

REFERENCIA, INTENSIÓN Y VERDAD

Acero, J. / Bustos, E. / Quesada, D. (1985):
Introducción a la Filosofía del Lenguaje.
Madrid: Cátedra (pp. 75-141).

4. REFERENCIA Y SIGNIFICADO

4.1. Introducción

Una de las funciones más importantes del lenguaje es comunicar cosas sobre el mundo. Dicho de otro modo, cuando hablamos, lo hacemos generalmente sobre aquello que nos rodea, tratando de informar a nuestros semejantes de lo que sabemos o creemos acerca de ello. Para eso usamos expresiones significativas, expresiones que nosotros y nuestros interlocutores pueden entender porque están dotadas de significado. Ahora bien, la referencia es precisamente el mecanismo que nos permite ligar esas expresiones dotadas de significado con el mundo, que permite que los demás entiendan que hablamos del mundo. No nos basta con hablar de un modo significativo, no sólo es suficiente que las expresiones que utilizamos sean comprendidas, es preciso que, además, se *refieran* a algo. Claro está que la referencia no sólo nos permite hablar del mundo exterior, también nos permite hablar de nuestro mundo interior (sentimientos, actitudes, creencias, etc.) e incluso de nuestro lenguaje; en este sentido podemos *referirnos* a nuestras intenciones o a nuestras palabras. La referencia y el significado son los dos pilares en que se basa el carácter intersubjetivo del (uso del) lenguaje. Es posible, y deseable, saber a qué me estoy refiriendo cuando hablo, o si dos personas se están refiriendo a la misma o a diferente cosa, etc. Es posible, y deseable, saber qué es lo que significo cuando digo algo, o si dos personas están significando lo mismo, o utilizando del mismo modo palabras con igual significado. La comprensión del lenguaje humano se basa, pues, tanto en la comprensión de lo que es el significado como en la comprensión de lo que es la referencia.

Desde el punto de vista filosófico, el estudio de la referencia es de primera importancia porque está relacionado con las teorías sobre la existencia. Mediante la utilización de la referencia expresamos nuestras opiniones sobre lo que hay, sobre lo que pensamos que existe, en un sentido amplio de 'existe'. Negar que determinados términos tengan una referencia es negar que determinadas cosas existan y, por tanto, delimitar una cierta ontología.

Históricamente, el término 'referencia' se introdujo haciendo pareja con el de 'sentido' en las teorías de G. Frege, pero hay multitud de términos que están relacionados con esta distinción. Así, J. S. Mill, antes que Frege, introdujo la distinción entre 'denotación' y 'connotación'. 'Denotación' es un precedente terminológico de 'referencia': los términos denotan objetos o clases de objetos, mientras que connotan propiedades. Otras distinciones relacionadas con la de 'referencia' y 'sentido' son la de 'referencia' como opuesta a 'significado', debida a B. Russell, o la de 'extensión' e 'intensión', de R. Carnap. También se ha utilizado a veces la expresión 'significado referencial' para distinguir una clase especial de significado opuesto al significado emotivo o significado afectivo. Todas estas distinciones que se desarrollan en conexión con teorías particulares sobre el significado y la referencia y, por tanto, quedan al margen de un tratamiento sistemático.

No obstante, antes de fijar una terminología definitiva para la discusión de los problemas que plantean la referencia y el significado es conveniente que examinemos brevemente la

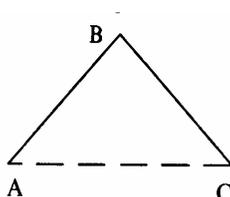
teoría del significado que ha dejado más profunda huella en la moderna filosofía del lenguaje, la teoría de G. Frege. Esta teoría se inscribe dentro del realismo filosófico, que adquiere particulares características en el campo del estudio del lenguaje. Esta es la razón de que esbozemos previamente un marco general en el que situar las teorías lingüísticas dentro de categorías filosóficas de acuerdo con determinados criterios que hacen apelación a hechos generales sobre el signo lingüístico, como la relación entre éste y lo que designa. De este modo, los rasgos fundamentales de las teorías de Frege quedan mejor definidos, permitiendo una comprensión más nítida. La teoría del sentido de Frege es complementada por su teoría de la referencia, realización intelectual de igual o mayor envergadura. Con sus teorías sobre el sentido y la referencia Frege sienta las bases fundamentales de la moderna filosofía del lenguaje determinando los problemas principales que serán debatidos a partir de su obra.

Uno de estos problemas en particular será tratado por dos importantes filósofos: B. Russell y P. F. Strawson. Se trata del problema de la comprensión de la semántica de las descripciones, un tipo especial de expresiones que tiene especial incidencia en la lógica y en el habla cotidiana. A este tema dedicamos un apartado histórico especial por la trascendencia que tuvo en la conformación de las actuales teorías sobre la referencia. Por otro lado, el núcleo indispensable de lo que es preciso conocer para abordar estas teorías modernas y sus cuestiones más características se expone en los apartados dedicados a las diferentes clases de referencia de las expresiones lingüísticas y a dos problemas con los que han batallado los filósofos de este siglo: el vacío y la opacidad referenciales.

4.2. Realismo y teoría semántica

Como es bien sabido, la siguiente representación gráfica se utiliza para representar la relación de significación y sus elementos:

Fig. 1



En este triángulo, utilizado entre otros por los investigadores Odgen y Richards, los puntos A, B y C representan respectivamente al signo, al elemento conceptual y al *significatum* (siguiendo la terminología de J. Lyons, Semántica, 4,1). El *signo* ha de identificarse con una cierta realidad física (una sucesión de ondas sonoras en el caso del lenguaje hablado), es decir, con lo que desde la obra de F. de Saussure se denomina *significante*. El punto B, que representa al componente conceptual, tiene una función mediadora fundamental entre el signo y su *significatum*. El *significatum* es lo aludido o designado por el signo lingüístico, es decir, aquello en lugar de lo cual está el signo en virtud de una convención. El modo en que las diferentes teorías sobre el significado han relacionado este componente conceptual con los demás componentes ha servido de criterio para calificar a éstas, de acuerdo con una terminología filosófica tradicional, de *realistas* o *nominalistas*. De acuerdo con esta¹

¹ Hay otra acepción filosófica de 'realismo' según la cual son realistas los filósofos que mantienen que hay una realidad objetiva independiente de la conciencia del individuo. En este sentido el realismo se opone al idealismo filosófico, que descrea de esa independencia.

terminología, ‘realismo’ designa la posición teórica que mantienen los que creen en la existencia de una realidad objetiva de los conceptos, ideas y/o significados de las expresiones lingüísticas.

