similaridad comparativa de Lewis. Así, es natural que los patrones de inferencia que vimos que eran inválidos en la teoría de Stalnaker lo sean también en la de Lewis; así ocurre efectivamente<sup>27</sup>. Pero inversamente, es de esperar que ciertos patrones de inferencia sean válidos en la teoría de Stalnaker y no en la de Lewis; también es así, y estos patrones de inferencia serán objeto de discusión acto seguido.

### El tercio excluso condicional y la suposición del límite

El aparato semántico de Stalnaker genera una relación de buen orden que intuitivamente se interpreta como una relación de similaridad entre mundos posibles. La relación análoga generada por el aparato de Lewis es un cuasi-orden que permite cadenas infinitas de descendentes. La suposición de que la relación de similaridad es antisimétrica que hace la teoría de Stalnaker, la suposición de que no hay dos mundos diferentes igualmente similares a uno dado, es lo que Lewis llama 'suposi-

---

<sup>24</sup> Lewis la define para todos los mundos, pero para visualizar las diferencias con la relación inducida por una función de selección de Stalnaker entre los mundos accesibles, esta restricción es más cómoda; por lo demás, este punto no es esencial.

<sup>25</sup>  $<_i$  se define así:  $j <_i k$  si y sólo si  $j \leq_i k$  y no es el caso que  $k \leq_i j$ .

<sup>26</sup> El mencionado C1, simplificando en *Counterfactuals*.

<sup>27</sup> Cfr. *Counterfactuals*, pp. 31- 36, y «Counterfactuals and comparative possibility», p. 17.

ción de Stalnaker' y que a veces también se llama 'suposición de unicidad'. La suposición de que la relación de similaridad no tiene cadenas infinitas descendentes, la suposición de que siempre en todo conjunto (no vacío) de mundos hay un mundo más similar al mundo base que todos los demás, que también hace la teoría de Stalnaker, es lo que Lewis llama 'suposición del límite'.

La suposición de unicidad hace válido un esquema que involucra la conectiva condicional, conocido como 'principio de tercio excluso condicional', y que no es válido sin asumir dicha suposición:

$$(A > C) \vee (A > \neg C);$$

en efecto, si hacemos la suposición de unicidad, hay un único mundo máximamente similar al real (o al mundo base, en general) en el que vale  $A$ ; entonces, por el principio de tercio excluso habitual, en ese mundo vale  $C$  o  $\neg C$  y por lo tanto vale o  $A > C$  o  $A > \neg C$ ; si negamos la suposición, el conjunto de mundos máximamente similares al actual en que  $A$  vale puede no ser un unitario, puede haber en él dos mundos tales que en uno de ellos valga  $C$  y en el otro  $\neg C$ . El principio parece bastante plausible; en un contexto dado, seguramente todos optaremos por uno de los dos extremos de la disyunción: si llueve, me mojaré o bien si llueve, no me mojaré; si hubiese estudiado, habría suspendido, o bien si hubiese estudiado, no habría suspendido.

Un ejemplo habitual para argumentar que la suposición de unicidad es ilegítima es el siguiente par de contrafácticos ideado por Quine:

Si Bizet y Verdi hubiesen sido compatriotas, Bizet habría sido italiano

Si Bizet y Verdi hubiesen sido compatriotas, Bizet no habría sido italiano

No parece haber razón para que en un mundo máximamente similar al real en el que Bizet y Verdi sean compatriotas, o ambos sean italianos o ambos sean franceses; un mundo en que ambos son italianos parece tan similar al nuestro como uno en que ambos son franceses, bajo una idea razonable de similaridad; así, la suposición de unicidad no valdría en general, y tampoco el principio de tercio excluso condicional: los dos contrafácticos de Quine serían falsos, pues los consecuentes no valdrían en todos los mundos máximamente similares al actual en el que Bizet y Verdi hubiesen sido compatriotas. La semántica de Lewis está diseñada especialmente para contemplar casos como el de los contrafácticos de Quine; la relación de similaridad entre los mundos no es lineal (ni siquiera es un orden), y dos mundos distintos pueden ser igualmente similares al base; un contrafáctico es verdadero cuando el consecuente es verdadero en *todos* los mundos de cierto ámbito de similaridad en que el antecedente sea verdadero, y esto es lo que parece no ocurrir con ninguno de los contrafácticos de Quine.

Sin embargo, también hay argumentos en favor de una teoría que adopte la suposición de unicidad. Del principio de tercio excluso condicional se sigue la consecuencia altamente plausible de que la negación de un contrafáctico (con antecedente posible) es equivalente a la negación interna de su consecuente, esto es, si  $A$  es posible,  $\neg (A > C) \Leftrightarrow (A > \neg C)$  (la dirección  $\leftarrow$  usa las propiedades de la lógica condicional comunes a las teorías de Stalnaker y Lewis). En efecto, el modo habitual de negar un contrafáctico y, en general, un condicional, es negar internamente su consecuente; si no es cierto que si llueve, me mojo, es porque si llueve no

me mojo. El propio Lewis observa lo siguiente: «dado el Tercio Excluido Condicional, no podemos decir con verdad cosas como la siguiente:

*No es el caso que si Bizet y Verdi fuesen compatriotas, Bizet habría sido italiano; y no es el caso que si Bizet y Verdi fuesen compatriotas, Bizet no habría sido italiano; sin embargo, si Bizet y Verdi fuesen compatriotas, Bizet habría sido o no habría sido italiano.*

Esto es:

$\sim (\varphi \Box \rightarrow \psi) \ \& \ \sim (\varphi \Box \rightarrow \sim \psi) \ \& \ (\varphi \Box \rightarrow \psi \vee \sim \psi)$

Yo quiero decir esto, y pienso que es probablemente verdadero; mi propia teoría fue diseñada para hacerlo verdadero. Pero a primera vista, debo admitirlo, suena a contradicción. La teoría de Stalnaker respeta, y la mía no, la opinión de cualquier hablante del lenguaje común que cuida de insistir en que es una contradicción»<sup>28</sup>.