Según la mayoría de los filósofos del lenguaje, a toda teoría sobre éste se le debe exigir una definición sobre la función que en ella tiene el elemento conceptual, si es que tiene alguna; sobre las relaciones que unen a ese elemento conceptual por una parte con las entidades físicas que constituyen los signos (o significantes) y, por otra, con los *significata*, las realidades a las que los signos remiten. Esta definición es un requisito fundamental para poder caracterizar el tipo de conexión que la teoría del lenguaje en cuestión establece entre el lenguaje y la realidad.

De acuerdo con nuestras intuiciones más inmediatas, generalmente pensamos que hay un mundo exterior que contiene entidades de diversas clases. A estas entidades se las suele dividir en dos grandes clases: individuos y propiedades. Por otro lado, entendemos Por *concepto* un tipo de representación mental que nos hacemos de las cosas. A determinados conceptos de individuos les atribuiríamos diferentes propiedades: unos conceptos son simples, otros complejos, concretos, abstractos, singulares, comunes, universales, Particulares, etc. No es el mismo el concepto de *Barcelona* que el de *ciudad*, el concepto de *mesa* que el de *desesperación*. Parte de las distinciones que a lo largo de la historia de la filosofía se han establecido entre los conceptos han sido asumidas en la gramática tradicional, como por ejemplo la conocida distinción entre nombres comunes y nombres propios, que refleja la distinción ontológica entre conceptos generales y singulares. En este contexto ha tenido una especial relevancia la distinción entre conceptos universales y particulares, porque a propósito de ella han discutido los filósofos nominalistas y realistas. Considérense, por ejemplo las siguientes oraciones:

- (1) Sócrates es un hombre.
- (2) El calor es agobiante.
- (3) El campeón de Francia es muy rápido.

En estos tres ejemplos, *hombre*, *agobiante*, *rápido* son conceptos universales, porque constituyen propiedades generales de las que en las respectivas oraciones se dice que se aplican a determinados individuos. El problema filosófico residía en el estatus ontológico de las propiedades en cuestión: ¿son de naturaleza mental, psicológica o tienen una autonomía que las hace objetivas, incluso hasta el punto de hacerlas independientes de los términos que las designan? ¿Son los correspondientes conceptos una elaboración psicológica a partir de experiencias concretas o, por el contrario, son extramentales, entidades de las que se adquiere conciencia en el proceso de conocer?

Con todos los matices y reservas necesarios cuando se trata de caracterizar una posición filosófica, en general los filósofos realistas se han inclinado por responder afirmativamente a las segundas partes de las interrogaciones disyuntivas. Desde la posición extrema del realismo trascendental platónico al realismo inmanentista de Aristóteles (la forma como principio universal de individuación), pasando por realistas medievales como Agustín de Hipona o Tomás de Aquino, hasta modernos filósofos del lenguaje como G. Frege o B. Russell, los teóricos del realismo han mantenido que los universales son objetivos, no mentales, independientes de la conciencia y del lenguaje humanos. Los nominalistas, en cambio, han tendido a estar de acuerdo con una respuesta afirmativa a las primeras partes de las preguntas: sólo hay entidades individuales y los conceptos universales se pueden reducir a ellas.

Dentro de esta caracterización tan general del nominalismo no sólo quedan contenidas teorías sobre el lenguaje como las de Ockam, Hobbes o Hume, sino también posturas intermedias o débiles de nominalismo a las que se suele calificar de *conceptualistas*, como las teorías que mantuvieron J. Locke o G. Berkeley. La diferencia principal entre un nominalista radical como G. Ockam y uno moderado o conceptualista, como J. Locke, es que mientras el primero no admitía que las entidades a las que se aplica un concepto universal (un predicado) tuvieran otra cosa en común que el hecho de que les fuera aplicable el mismo término, el segundo sostenía que lo común a esas realidades era una construcción mental, operada mediante abstracción a partir de sus representaciones o cualidades (perceptivas). Según Locke, los conceptos (universales) son el fruto de un proceso de abstracción que parte del conocimiento de las entidades particulares, proceso de abstracción que sigue leyes psicológicas, pero en modo alguno son independientes u objetivos en el sentido que mantienen los realistas.

Resumiendo la situación con respecto a las teorías filosóficas sobre el lenguaje, éstas se pueden caracterizar como realistas o nominalistas (radicales y moderadas o conceptualistas). Las primeras conceden una existencia objetiva a los conceptos, las segundas niegan rotundamente esa objetividad, reduciendo los conceptos a construcciones mentales, o eliminándolos por completo de la teoría.

Para las teorías semánticas realistas la significación es una relación convencional (o natural, como en el caso del *Cratilo*, de Platón) entre signos, esto es, entidades lingüísticas concretas, y entidades conceptuales y/o reales que existen independientemente de aquéllos. En tales teorías suele ocupar un lugar central la función semiótica de *designación*. Las dos principales categorías gramaticales o filosófico-gramaticales que se estudian en ellas, la de nombre (en un sentido amplio) y predicado, nombran o designan. Los nombres designan individuos (entidades individuales) o clases de individuos; los predicados designan *atributos*, esto es, propiedades que *se atribuyen* a lo designado por los nombres. Estos atributos pueden ser de dos clases: cualidades y relaciones. Por ejemplo:

- (4) Barcelona tiene cuatro millones de habitantes
- (5) Barcelona es una ciudad con más habitantes que Gerona

son dos oraciones, la primera de las cuales enuncia una propiedad que satisface (*conviene a* se decía en la filosofía tradicional) un determinado individuo, el designado por el nombre 'Barcelona'. En cambio, la segunda afirma que se da una determinada relación entre dos entidades, las designadas por los nombres 'Barcelona' y 'Gerona'. El predicado 'es una ciudad con más habitantes que' es satisfecho por (o conviene a) dos entidades, Barcelona y Gerona; pero adviértase que no en cualquier orden, sino precisamente en el caso en que 'Barcelona' es el sujeto de la oración y 'Gerona' el segundo término de la comparación. Por eso se dice que la relación es satisfecha por una *secuencia*, esto es, por un conjunto ordenado de entidades. En nuestro ejemplo, el predicado de la oración (5) no sólo es satisfecho, por la secuencia <Barcelona, Gerona>, sino también por otras secuencias como <Gerona, Sitges>, <Londres, Barcelona>, etc.

Con los nombres en sentido amplio (términos) y los predicados se construyen las oraciones, que son las entidades lingüísticas básicas del análisis semántico. Las oraciones declarativas designan estados de cosas, esto es, señalan cuáles son los hechos, cómo son.

La oración

(6) Juan es mejor alumno que Pedro y Jorge

está construida a partir de tres términos y una relación y se puede simbolizar del siguiente modo : $R(t_1, t_2, t_3)$. En general, muchas oraciones de este -tipo simple se pueden simbolizar como $P(t_1...t_n)$. La oración enuncia un determinado hecho, un estado de cosas que es una función de lo que designan sus componentes y de la forma en que éstos están combinados en la oración. Por ejemplo,

(7) Pedro es mejor alumno que Juan y Jorge

designa, y por tanto, significa algo distinto de (6) en virtud de la diferente ordenación de sus términos porque, como es bien sabido, $R(t_1, t_2, t_3)$ no es lo mismo que $R(t_2, t_1, t_3)$. En estas teorías generalmente se suele afirmar que una oración es verdadera cuando el estado de cosas que designa es un hecho, pero el modo en que se analizan los hechos varía mucho de una teoría a otra.

El significado de las oraciones compuestas se construye a partir del significado de las oraciones simples que las componen, con la ayuda de determinadas reglas de interpretación de las partículas que las unen. Así,

(8) Pedro es mejor alumno que Juan y Juan es mejor alumno que Jorge

es una oración que designa lo que designan (9) y (10):

(9) Pedro es mejor alumno que Juan

(10) Juan es mejor alumno que Jorge

y la conjunción 'y' o, mejor dicho, su interpretación o significado.

Por lo tanto, los términos, los predicados y las oraciones están unidos por la relación de designación con entidades como los objetos en un sentido amplio-, los atributos y los estados de cosas o hechos. A ella se reduce, casi por completo, la relación de significación que se -da entre el signo, el concepto y el *significatum*. Ahora bien, en las teorías semánticas realistas a veces se identifican los objetos y los atributos con conceptos (individuales y genera-les) y estados de cosas con proposiciones o, como es más adecuado denominarlas, *ideas* (véase el capítulo 2), es decir, se supone que hay un segundo polo en la relación de designación que es una entidad abstracta, que en algunas versiones es de naturaleza trascendental.

La dificultad que tienen estas teorías al hacer intervenir el nivel conceptual en la relación de significación es la de proveer un criterio de identificación para los componentes de ese nivel. Por ejemplo, si las oraciones expresan ideas, la teoría ha de proporcionar los medios para poder distinguir entre unas ideas y otras, ha de poder determinar cuándo dos oraciones expresan una misma idea, cuándo una oración expresa una idea verdadera, etc. Las entidades abstractas, como las ideas, no se pueden señalar como se pueden indicar los objetos físicos, por lo que las teorías semánticas realistas tienen que acudir, y acuden, a otros procedimientos para su identificación.

El ejemplo de los nombres es considerado en filosofía del lenguaje como paradigmático: en una buena parte de los casos podemos señalar entidades físicas que corresponden a nom-

bres, como en el caso de 'Barcelona', 'Juan Carlos I', 'Montserrat', etc. En el caso de otros nombres no podemos señalar ninguna entidad física que les corresponda, por ejemplo, 'Pegaso', 'el siete', 'el Coco', etc. Las semánticas realistas tienen que proporcionar una especificación de la relación de designación que cubra todos estos casos. Si en esa especificación interviene el nivel conceptual, ha de proporcionar los medios para distinguir, en el caso de los nombres, entre lo que nombran diferentes clases de nombres, nombres a los que sólo corresponde una entidad individual, nombres a los que no les corresponde ninguna, más de una, etc. Lo mismo sucede en el caso de los predicados, en el que la relación de designación los ha de poner en relación con los conceptos generales (los universales de la filosofía medieval), y en el de las oraciones, en el que ha de establecer las diferentes relaciones que éstas mantienen con las ideas.

4.3. La teoría de G. Frege sobre el significado

Se considera que G. Frege (1848-1925) no sólo fue el fundador de la lógica contemporánea, sino que también fue el gran iniciador de la moderna filosofía del lenguaje. Las principales ideas de Frege sobre el lenguaje, aunque no están expuestas de un modo sistemático, como lo están sus teorías lógicas, se encuentran en sus artículos «Sobre el sentido y la referencia» (1892) y «Sobre el concepto y el objeto» (1892). Estas ideas están indisolublemente unidas a, e incluso motivadas por, su filosofía de la lógica y de las matemáticas:

Frege fue llevado a ellas en el curso de sus investigaciones sobre los fundamentos de la matemática, aunque posteriormente se convirtieron en el núcleo de su teoría sobre el signo lingüístico. El punto de partida de las reflexiones sobre el lenguaje de G. Frege es la relación de identidad, relación que tiene una especial relevancia lógica. ¿Es esta relación una relación entre nombres, o entre objetos designados por nombres? Si se tratara de una relación entre objetos, razonaba Frege, un enunciado que formulara una identidad entre dos objetos nombrados por diferentes nombres, por ejemplo $a = b$, sería equivalente (o sinónimo) a un enunciado como el que afirma la identidad de un objeto consigo mismo, por ejemplo $a = a$. Esto tendría como consecuencia que ' $a = b$ ' sería un enunciado vacío de contenido empírico, informativo, pues tendría el carácter analítico que se atribuye a ' $a = a$ '. Pero esto no es así, puesto que ' $a = b$ ' puede ser una aserción informativa, hablar del mundo y no del lenguaje, como se pone de relieve en el ejemplo clásico de Frege:

(11) el lucero del alba es el lucero vespertino

Con este ejemplo, Frege quiso mostrar claramente que (11) no formula simplemente la identidad de un objeto consigo mismo (la del planeta Venus en este caso, al que se refieren tanto la expresión 'el lucero del alba' como 'el lucero vespertino'), a diferencia de la oración

(12) el lucero del alba es el lucero del alba

sino que (11) constituye una afirmación informativa sobre el mundo, una afirmación que nos dice que se da un determinado hecho. El carácter informativo de (11) viene dado por *la forma* en que sus expresiones nominales *se refieren* al planeta Venus. Este carácter sintético reside en el hecho de que la expresión nominal 'el lucero del alba' se refiere al planeta Venus de un modo diferente a como se refiere a él la expresión 'el lucero vespertino'. No obs-

tante, esta forma de referirse no es arbitraria, sino que se corresponde con una forma de darse el objeto, con una propiedad bajo la cual el lenguaje nos presenta el objeto.

De este ejemplo y de ejemplos semejantes a éste, Frege concluyó que las expresiones nominales tienen dos funciones semióticas diferentes que es preciso distinguir cuidadosamente. Por una parte, esas expresiones *se refieren* a un objeto y, por otra, expresan un sentido, un peculiar modo de darse el objeto que designan. El objeto a que se refiere la expresión nominal es *su referencia*, mientras que el sentido es ese particular modo en que el lenguaje nos presenta el objeto, modo que ha de corresponderse con una determinada propiedad que el objeto posee. Así, la referencia de la expresión nominal ‘el lucero del alba’ es el planeta Venus y el sentido de esa expresión es la propiedad de que Venus sea la estrella que brilla al amanecer.

Según Frege, toda expresión nominal, gramaticalmente correcta, tiene un sentido, aunque no toda expresión nominal posea una referencia. Por ejemplo, tienen sentido y referencia las expresiones nominales ‘el número natural que sigue al tres’, ‘Barcelona’, ‘el Tibidabo’, etc. En cambio, tienen sentido pero no referencia las expresiones nominales ‘el mayor número natural’, ‘el último número primo’, ‘el Cancerbero’, ‘don Quijote’, etc. En este último caso, las expresiones lingüísticas tienen como sentido propiedades que no cumple o satisface ninguna entidad individual, lo cual no es obstáculo para que las utilicemos en ocasiones de forma correcta. Según la teoría de Frege, es claro que toda expresión nominal que tenga una referencia ha de tener un sentido, pero que no siempre ocurre lo contrario. Igualmente se sigue de su teoría que dos expresiones nominales que tengan el mismo sentido han de tener la misma referencia, si es que la tienen, pero no al revés. Dos expresiones con diferente referencia han de tener necesariamente diferente sentido. Así, por ejemplo, ‘el lucero del alba’ y ‘el lucero vespertino’ son expresiones nominales que tienen distinto sentido pero la misma referencia; ‘el mayor número natural’ y ‘el último número primo’ son expresiones que tienen diferente sentido, aunque no tengan ninguna referencia; ‘el rey de España’ y ‘el monarca español’ son expresiones con el mismo sentido y, por tanto, con la misma referencia; ‘el rey de España’ y ‘el presidente del Gobierno español’ son expresiones con diferente referencia en la actualidad y, por tanto, han de expresar sentidos diferentes.

Frege advirtió el problema que para su teoría suponen los nombres propios. Si bien es fácil identificar la propiedad que constituye el sentido de las expresiones nominales complejas, como las descripciones, no es tan fácil en el caso de los nombres propios. Aunque en su origen los nombres propios pudieron constituir descripciones de propiedades que poseían los individuos que los portaban (y así sucede en algunas lenguas de pueblos primitivos), esto no sucede en la actualidad en la mayoría de las lenguas naturales. ¿Cuál es el sentido, entonces, de los nombres propios? Una solución propuesta por Stuart Mill, antes de Frege, era que los nombres propios no poseen un sentido (connotación, según la terminología de Mill), sino que solamente poseen una referencia (una denotación), que su función es puramente designativa. Sin embargo, Frege optó por afirmar que los nombres propios poseen un sentido peculiar: la descripción o la propiedad con que los hace equivaler quien los usa. Así, si el filósofo Aristóteles es conocido por un hablante bajo la propiedad de ser el maestro de Alejandro Magno, el sentido del nombre ‘Aristóteles’ equivaldrá, para ese individuo, al de la expresión nominal ‘el maestro de Alejandro Magno’. El sentido de los nombres propios variará por lo tanto según se los haga equivaler con unas u otras expresiones nominales que sean correferenciales con ellos, esto es, que designen al mismo individuo que ellos. La solución propuesta por Frege, que se opone a la de Stuart Mill, ha dado origen a

una polémica de amplias consecuencias en la filosofía del lenguaje del siglo XX, polémica que aún se encuentra lejos de una conclusión definitiva.

La comprensión del sentido, del modo en que pueden darse diversos objetos (posibles) o entidades individuales, no depende del conocimiento directo de éstos, sino del conocimiento del lenguaje: se puede captar el sentido de una expresión nominal sin conocer a su referente o, dicho de otro modo, podemos *saber* a qué se refiere una expresión lingüística sin que conozcamos su referencia. Ahora bien el sentido de una expresión lingüística ayuda a encontrar su referente, si es que se le busca. Como dice Frege, «lo ilumina parcialmente», pues establece una de las posibles formas de su determinación. No conocemos nunca todas las posibles determinaciones de un objeto, por lo cual no podemos tener un conocimiento completo de él ; incluso puede que, al ser el objeto solamente posible, él sentido no determine, ni siquiera parcialmente, un objeto o individuo realmente existente, como en el caso de la expresión nominal ‘el quinceavo planeta del sistema solar’.