En «A defense of conditional excluded middle», de 1981, Stalnaker esboza una ampliación de su teoría semántica que intenta hacer plausibles sus peculiares características formales. A propósito de la suposición de unicidad, dice: «se trata, sin duda, de una suposición altamente imposible acerca del tipo de relación de similitud que usamos para interpretar los condicionales, y es una suposición que la teoría semántica abstracta que quiero defender hace. Pero como muchas suposiciones idealizadas hechas en la teoría semántica abstracta, puede relajarse en la aplicación de la teoría. En general, para aplicar una teoría semántica a la interpretación del lenguaje tal como es usado, uno no necesita asumir que todo determinante semántico está completa y precisamente definido»<sup>29</sup>. La idea de Stalnaker consiste en lo siguiente: sin modificar su teoría semántica abstracta, cambiar la suposición acerca del determinante semántico principal, la función de selección; reconociendo que la relación de similaridad no es un orden lineal de los mundos accesibles, Stalnaker postula que una función de selección admisible corresponde a una elección para cada antecedente *A* de *uno* de los mundos máximamente similares al base en que *A* es verdadero. Cada función de selección admisible corresponderá, según Stalnaker, a una resolución de la indeterminación implícita en todo condicional, y a una especificación de la relación de similaridad privilegiada por el contexto.

Reconociendo la infradeterminación de la selección de toda función de selección, Stalnaker ofrece una modificación de las condiciones de verdad de los enunciados de su lógica condicional inspirándose en el método de supervaluaciones de van Fraassen<sup>30</sup>. Un enunciado será verdadero de acuerdo con una supervaluación si es verdadero de acuerdo con toda función de selección admisible (de acuerdo con toda valuación normal); un enunciado será falso en una supervaluación si es falso en toda valuación (para toda función de selección); no será ni verdadero ni falso si es verdadero de acuerdo con algunas funciones de selección admisibles y falso de acuerdo con otras<sup>31</sup>.

<sup>28</sup> Cfr. *Counterfactuals*, p. 80.

<sup>29</sup> «A defense of conditional excluded middle», en Harper *et al.* (eds.): *Ifs*; p. 89; ver también *Inquiry*, p. 134.

<sup>30</sup> Tal como este autor lo expone en «Singular terms, truth-value gaps and free logic», *Journal of philosophy*, 63 (1966), pp. 481-495.

<sup>31</sup> Cfr. «A defense of conditional excluded middle», p. 93; *Inquiry*, p. 134; ver también *Counterfac-*

De acuerdo con una supervaluación, los contrafácticos de Quine no son ni verdaderos ni falsos, pero cada uno es verdadero para ciertas valuaciones clásicas o determinaciones de una función de selección admisible (o determinaciones completas del contexto). El principio de tercio excluido condicional sigue siendo válido de acuerdo con las nuevas condiciones de verdad basadas en supervaluaciones;  $(A > C) \vee (A > \neg C)$  es verdadero de acuerdo con toda función de selección admisible, pues en todo mundo  $\alpha$  y para toda función de selección  $f$ ,  $A > C$  es verdadero o  $A > \neg C$  es verdadero, debido al hecho de que  $C \vee \neg C$  es verdadero en  $f(A, \alpha)$ . Stalnaker puede concluir que «es porque se pretende que los antecedentes contrafácticos representen situaciones posibles únicas que los ejemplos que muestran que puede no ser así son un problema. Uno debería responder a este problema, según creo, no revisando las condiciones de verdad para los condicionales de modo que el problema no aparezca, sino más bien reconociendo lo que de todos modos debemos reconocer: que en la práctica hay un gran potencial de indeterminación en las condiciones de verdad de los contrafácticos»<sup>32</sup>. La suposición del límite consiste en asumir que todo antecedente posible es verdadero en un mundo máximamente cercano al actual. La teoría de Stalnaker hace esta suposición: el cometido de una función de selección admisible consiste precisamente en seleccionar un mundo máximamente similar en que valga el enunciado argumento. La teoría de Lewis no hace la suposición del límite: sus condiciones de verdad sólo exigen para que un contrafáctico  $A \Box \rightarrow B$  sea verdadero que  $A \supset B$  sea verdadero en todos los mundos de alguna esfera o ámbito de similaridad (en el que  $A$  sea verdadero en algún mundo); esto equivale a decir que «cuando tomamos [esferas] más y más pequeñas sin nunca acabar, llegamos finalmente a unas en las que el consecuente vale en todo mundo en que vale el antecedente»<sup>33</sup>. La suposición del límite corresponde al siguiente principio de inferencia para condicionales de Stalnaker: para toda clase de proposiciones  $\Gamma$  y proposiciones  $A$  y  $C$ , si  $\Gamma$  implica semánticamente  $C$ , entonces  $\{A > B : B \in \Gamma\}$  implica semánticamente  $A > C$ : si los  $B$ s de  $\Gamma$  valen todos en un mismo mundo (el más similar al real en que vale  $A$ ), entonces  $C$  vale en ese mundo, pues  $\Gamma$  implica  $C$ ; si la suposición del límite no es válida, los  $B$ s de  $\Gamma$  pueden valer en distintos mundos en que  $A$  sea verdadero cada vez más próximos al real, en ninguno de los cuales valga  $C$  (aunque los  $B$ s como conjunto impliquen  $C$ ).