Otra de las ideas básicas de la moderna filosofía del lenguaje que Frege formuló es que el sentido y la referencia de las expresiones lingüísticas complejas es una función (esto es, depende por completo y de modo unívoco) del sentido y la referencia de las expresiones componentes. Esta idea, conocida como el principio de composicionalidad, está en la base de la generalización de la distinción entre sentido y referencia a las expresiones lingüísticas que no son puramente nominales, es decir, a los predicados y a las oraciones.

En el caso de las expresiones oracionales, Frege sostuvo que su referencia está constituida por *el valor de verdad* que poseen, esto es, lo verdadero o lo falso, mientras que su sentido es la *idea* que expresan (pensamiento en su terminología). El argumento que aduce Frege para mantener que los valores veritativos son la referencia de las oraciones y que las ideas son sus sentidos se basa en el principio de la sustituibilidad. Según este antiguo principio leibniziano una expresión se ha de poder sustituir por otra, en el seno de una oración, sin que en ella cambie el valor de verdad, siempre que la primera designe lo mismo que la segunda. A este principio se le conoce precisamente como el principio de la sustituibilidad *salva veritate*. Luego, razonó Frege, dado que la referencia de una oración depende de la de sus componentes, si cambiamos uno de sus elementos por otro con la misma referencia, la referencia de la oración completa no ha de cambiar en absoluto. Por ejemplo, si en la oración

(12) el lucero del alba es el planeta Venus

cambiamos una de sus expresiones componentes, ‘el lucero del alba’ por otra correferencial, como ‘el lucero vespertino’ tendremos

(13) el lucero vespertino es el planeta Venus.

¿Qué es lo que ha cambiado de (12) a (13)? Evidentemente, no ha cambiado el valor de verdad, puesto que, si (12) es verdadera, (13) también lo es. Como era de prever, al no cambiar la referencia de las expresiones que se intercambian, queda inalterada la referencia de la oración completa. La referencia de la oración ha de ser, por tanto, el valor de verdad. Sin embargo, entre (12) y (13) hay una clara diferencia: expresan diferentes pensamientos o ideas. Esto es natural y previsible pues, aunque las expresiones que se han intercambiado tienen la misma referencia, no tienen el mismo sentido. Lo que ha resultado alterado es el sentido de la oración, es decir, la idea que ésta expresa.

El valor de verdad de una oración depende de la referencia de sus elementos, mientras que la idea o pensamiento que expresa depende del sentido de sus componentes. Así pues, el valor de verdad es la referencia de la oración, mientras que la idea que expresa constituye su sentido.

Las oraciones con respecto a su sentido y a su referencia se comportan del mismo modo que las expresiones nominales: pueden tener sentido, pero no referencia y pueden tener sentido y referencia, lo que no pueden tener es referencia, pero no sentido. La oración

(14) El mayor número primo es divisible por 7

tiene sentido, está construida de un modo perfectamente gramatical, expresa una idea, pero no obstante carece, según Frege, de referencia, de valor de verdad, ya que uno de sus componentes, la expresión 'el mayor número primo' no designa nada. En cambio, la oración

(15) El número natural siguiente al 10 es divisible por 7

es una oración con sentido y referencia; con sentido pues es una oración bien construida y con referencia porque todos sus componentes la tienen (designa lo falso).

Una consecuencia obvia de la teoría de Frege es que todas las oraciones verdaderas tienen una misma referencia, lo verdadero, y lo mismo sucede con las oraciones falsas, pues todas designan lo falso. Frege consideraba que los valores de verdad eran objetos con los cuales las oraciones estaban en el mismo tipo de relación que las expresiones nominales con sus referencias: del mismo modo que una referencia, una realidad individual, nos puede ser presentada lingüísticamente bajo muy diferentes propiedades, del mismo modo lo verdadero y lo falso son objetos que se nos pueden aparecer de muy diferentes formas, en conexión con muy diferentes ideas o sentidos. Cada referencia de una oración tiene, por así decirlo, un modo peculiar de descomposición, de dividirse en componentes. Aunque estos componentes son en general expresiones nominales y predicativas, en las oraciones compuestas se da el caso de que los componentes en cuestión puedan ser oraciones. En las oraciones compuestas coordinadas no hay ningún problema, pues por el principio de composicionalidad, la referencia de una oración completa depende de la referencia de las oraciones componentes, esto es, de su valor de verdad. Por ejemplo, el valor de verdad de la oración compuesta:

(14) La inflación aumentó un 15% y el paro subió al 13%

depende de los valores de verdad de sus dos oraciones componentes : será la verdad cuando *ambas* oraciones simples sean verdaderas y el de falsedad cuando alguna de ellas sea falsa. ¿Pero qué sucede en el caso de las oraciones compuestas subordinadas? En estas oraciones compuestas, parte de los componentes son expresiones nominales y predicativas, y parte son oraciones. ¿Cómo se combinan unas y otras para determinar el valor de verdad de toda la oración? La alternativa más congruente con la teoría de Frege parecería consistir en que las oraciones son también un tipo de expresiones nominales, que designan lo verdadero o lo falso, por lo que toda oración compuesta lo está por nombres y predicados. Sin embargo, el propio Frege advirtió que se presentan varias dificultades si las oraciones subordinadas se consideran como nombres de los valores veritativos. La primera de ellas es que existen oraciones en el lenguaje natural que no se refieren a valores de verdad sino a individuos. El ejemplo que Frege utilizó es la oración

(15) Quien descubrió la forma elíptica de las órbitas planetarias murió en la miseria.

En esta oración compuesta, la oración subordinada ‘quien descubrió la forma elíptica de las órbitas planetarias’ refiere no a lo verdadero o a lo falso, sino a un individuo, el astrónomo Kepler, y se comporta en todo como una expresión nominal normal. De hecho, este tipo de oraciones se pueden hacer equivaler fácilmente a expresiones nominales descriptivas; por ejemplo, la oración subordinada en (15) puede sustituirse por la expresión ‘el individuo que descubrió la forma elíptica de las órbitas planetarias’.

Un segundo tipo de excepción a la tesis de que la referencia de las oraciones es su valor de verdad es el que proporcionan los denominados estilo directo e indirecto. Utilizamos el estilo directo cuando citamos textualmente las palabras de otra persona (y por ello las ponemos entre comillas), mientras que usamos el indirecto cuando las parafraseamos. Por ejemplo, de las oraciones

(16) El presidente dijo: «es preciso aumentar la inversión»

(17) El presidente dijo que era preciso aumentar la inversión la primera pertenece al estilo

directo, mientras que la segunda está en el indirecto. En la oración (16), la oración subordinada aparece entrecomillada, para indicar que se menciona (véase 2.2), que quién utiliza (16) se refiere a las palabras de otra persona ; en consecuencia, la oración entrecomillada se refiere a sí misma y no a un valor de verdad, su referencia son palabras y no valores veritativos.

En cambio, en (17) no hay expresiones mencionadas, sino una referencia indirecta a lo que dijo una persona. Si la referencia de la oración subordinada fuera en este caso un valor de verdad, por ejemplo, lo verdadero, por el principio de sustituibilidad, *salva veritate*, podríamos intercambiar esa oración por otra con el mismo valor de verdad sin que se alterara la referencia o valor de verdad de la oración total. Sin embargo, esto no es así, puesto que si sustituimos la expresión ‘(que) era preciso aumentar la inversión’ por otra oración verdadera como ‘(que) dos más dos son cuatro’, el resultado, esto es:

(18). El presidente dijo que dos más dos son cuatro

puede resultar una oración no verdadera, esto es, puede cambiar de referencia. La conclusión de Frege es que en estos casos de cita indirecta la referencia de la oración subordinada es un pensamiento o idea, esto es, que su referencia es su sentido: «en el estilo indirecto se habla del sentido, por ejemplo, del discurso de otro. Se ve claramente que, incluso en este modo de hablar, las palabras no tienen su referencia usual, sino que se refieren a lo que habitualmente es su sentido» («Sobre sentido y referencia», pág. 53).

La solución de Frege no sólo se aplica a los casos de cita indirecta, sino también a otros en que la oración subordinada es una oración completiva. Hay predicados como *creer*, *saber*, *opinar*, *sospechar*, *temer*, etc., que toman como complementos oraciones a las que no les es aplicable el principio de sustituibilidad *salva veritate*. Por ejemplo, en la oración

(19) Galileo creía que la tierra es redonda

no se puede sustituir la oración subordinada ‘la Tierra es redonda’ por otra igualmente verdadera, como ‘el sistema solar tiene nueve planetas’, sin que pueda cambiar el valor de verdad de toda la oración (19). De hecho, mientras que (19) es verdadera, la oración

(20) Galileo creía que el sistema solar tiene nueve planetas

es falsa. También en estos casos la referencia de la oración subordinada es una referencia indirecta, esto es, un pensamiento y no un valor de verdad.

Frege considera en su artículo «Sobre el sentido y la referencia» casos más complejos en que la referencia de una oración no es un valor de verdad. No obstante, estos tres que hemos mencionado (el de las oraciones subordinadas nominales, el de las citas, directa e indirecta, y el de los contextos indirectos pueden servir como ejemplo de que la distinción entre sentido y referencia no siempre se puede aplicar fácilmente al nivel oracional. Para comprender la aplicación de la distinción entre sentido y referencia a las expresiones predicativas hay que remontarse a una distinción fregeana de carácter más general: la que divide las expresiones en completas, o saturadas, e incompletas o no saturadas. Una expresión completa es una expresión a la que no le falta ningún elemento para la determinación de su referencia, sea ésta un individuo o un valor de verdad (de acuerdo con la terminología de Frege, un objeto). Por ejemplo, son expresiones saturadas ‘Juan Carlos I’, ‘el rey de España’, ‘quién reina en España en 1981’, ‘el rey Juan Carlos I es el rey de España’, etc. Todas ellas designan un objeto individual que es la referencia de tales expresiones. Sin embargo, son expresiones no saturadas las expresiones ‘el rey de x’, ‘quién reina en x en 1981’, ‘x es el rey de España’, etc. A todas estas expresiones les falta un elemento, el ocupado por la variable x, para poder tener una referencia. La referencia que puedan tener depende de lo que pongamos en el lugar de esa x. Por ejemplo, si en la expresión ‘el rey de x’ ponemos en lugar de la x el nombre ‘España’, la expresión así completada designará a Juan Carlos I; si en cambio sustituimos la x por ‘Suecia’ la expresión referirá a Carlos Gustavo de Suecia. Del mismo modo, si sustituimos en la expresión oracional ‘x es el rey de España’ la variable x por ‘Juan Carlos I’ la oración tendrá como referencia lo verdadero, mientras que si la sustituimos por ‘Adolfo Suárez’ tendrá como referencia lo falso.

Las expresiones incompletas se pueden considerar como un tipo de funciones, funciones que toman como argumentos expresiones saturadas o completas y dan como valores la referencia de las expresiones así completadas. Dentro de este tipo de funciones tienen particular importancia las funciones que dan como resultado valores veritativos, pues a ellas pertenecen las expresiones predicativas (pero no sólo ellas). Toda oración simple y enunciativa se puede concebir, de acuerdo con Frege, como el resultado de la aplicación de una expresión funcional, el predicado, a una o varias expresiones nominales, aplicación que da como resultado o bien lo verdadero o bien lo falso (siempre que las expresiones nominales tengan referencia, claro está). Si el hecho de adjuntar esa expresión predicativa a una o más expresiones nominales tiene como resultado lo verdadero, se puede decir que un objeto o varios (los referidos por las expresiones nominales) caen bajo un determinado concepto, que es lo referido por la expresión predicativa. Según Frege, «un concepto es una función cuyo valor es siempre un valor veritativo» («Función y concepto», 1891). Una oración como

(21) Zaragoza está entre Madrid y Barcelona

es una oración que consta de tres expresiones nominales, 'Zaragoza', 'Madrid' y 'Barcelona', y una expresión predicativa o término conceptual, 'está entre'. El concepto al que refiere 'está entre' es en este caso una relación, pues requiere más de un argumento para producir un valor de verdad como referencia. Como hay al menos tres argumentos (la secuencia <Zaragoza, Madrid, Barcelona>) que hacen a la oración verdadera, el concepto al que se refiere 'está entre' no es un concepto vacío, pues se aplica con verdad al menos a tres cosas. Por el contrario, el concepto designado por la expresión 'es el mayor número natural' es un concepto vacío, porque no hay ningún argumento al que, aplicándole la expresión predicativa, dé como resultado la verdad de una oración.