Los ejemplos en que Lewis basa su rechazo de la suposición del límite son de la forma siguiente: «supongamos que adoptamos la suposición contrafáctica de que en este punto aparece una línea que mide más de una pulgada (de hecho tiene algo menos de una pulgada). Hay mundos con una línea que mide 2"; mundos presumiblemente más próximos al nuestro con una línea que mide 11/2"; mundos presumiblemente más cercanos al nuestro con una línea que mide 11/4"; mundos presumiblemente aun más cercanos... Pero ¿cuánto mide la línea en los mundos *más cercanos*

---

*tuals*, pp. 81- 82 y «Counterfactuals and comparative possibility», pp. 7-8, para un ejemplo similar.

<sup>32</sup> Cfr. «A defense of conditional excluded middle», p. 90; *Inquiry*, p. 137.

<sup>33</sup> Cfr. *Counterfactuals*, p. 21; recordemos que las esferas están nidificadas, es decir, totalmente ordenadas por inclusión, de modo que todas las esferas con mundos tanto o más similares al real que todos los de una dada están incluidos en esta última.

con una línea que mide más de una pulgada? Si mide  $1 + x$  para algún  $x$  por pequeño que sea, ¿por qué no hay otros mundos aun más próximos al nuestro en los que mide  $1 + \frac{1}{2}x$ , una longitud aun más próxima a la real? Cuanto más corta hagamos la línea (por encima de  $1$ ), más cerca estaremos de su longitud real; por tanto, presumiblemente, más cerca estaremos del mundo real. Dado que no hay una longitud más corta que todas las demás por encima de  $1$ , no hay mundo más cercano al nuestro que todos los demás entre los mundos con una línea de más de una pulgada»<sup>34</sup>.

En «A defense of conditional excluded middle», Stalnaker defiende la aceptabilidad de la suposición del límite explorando el alcance del argumento de Lewis. Stalnaker introduce una restricción en la idea intuitiva del cometido de una función de selección: una función de selección selecciona un mundo posible en el que el antecedente es verdadero, y que por otra parte es tan parecido al mundo real, *en respectos relevantes*, como sea posible. «La función de selección puede ignorar respectos de similaridad que no son relevantes para el contexto en el que se hace la aserción condicional. Incluso si, en términos de alguna noción general de similaridad global,  $i$  es claramente más similar al mundo real que  $j$ , si los aspectos en que es más similar son irrelevantes, entonces  $j$  puede ser tan buen candidato para la selección como  $i$ »<sup>35</sup>. Puede haber relaciones de similaridad aceptables desde un punto de vista intuitivo tales que varios mundos diferentes, enteramente iguales al real salvo en que la línea tiene en cada uno una longitud diferente por encima de una pulgada, son igualmente similares al mundo real; una función de selección admisible para el antecedente 'la línea mide más de una pulgada' escogería uno cualquiera de estos mundos. Stalnaker observa que en un contexto en el que sí fuese relevante cualquier mínima diferencia en la longitud de la línea contrafáctica sería simplemente inapropiado usar el antecedente en la forma 'si la línea midiera más de una pulgada', pues sería como usar el antecedente 'si la línea tuviese la mínima longitud por encima de una pulgada'; éste sería un antecedente imposible, a no ser que consideráramos, por ejemplo, que la mínima longitud por encima de una pulgada la fija algún factor del contexto, como la técnica de impresión.

Una objeción de Stalnaker contra el argumento de Lewis\* y sus condiciones de verdad para contrafácticos, todos los enunciados de esta forma serían verdaderos: «si la línea hubiese medido más de una pulgada, no habría medido  $x$  pulgadas», donde  $x$  es un número real. Esto es de por sí bastante implausible, pero lo es más si tenemos en cuenta que Lewis define en sus trabajos una conectiva que pretende dar cuenta de los condicionales de la forma «si  $A$ , entonces podría ser el caso que  $B$ » del siguiente modo:  $A \diamond \rightarrow B$  (leído 'si  $A$ , entonces podría ser que  $B$ ') =<sub>def</sub>  $\neg (A \square \rightarrow \neg B)$ ; de acuerdo con esta definición, ningún enunciado de la forma «si la línea hubiese medido más de una pulgada entonces podría haber medido  $x$  pulgadas» (donde  $x$  es un número real) sería verdadero, lo cual parece intuitivamente absurdo<sup>36</sup>. Esto último no es tanto un argumento contra la negación de la suposición del límite como la sugerencia de que negar la suposición del límite no es consistente

<sup>34</sup> Cfr. *Counterfactuals*, pp. 20-21; ver también «Counterfactuals and comparative possibility», p. 9, para un ejemplo similar.

<sup>35</sup> Cfr. «A defense of conditional excluded middle», p. 97; *Inquiry*, p. 141.

<sup>36</sup> Cfr. «A defense of conditional excluded middle», p. 98. e *Inquiry*, p. 142.

con el análisis de Lewis de los condicionales que contienen un 'podría' precediendo al consecuente<sup>37</sup>.