Los conceptos no son objetos. Esto quiere decir, según la teoría de Frege, que no se puede uno referir a ellos mediante una expresión nominal. La expresión 'el concepto x' no se refiere, según Frege, a un concepto, sino a un objeto. De igual modo, la oración

(22) El concepto 'es rey de España' es igual al concepto 'es jefe del Estado español'

no es un enunciado sobre la identidad entre dos conceptos, sino entre dos objetos. El concepto, dice Frege, «es de naturaleza esencialmente predicativa», por lo que no puede constituir la referencia de un nombre. Siendo esto así, ¿qué es lo que enuncia (22)? ¿cuándo puede afirmarse que dos conceptos son idénticos?

Según Frege, «la relación lógica fundamental es la de *caer un objeto bajo un concepto*: a ella pueden reducirse todas las relaciones entre conceptos» («Consideraciones sobre sentido y referencia», 1892-1895). Dos objetos son idénticos cuando caen exactamente bajo los mismos conceptos, lo cual no es más que un resultado de la aplicación del principio leibniziano de la indiscernibilidad de los idénticos. Dos conceptos, por su parte, pueden aplicarse exactamente a los mismos objetos, como en el caso de los mencionados en (22), 'es rey de España' y 'es jefe del Estado español', pero ello no quiere decir que los dos conceptos sean idénticos (Frege dudó sobre este punto: cfr. sus «Consideraciones sobre sentido y referencia» y «Sobre concepto y objeto»). Cuando dos conceptos se aplican a los mismos objetos están en una relación *parecida* a la de identidad, pero que no es la identidad misma. Es a esa relación a la que se refiere (22), puesto que es un enunciado que dice precisamente que 'es rey de España' y 'es jefe del Estado español' se aplican al mismo individuo.

Una de las posibles cosas que nos permitiría hablar de identidad entre conceptos sería el hecho de disponer de un criterio para saber cuándo dos expresiones predicativas tienen el mismo sentido. Pues si dos expresiones de esta clase tienen el mismo sentido han de tener forzosamente la misma referencia, de tal modo que han de referirse al mismo concepto. Se podría afirmar entonces que dos conceptos son iguales (al menos) cuando son referidos por expresiones predicativas que tienen el mismo sentido. Pero del mismo modo que en el caso de las expresiones nominales y oracionales, Frege no nos legó ninguna solución al problema de encontrar ese criterio, pues en su obra no se encuentra prácticamente ninguna alusión al sentido de las expresiones predicativas.

Como resumen de la teoría sobre el significado que Frege esbozó, el siguiente cuadro expone, en la medida de lo posible, cuáles son el sentido y la referencia de las expresiones nominales, predicativas y oracionales:

Expresiones	Sentido	Referencia
Nominales	Propiedades, («modo de presentación» lingüística)	Objetos (entidades individuales)
Predicativas	?	Conceptos
Oracionales	Ideas	Valores de verdad

Al reflexionar sobre este esquema hay que tener en cuenta que 1) las propiedades que son el sentido de las expresiones nominales están ligadas, para Frege, con descripciones lingüísticas (que, en el caso de los nombres propios, pueden ser varias; 2) que la categoría de objeto incluye no sólo las entidades individuales que tienen existencia material, sino también algunas de índole abstracta, como los números, los valores de verdad, los momentos de tiempo, etc., y 3) que, en algunas ocasiones, la referencia de una oración puede ser su sentido, esto es, la idea que expresa.

4.4. La teoría de las descripciones de B. Russell

B. Russell mantuvo una teoría del significado en que esta noción está estrechamente ligada a la de referencia. Su doctrina de las expresiones denotativas (en la cual se incluye la llamada ‘teoría de las descripciones’) es un instrumento conceptual diseñado para la solución de alguno de los problemas, planteados por la referencia, que preocupaban a G. Frege, sin recurrir a la noción de sentido, que Russell rechazó. El principio básico de esta doctrina es que los denominados «sintagmas denotativos», que hasta cierto punto equivalen a lo que hasta ahora hemos venido llamando «expresiones nominales», no tienen significado por sí mismos, sino que sólo lo adquieren cuando se encuentran en el seno de oraciones. Son ejemplos de sintagmas denotativos las expresiones ‘un hombre’, ‘el hombre del pelo gris’, ‘quienquiera que fuese’, ‘el que vino a preguntar’, etcétera.

Los sintagmas denotativos se oponen a los nombres propios. Los sintagmas denotativos no tienen significado, pero sí los nombres propios: el significado de éstos es aquello que designan. En el caso de los nombres propios su significado es su referencia. B. Russell soluciona con esta distinción el problema que plantean las afirmaciones de que algo no existe. Así,

(23) el número primo más alto no existe

no es un enunciado sobre un número que fuera el significado-referencia de la expresión ‘el número primo más alto’, pues este sintagma no tiene en sí mismo significado. El razonamiento de B. Russell es más o menos el siguiente: la oración ‘el número primo más alto no existe’ es significativa, luego la expresión-sujeto también lo es; ahora bien, su significado no puede ser su referencia porque tal referencia no existe, por lo tanto la expresión en cuestión no es un nombre propio, es un sintagma denotativo que no tiene significado por sí mismo.

Según Russell, es preciso analizar los sintagmas denotativos para no dejarse engañar por su forma gramatical: mediante el análisis que él propone oraciones como la anterior quedan convertidas en

(24) no es cierto que haya un número primo y sólo uno tal que sea el más alto de los números primos.

Entre los sintagmas denotativos tienen especial importancia las descripciones definidas, expresiones que aparecen encabezadas por los artículos determinados en singular, como en 'el número primo más alto' del ejemplo (23). Las descripciones definidas en cuestión no son nombres propios. Consideremos el ejemplo empleado por el mismo Russell

(25) Scott es el autor de Waverley.