### Condicionales contrafácticos y condicionales indicativos

La semántica propuesta por Stalnaker pretendía, como vimos, describir el comportamiento semántico de todas las construcciones condicionales. Lewis limita su semántica a los condicionales contrafácticos, que se corresponden aproximadamente con los condicionales en los que el verbo del antecedente va en subjuntivo; para mostrar que hay una diferencia semántica entre estos condicionales y aquellos en los que el verbo va en indicativo, Lewis apela a un ejemplo ofrecido por Ernst Adams<sup>38</sup>: mientras que 'si Oswald no mató a Kennedy, entonces algún otro lo hizo' parece verdadero, 'si Oswald no hubiese matado a Kennedy, entonces algún otro lo habría hecho' (la forma subjuntiva del primer condicional) parece falso. Otro ejemplo de este tipo puede ser el siguiente: mientras que 'si Shakespeare no escribió *Ricardo III*, entonces lo hizo Francis Bacon' puede ser verdadero (y lo es para algunos), 'si Shakespeare no hubiese escrito *Ricardo III*, entonces lo habría hecho Francis Bacon' es seguramente falso. La teoría de Stalnaker daría, a primera vista, el mismo valor de verdad para ambos condicionales.

En varios lugares Lewis sugiere que este defecto aparente de la teoría de Stalnaker tiene su causa en la motivación intuitiva de dicha teoría. Recordemos que el análisis de Stalnaker de la semántica de los condicionales en términos de mundos posibles y funciones de selección lo motivaba la observación de Ramsey sobre el modo de evaluación de la verdad de un condicional (el «test de Ramsey»): añadir a nuestro acopio de creencias el antecedente, introduciendo las mínimas revisiones necesarias, y ver si nos vemos compelidos a creer el consecuente. Según Lewis, el test de Ramsey no es adecuado para condicionales subjuntivos; al evaluar el valor de verdad de un contrafáctico no añadimos meramente el antecedente a nuestro acopio de creencias; en el ejemplo anterior, al evaluar 'si Oswald no hubiese matado a Kennedy, algún otro lo habría hecho' no suponemos meramente que Oswald no mató a Kennedy, conservando creencias consistentes con ésta como la de que alguien mató a Kennedy; si fuese así, nos veríamos compelidos a afirmar que si Oswald no hubiese matado a Kennedy algún otro lo habría hecho, pero esto es precisamente lo que no aceptamos. Según Lewis, los mundos más probables para el evaluador no son necesariamente los más cercanos o más similares; los mundos en que Kennedy no es asesinado son mundos más próximos al nuestro que los mundos en que lo mata alguien distinto de Oswald<sup>39</sup>.

---

<sup>37</sup> Lewis también había argumentado que la teoría de Stalnaker no puede ofrecer un análisis satisfactorio de los condicionales del tipo  $A \diamond \rightarrow B$ ; Stalnaker responde a estas objeciones en las páginas finales de «A defense of conditional excluded middle»; esta discusión tiene menos interés que la relativa al tercio excluido condicional y la suposición del límite, y en parte, según creo, sólo tiene sentido para un hablante nativo del inglés.

<sup>38</sup> En «Subjunctive and indicative conditionals», *Foundations of language*, 6 (1970), pp. 89-94.

<sup>39</sup> Cfr. *Counterfactuals*, p. 71.

También fue la inspiración en el test de Ramsey lo que sugirió a Stalnaker una interpretación probabilística de su teoría semántica para la que obtuvo un importante resultado en 1970, en su artículo «Probability and conditionals». Stalnaker partió de la idea intuitiva de que un sistema de creencias es representable por una función de probabilidad que asigna un número entre 0 y 1 a cada fórmula de un lenguaje –indicativo del grado de creencia en la proposición significada por la fórmula, o de su «probabilidad subjetiva»–, e impuso ciertas condiciones intuitivamente plausibles al valor de una función de probabilidad para las fórmulas que contuviesen las conectivas usuales y la conectiva condicional  $>$ , contemplando en especial el caso en que la probabilidad de un antecedente fuese 0 (o el caso, intuitivamente de un condicional contrafáctico). La más importante de estas condiciones estaba directamente inspirada en el test de Ramsey: Stalnaker postuló que para toda función de probabilidad  $\text{Pr}$  (y para cualesquiera proposiciones  $A, C$ ),  $\text{Pr}(A > C) = \text{Pr}(C \setminus A)$ <sup>40</sup> (en un eslogan acuñado por Lewis: que la probabilidad de un condicional es una probabilidad condicionada); en efecto,  $\text{Pr}(C \setminus A)$  es un análogo formal del grado de creencia en  $C$  dado que aceptemos  $A$ , que por la sugerencia de Ramsey es directamente proporcional al grado de creencia en el condicional  $A > C$ . El resultado obtenido por Stalnaker a que nos referíamos fue una prueba de que los teoremas de su lógica condicional C2 son exactamente las fórmulas  $\varphi$  tales que para toda fórmula  $\psi$  y toda función de probabilidad  $\text{Pr}$  que cumplierse los requisitos formales impuestos,  $\text{Pr}(\varphi \setminus \psi) = 1$ .

En 1972 Lewis obtuvo un resultado (publicado por primera vez en su artículo «Probabilities of conditionals and conditional probabilities», de 1976) que ponía en duda el valor del hallazgo de Stalnaker, y especialmente el uso que hace de la hipótesis de Stalnaker, directamente inspirada en el test de Ramsey. El resultado principal de Lewis es fácil de exponer; supongamos que hay una conectiva  $\Rightarrow$  en un lenguaje dado tal que para ella vale la hipótesis de Stalnaker; entonces se puede probar que

$$(1) \text{Pr}(A \Rightarrow C \mid B) = \text{Pr}(C \mid A \wedge B), \text{ si } \text{Pr}(A \wedge B) \text{ es positiva}^{41}$$

Por (1) tenemos que

$$(2) \text{Pr}(A \Rightarrow C \mid C) = \text{Pr}(C \mid A \wedge C) = 1$$

<sup>40</sup> Cfr. «Probability and conditionals», en Harper *et al.* (eds.): *Ifs*; p. 120. A veces se llama a esta ecuación ‘hipótesis de Stalnaker’ o ‘axioma de Stalnaker’.  $\text{Pr}(C \mid A)$ , la probabilidad condicionada de  $C$  supuesto que se dé  $A$ , se define así:  $\text{Pr}(C \mid A) = \text{def } \text{Pr}(C \wedge A) / \text{Pr}(A)$ , si  $\text{Pr}(A)$  es positivo; un ejemplo muestra el sentido de la definición; pongamos que tenemos un universo de seis posibilidades, con los valores de verdad en cada una de ellas de dos proposiciones dadas,  $A$  y  $B$ , así:

|   |   |   |   |   |   |   |
|---|---|---|---|---|---|---|
|   | 1 | 2 | 3 | 4 | 5 | 6 |
| A | 0 | 0 | 1 | 1 | 1 | 1 |
| B | 1 | 0 | 0 | 1 | 1 | 1 |

En este caso,  $\text{Pr}(A) = 2/3$ ,  $\text{Pr}(B) = 2/3$ ,  $\text{Pr}(A \wedge B) = 1/2$  y  $\text{Pr}(B \mid A) = \text{Pr}(A \wedge B) / \text{Pr}(A) = 3/4$ : de las cuatro posibilidades en que  $A$  se da,  $B$  se da en tres de ellas.

<sup>41</sup> Una prueba de (1) es como sigue: sea  $\text{Pr}'$  la función que resulta de «condicionalizar» sobre  $B$ ; así, para todo  $D$ ,  $\text{Pr}(D \mid B) = \text{Pr}'(D)$ , y en particular  $\text{Pr}(A \Rightarrow C \mid B) = \text{Pr}'(A \Rightarrow C)$ ; por la hipótesis

$$(3) \Pr(A \Rightarrow C \mid \neg C) = \Pr(C \mid A \wedge \neg C) = 0$$

El teorema de expansión de casos dice que

$$(4) \Pr(D) = \Pr(D \mid C) \cdot \Pr(C) + \Pr(D \mid \neg C) \cdot \Pr(\neg C)$$

Tomando  $D = A \Rightarrow C$  y por (4) y la hipótesis de Stalnaker, tenemos

$$(5) \Pr(C \mid A) = 1 \cdot \Pr(C) + 0 \cdot \Pr(\neg C) = \Pr(C),$$

es decir, que la probabilidad de C condicionada a A es la misma que la probabilidad absoluta de C, para toda función de probabilidad, todo antecedente A y todo consecuente C<sup>42</sup>. Esto es absurdo, pues naturalmente hay antecedentes que al ser supuestos modifican la probabilidad de sus consecuentes. Asumir la hipótesis de Stalnaker, junto a ciertos resultados comunes de teoría de la probabilidad, ha llevado al resultado de que el lenguaje que tomamos como parámetro es un lenguaje trivial, de escasísima fuerza expresiva.

El resultado de Lewis puede considerarse una prueba de que no hay en ningún lenguaje aceptable (en ningún lenguaje en que al menos un consecuente no sea independiente de un antecedente) un conectiva para la que valga la hipótesis de Stalnaker<sup>43</sup>; como la hipótesis era un análogo directo del test de Ramsey, Lewis interpreta su resultado como una prueba de que los condicionales de Stalnaker no son condicionales para los que sea adecuado el test de Ramsey<sup>44</sup>. La alternativa que resta es que la semántica de Stalnaker es una semántica de los condicionales subjuntivos o contrafácticos, pero no una semántica apropiada para los condicionales indicativos (como era de esperar, por ejemplos como el de Adams).

Según Lewis, la hipótesis de Stalnaker tiene una justificación si en ella introducimos una modificación sustancial. Siguiendo a Adams, Lewis observa que  $\Pr(C \mid A)$  puede ser una medida de la asertabilidad de los condicionales indicativos de la forma  $A \rightarrow C$ , esto es, del grado en que el hablante está dispuesto a afirmar el condicional  $A \rightarrow C$ , aunque no sea una medida de su grado de creencia en la verdad de éste (o probabilidad subjetiva); sólo suponer esto último conduciría al absurdo: sólo suponer que hay una conectiva cuyas condiciones de verdad son tales que garantizan la validez para ella de la hipótesis de Stalnaker, para toda función de probabilidad subjetiva. En otros términos, el test de Ramsey sólo sería adecuado para evaluar el grado de justificación que un hablante cree tener para proferir un condicional indicativo en un contexto conversacional típico. «El método de los experimentos mentales», dice Lewis, «da las condiciones de asertabilidad no para contrafácticos sino para condicionales indicativos (...). Por ejemplo, yo afirmo que si Oswald no mató a Kennedy, entonces algún otro lo hizo; lo hago porque  $\Pr(\text{'Algún otro lo hizo'} \mid \text{'Oswald no lo hizo'})$  es alta. Y niego que si Oswald no mató a Kennedy entonces Kennedy no fue asesinado; esto es porque  $\Pr(\text{'Kennedy no fue asesinado'} \mid \text{'Oswald no le mató'})$  es baja (los indicativos son así opuestos en asertabilidad a

de Stalnaker,  $\Pr(A \Rightarrow C) = \Pr(C \wedge A) / \Pr(A)$ ; pero

$$\Pr(C \wedge A) / \Pr(A) = \frac{\Pr(C \wedge A \mid B)}{\Pr(A \mid B)} = \frac{\Pr(C \wedge A \wedge B) / \Pr(B)}{\Pr(A \wedge B) / \Pr(B)} = \frac{\Pr(C \wedge A \wedge B)}{\Pr(A \wedge B)} = \Pr(C \mid A \wedge B)$$

<sup>42</sup> Cfr. «Probabilities of conditionals and conditional probabilities», en *Philosophical papers*, vol. II, pp. 135-136.

<sup>43</sup> Cfr. «Probabilities of conditionals and conditional probabilities», pp. 145-146.

<sup>44</sup> Cfr. «Probabilities of conditionals probabilities», pp. 148-149.

los contrafácticos correspondientes). La observación de Adams acerca de las condiciones de asertabilidad para los condicionales indicativos es compatible con varias posturas alternativas acerca de sus condiciones de verdad, o su carencia de ellas»<sup>45</sup>.

Si la tesis de Adams es válida (esto es, si la asertabilidad de un condicional indicativo es proporcional a su probabilidad condicionada) y sin embargo no hay ninguna conectiva lingüística que pueda interpretarse de tal modo que implique la hipótesis de Stalnaker para toda función de probabilidad subjetiva o sistema de creencias (como probaría el resultado de Lewis), resulta tentadora la idea de que los condicionales indicativos no tienen en absoluto condiciones de verdad fijas; ciertamente, tenemos indicios de que su significado puede ser extremadamente dependiente del contexto: la verdad de 'si Shakespeare no escribió Ricardo III, entonces Bacon lo hizo' depende en grado extremo de lo que consideramos altamente probable y consistente con 'Shakespeare no escribió Ricardo III', que es lo que tenemos en cuenta al evaluar el condicional; en el caso de un contrafáctico, nuestras creencias no parecen ser tan relevantes al evaluar su verdad. El propio Adams ha sostenido que los condicionales indicativos no tienen condiciones de verdad, sino sólo de asertabilidad.

Lewis, apoyándose en H. P. Grice, ha argumentado que aunque las condiciones de asertabilidad de un condicional indicativo no coincidan con las de un condicional material<sup>46</sup>, las condiciones de verdad de los condicionales indicativos son simplemente las condiciones de verdad de los condicionales materiales correspondientes<sup>47</sup>. Su grado de asertabilidad no sería idéntico en razón de ciertos aspectos de implicatura conversacional que regirían el uso de los condicionales indicativos; por ejemplo, «el hablante no debería afirmar el condicional si lo cree verdadero predominantemente porque cree que su antecedente es falso, de modo que la probabilidad de su verdad reside principalmente en su probabilidad de ser vacuamente verdadero. En esta situación, ¿por qué afirmar el condicional en lugar de negar el antecedente? Hacerlo es absurdo; y si es absurdo, es también algo peor: es desorientador. El oyente, confiando en que el hablante no afirma absurdos, asumirá que no lo ha hecho»<sup>48</sup>.

En favor de su tesis, Lewis aduce este resultado: al restar a la asertabilidad de  $A \supset C$  la probabilidad de que el antecedente sea falso incrementada por el grado en que la probabilidad de que el condicional sea falso es un parte de la probabilidad de que el antecedente sea verdadero, obtenemos, formalmente,  $\Pr(A \supset C) - \Pr(\neg A) \cdot (\Pr(\neg C \wedge A) \mid \Pr(A))$ ; simplificando esta expresión de acuerdo con resultados de teoría de la probabilidad se prueba que  $\Pr(A \supset C) - \Pr(\neg A) \cdot (\Pr(\neg C \wedge A) \mid \Pr(A))$ ; es igual a  $\Pr(C \setminus A)$ , la probabilidad condicionada de C

---

<sup>45</sup> *Counterfactuals*, p. 72, nota; una lista de argumentos en favor de la tesis de Adams de que la asertabilidad de un condicional indicativo es proporcional a la probabilidad condicionada de su consecuente dado su antecedente se halla en Jackson: *Conditionals*, pp. 8-16.

<sup>46</sup> Nótese que, en general,  $\Pr(A \supset C) \leq \Pr(C \setminus A)$ :  $\Pr(\text{'Tiro el dado y sale un número par } \supset \text{ sale un seis'}) = 2/3$ , pero  $\Pr(\text{'sale un seis'} \mid \text{'sale un número par'}) = 1/2$ .

<sup>47</sup> En «Probabilities of conditionals and conditional probabilities», p. 142 y ss.; ver también *Counterfactuals*, p. 72, nota.

<sup>48</sup> «Probabilities of conditionals and conditional probabilities», p. 142.

dado A, que por hipótesis es la probabilidad del condicional indicativo  $A \rightarrow C$ . Así, una cuantificación intuitivamente plausible de los factores griceanos de disminución de la asertabilidad de un condicional indicativo ratifica la tesis de Adams.

En su artículo «Indicative conditionals», de 1975, Stalnaker se enfrentó a los problemas planteados por la consideración de las diferencias semánticas aparentes entre condicionales indicativos y condicionales subjuntivos o contrafácticos de un modo similar a como vimos que abordaba las objeciones al principio de Tercio Excluido Condicional. La forma en que da cuenta de esas diferencias en su reciente libro *Inquiry* permanece, en lo esencial, muy próxima al espíritu de sus primeros trabajos sobre condicionales. Concluiremos este apartado con un conciso resumen de los principales puntos de vista de Stalnaker sobre este tema.

En «A theory of conditionals», Stalnaker sugería que «el modo tiende a indicar algo acerca de la actitud del hablante, pero de ninguna forma afecta al contenido proposicional del enunciado»<sup>49</sup>. En «Indicative conditionals», Stalnaker matiza y corrige esta tesis dentro del marco de su teoría semántica abstracta. La idea básica del artículo consiste en observar que todo condicional (como todo enunciado) se profiere en un contexto que presupone un conjunto de situaciones compatibles con lo que aceptan el hablante y los oyentes; este conjunto es representable por un conjunto de mundos posibles, el conjunto de los mundos posibles tales que todas las proposiciones verdaderas en ellos son compatibles con lo que se acepta en la conversación; Stalnaker llama a este conjunto ‘conjunto del contexto’.

En los usos indicativos de los condicionales, parece plausible suponer que el contexto impone un constrictión pragmática sobre la función de selección apropiada para ofrecer su valor de verdad: «si el condicional está siendo evaluado en un mundo en el conjunto del contexto, entonces el mundo seleccionado debe estar también, si es posible, dentro del conjunto del contexto (cuando C es el conjunto del contexto, si  $i \in C$ , entonces  $f(A, i) \in C$ )»<sup>50</sup>. Un mismo condicional indicativo proferido en diferentes contextos podrá tener diferentes valores de verdad, de acuerdo con las diferentes funciones de selección apropiadas en cada caso. Así, Stalnaker reconoce la dependencia del contexto de la verdad de los condicionales indicativos, pero sostiene que la estructura abstracta de sus condiciones de verdad es siempre la misma; una determinación completa del contexto, correspondiente a una función de selección admisible cuyo recorrido está incluido en el conjunto del contexto, fija sin ambigüedad el valor de verdad de un enunciado condicional indicativo. Por lo demás, las condiciones de verdad de un condicional indicativo no son las de un condicional material; Stalnaker ofrece un contraejemplo similar al siguiente: si yo sé que no cometí cierto asesinato, sé que es falso que si el mayordomo no lo cometió, lo cometí yo (observese que bajo la mínima revisión de mis creencias compatible con que el mayordomo no cometiera el asesinato, yo sigo sin haberlo cometido). Pero el condicional material ‘el mayordomo no lo cometió  $\supset$  lo cometí yo’ puede ser verdadero, si el mayordomo fue el asesino<sup>51</sup>.

<sup>49</sup> Cfr. «A theory of conditionals», p. 54, nota 3.

<sup>50</sup> Cfr. «Indicative conditionals», en Harper *et al.* (eds.); *Ijs*; p.199.

<sup>51</sup> Cfr. «Indicative conditionals», p. 203; Stalnaker ofrece otros ejemplos de este tipo en *Inquiry*, p.

Según Stalnaker, «el subjuntivo en inglés y algunas otras lenguas es un instrumento convencional para indicar que las presuposiciones se dejan en suspenso, lo cual quiere decir que en el caso de los enunciados condicionales subjuntivos, la función de selección puede llegar fuera del conjunto del contexto»<sup>52</sup>. El uso del subjuntivo es un modo de eliminar parte de la dependencia del contexto del valor de verdad de un enunciado condicional; el valor de verdad de 'si Shakespeare no hubiese escrito Ricardo III, entonces Bacon lo habría hecho' puede juzgarse «objetivamente», sin tener en cuenta cuáles son las creencias que aceptamos en torno a la escritura de Ricardo III que sobreviven cuando añadimos a nuestro acopio de creencias la de que Shakespeare no lo escribió. 'Si Shakespeare no escribió Ricardo III, entonces Bacon lo hizo' puede ser verdadero si simpatizamos con la tesis de que Bacon fue el auténtico autor de las obras atribuidas a Shakespeare; pero puede ser falso si simpatizamos con la tesis de que el auténtico autor de las obras de Shakespeare fue el conde Rutland.

En *Inquiry* (1984), Stalnaker introduce la noción de condicional abierto; un condicional abierto se caracteriza epistémicamente: es un condicional cuya aceptación por un hablante le llevará a aceptar el consecuente si llega a creer que el antecedente es verdadero, un condicional que indica una disposición del que lo acepta a cambiar sus creencias<sup>53</sup>. Los condicionales abiertos son (extensionalmente, podríamos decir) los indicativos, si bien éstos se caracterizan gramaticalmente. Si creemos verdadero 'si Shakespeare no escribió Ricardo III, entonces Bacon lo hizo', entonces estaremos dispuestos a aceptar que Bacon escribió Ricardo III si llegamos a estar seguros de que Shakespeare no lo hizo; por otro lado, aunque creamos que si Shakespeare no escribió Ricardo III, Bacon no lo habría hecho, no aceptaremos sin más que Bacon no escribió Ricardo III si llegamos a dar por cierto que Shakespeare no lo escribió.

En *Inquiry* Stalnaker se enfrenta con detalle a la tesis, a la que aludimos brevemente, de que los condicionales abiertos o indicativos no tienen condiciones de verdad (la tesis de que un condicional indicativo no expresa una proposición, la tesis de que su significado depende del contexto de modo radical e irreducible); en particular, examina un ejemplo de Allan Gibbard del que este autor se ha servido para apoyar esta tesis. El ejemplo de Gibbard es el siguiente: «Sly Pete y Mr. Stone están jugando al póquer en un barco del Mississippi. Ahora le toca a Pete jugar o pasar. Mi secuaz Jack ve las dos manos, y ve que la de Pete es mala, y que la de Stone es la ganadora. En este momento, la habitación es desalojada [de espectadores]. Pocos minutos después, Zack me pasa una nota que dice: 'si Pete jugó ganó' y Jack me pasa una nota que dice: 'si Pete jugó perdió'. Concluyo que Pete pasó»<sup>54</sup>.

Gibbard observa que las afirmaciones hechas por sus secuaces son ambas verdaderas, pues son afirmaciones sinceras basadas en creencias correctas; pero en-

---

107. El hecho de que todos ellos apelen a la propiedades de nociones epistémicas, como «saber», no parece en absoluto casual.

<sup>52</sup> «Indicative conditionals», p. 200.

<sup>53</sup> Cfr. *Inquiry*, p. 106.

<sup>54</sup> Cfr. *Inquiry*, pp. 208-209, y Gibbard: «Two recent theories of conditionals», en Harper *et al.* (eds.): *Ifs*; p. 231.

tonces el enunciado ( $\beta$ ) 'si Pete jugó, ganó', debe de expresar una proposición diferente cuando es proferido por cada uno de los secuaces, pues para uno es verdadero y para el otro es falso. Podríamos pensar que la diferencia en el significado de ambas preferencias depende de un factor contextual, del mismo modo que el significado de 'Yo lo hice allí' depende de quién lo profiere, a qué acción se refiere con 'hacerlo' y a qué lugar se refiere con 'allí'. «La diferencia en los contextos aquí, sin embargo», dice Gibbard, «tiene una característica extraña. Ordinariamente, cuando el contexto resuelve una ambigüedad pragmática, las características del contexto que la resuelven son conocimiento común al espectador y la audiencia (...). En el ejemplo, por contra, cualesquiera que sean las diferencias contextuales entre las preferencias, son desconocidas para la audiencia. Yo, la audiencia, sé exactamente lo mismo acerca de los dos contextos: que el enunciado es el contenido de una nota que me ha pasado uno de mis secuaces»<sup>55</sup>. Gibbard concluye que el contenido proposicional de los condicionales indicativos es radicalmente dependiente del contexto: las condiciones de verdad de sus preferencias no pueden especificarse de modo general en términos de un parámetro contextual determinable en cada preferencia (de hecho, la verdad de las preferencias de ( $\beta$ ) de Zack y Jack parece depender de toda su historia personal y sus creencias).

Stalnaker rechaza el supuesto de que una preferencia expresa una proposición sólo si la audiencia puede determinar exactamente su contenido; esto no ocurre tampoco, dice, en «un caso en el que la información necesaria para interpretar el enunciado la da el mismo proferir el enunciado, como cuando usted pregunta, en la oscuridad, '¿dónde está usted?', y yo respondo 'por aquí'»<sup>56</sup>. En un caso como éste, una vez que se oye la respuesta, se tiene información suficiente como para saber lo que se ha dicho; análogamente, los condicionales indicativos proporcionan información acerca del sistema de creencias del hablante o del contexto compartido por un grupo de hablantes, y en muchos casos en ello consiste su función comunicativa. De la nota de Zack, infiero que su situación epistémica es tal que estaría dispuesto a creer que Pete ganó si llegará a saber que jugó, aun sin saber por qué circunstancias llega a tal situación epistémica; de alguien que diga que si Shakespeare no escribió Ricardo III, entonces lo escribió Bacon, infiero que simpatiza con la tesis de que Bacon es el autor de las obras atribuidas a Shakespeare, aunque desconozca qué le ha llevado a esa creencia. De este modo, dice Stalnaker, puede aceptarse que los condicionales abiertos expresan proposiciones, si bien altamente dependientes del contexto.

Los defensores de la tesis de que los condicionales indicativos no expresan proposiciones, como Gibbard y Adams, han subrayado las diferencias semánticas aparentes entre condicionales indicativos y condicionales subjuntivos (o contrafácticos): sólo estos últimos serían, según ellos, condicionales que expresan proposiciones. En *Inquiry*, Stalnaker sostiene que estas diferencias encubren una relación profunda: «el contraste entre condicionales abiertos y condicionales que expresan una proposición puede verse como un caso especial de un contraste de más alcance

<sup>55</sup> «Two recent theories of conditionals», p. 232.

<sup>56</sup> *Inquiry*, p. 111.

entre conceptos modales subjetivos y objetivos. Posibilidad puede ser compatibilidad con lo conocido o compatibilidad con las necesidades objetivas. Probabilidad puede ser grado subjetivo de creencia o probabilidad objetiva (...). El contraste entre 'si él mató al Papa' y 'si él hubiese matado al Papa' es paralelo al contraste entre 'el Papa puede haber sido asesinado' y 'el Papa podría haber sido asesinado'<sup>57</sup>.

Para Stalnaker, los condicionales subjuntivos surgen como resultado de una proyección a partir de los condicionales abiertos, que indican una disposición a cambiar nuestras creencias; con un condicional subjuntivo (o contrafáctico) pretendemos expresar una convicción acerca de cómo es el mundo, abstrayéndola de nuestras creencias particulares. «Si pudiésemos distinguir y filtrar aquellos aspectos de nuestra situación epistémica que derivan más de nuestra perspectiva parroquiana que del modo como entendemos que es el mundo, podríamos ser capaces de explicar la aceptación de proposiciones condicionales en términos de los condicionales abiertos que serían aceptables en contextos idealizados que son abstraídos de esos aspectos»<sup>58</sup>. Del mismo modo que el discurso causal puede interpretarse como una proyección de nuestros hábitos sobre el mundo, en el sentido de Hume, el discurso contrafáctico puede interpretarse como una proyección sobre el mundo de nuestras disposiciones a cambiar nuestras creencias<sup>59</sup>.

Mario GOMEZ TORRENTE

---

<sup>57</sup> *Inquiry*, p. 112.

<sup>58</sup> *Inquiry*, p. 116

<sup>59</sup> Cfr. *Inquiry*, p. 103, para la analogía con la explicación de la causalidad por Hume y con otros ejemplos de reducciones filosóficas de lo objetivo a lo subjetivo. Lo dicho en estos últimos párrafos es un resumen condensadísimo de la tesis principal de Stalnaker en el capítulo 6 de *Inquiry*; una exposición detallada de la tesis de la proyección no es necesaria aquí, según creo, pues su relevancia para las cuestiones de la teoría semántica de los condicionales es indirecta.