Si 'el autor de Waverley' fuera un nombre propio, entonces o el enunciado equivaldría a 'Scott es Scott', que es un enunciado trivialmente verdadero (analítico), o sería falso. Pero como no ocurre ni una cosa ni otra, 'el autor de Waverley' es una descripción definida a la que hay que aplicar el análisis lógico. En general, los enunciados con descripciones definidas pueden ser de dos clases: afirmaciones de existencia y atribuciones de propiedades (incluyendo en esta última clase los enunciados de identidad). En el primer caso Russell analiza

(26) Existe el actual rey de Francia

como

(27) Hay exactamente un x tal que es rey de Francia en la actualidad

y, en el segundo caso,

(28) El actual rey de Francia es calvo

como

(29) Hay exactamente un x tal que es rey de Francia en la actualidad y es calvo.

Una consecuencia evidente de este modo de análisis es que (28) resulta falsa cuando no hay tal rey de Francia. Igualmente se sigue que Russell no admitía el principio de sustituibilidad para los nombres y las descripciones. Si tenemos un nombre y una descripción que aparentemente designan lo mismo, no podemos sustituir uno por otra en todo contexto. Así, si tenemos

(30) Jorge IV quería saber si Scott era el autor de Waverley

y (25), no podemos inferir, mediante sustitución,

(31) Jorge IV quería saber si Scott era Scott

pues siendo las premisas verdaderas, la conclusión es falsa. Según el modo de análisis de Russell, lo que Jorge IV quería saber es si una persona y sólo una escribió Waverley y si

esa persona era Scott. La teoría de Russell tiene, dicho sea de paso, una curiosa consecuencia para el argumento ontológico. Todas las versiones del argumento ontológico incluyen premisas con expresiones como ‘el ser más perfecto’, etc. De acuerdo con el análisis ruse-lliano la utilización de estas expresiones definidas implica ya la afirmación de la existencia de un ser: ‘el ser que reúne todas las perfecciones’ equivale a ‘hay un x y sólo uno tal que reúne todas las perfecciones’, por lo cual lo que se pretende demostrar se afirma ya en las premisas, invalidando la demostración una falacia de *petitio principii*.

	SIGNIFICADO	REFERENCIA
NOMBRES PROPIOS	su referencia	el objeto
DESCRIPCIONES DEFINIDAS	está incluido en la oración que resulte del análisis	no tienen

4.5. Referencia y presuposición según P. F. Strawson

Las teorías de Strawson son directamente opuestas a las de B. Russell en dos sentidos: según Strawson, cuando se afirma ‘el rey actual de Francia es sabio’, por un lado *no* forma parte de la aserción que exista un rey ‘y sólo un rey actual de Francia y, por otro, no se hace una afirmación ni verdadera ni falsa. Puede parecer deseable que todas las fórmulas hayan de ser verdaderas o falsas, pero esto no tiene por qué suceder en el lenguaje natural, en el que se emiten ejemplares de oraciones declarativas (enunciados en nuestra terminología) que no son ni verdaderos ni falsos. En realidad, hay algunos de éstos en el lenguaje ordinario a los que no se puede asignar uno de los dos valores de verdad ; tal como lo expresa Strawson, la cuestión de su valor de verdad no se suscita. La relación que según Strawson hay entre enunciados como ‘el rey de Francia es sabio’ y ‘existe el rey de Francia’ no es de *implicación lógica (entailment)*, sino de implicación no lógica (*implicación*). En obras posteriores a su artículo «On Referring», Strawson designó esta relación con el nombre de *presuposición*. Los enunciados con descripciones definidas *presuponen* que hay una referencia que les corresponde a dichas descripciones. Así pues, los enunciados presupuestos son condiciones necesarias de la verdad o falsedad de los enunciados que presuponen. La relación queda definida del siguiente modo: E presupone E’ si la verdad de E’ es condición necesaria de la verdad o falsedad de E. Si E’ es falsa, entonces E no es verdadera, pero tampoco es falsa. Para entender bien la posición de Strawson hay que tener en cuenta que utiliza una distinción entre oración (*sentence*) y enunciado (*statement*) diferente de la nuestra. Para Strawson una oración es una entidad lingüística, gramatical, resultado de aplicar las reglas de formación de una lengua a expresiones de esa lengua. Las oraciones son sucesiones de sonidos o trazos, etc. La diferencia estriba pues en que Strawson no distingue entre oración-tipo y ejemplar de una oración. Por otro lado, Strawson liga la noción de enunciado con el concepto pragmático de aserción. Un enunciado es *lo que se hace* al utilizar una oración en un momento y lugar determinado, de acuerdo con determinadas intenciones, las de hacer una afirmación. Como se ve, Strawson no distingue tampoco entre el acto y el resultado del acto. No obstante, de su teoría se sigue que lo que son verdaderos o falsos son los enunciados y no las oraciones. ‘El actual rey de Francia es sabio’ no es un enunciado verdadero o falso si alguien emite asertivamente esa oración en la actualidad; de hecho, ni siquiera constituiría esa proferencia un auténtico enunciado.

La diferencia fundamental entre Russell y Strawson está pues en que el primero creía que la significatividad era una condición suficiente y necesaria para la asignación de valor de verdad a las oraciones, mientras que el segundo añadía a la significatividad la propiedad de ser un enunciado: para que una oración sea verdadera o falsa no sólo ha de ser significativa, tener sentido, sino que además ha de ser utilizada por alguien en unas circunstancias y modos determinados. Las oraciones no son ni verdaderas ni falsas, sólo lo son los enunciados. En adelante, sin embargo, no utilizaremos el término ‘enunciado’ en el sentido de Strawson. Los puntos de vista de Strawson respecto a la referencia se pueden resumir como sigue: en realidad no son las expresiones referidoras (entre las que hay que incluir a las expresiones denotativas y a los nombres propios) las que se refieren a algo o tienen una referencia, sino que son los hablantes quienes, mediante el uso de esas expresiones, hacen referencia a objetos, clases, etc. Cuando se usan las expresiones referidoras como componentes de enunciados para hacer aserciones, esas expresiones están en una relación de presuposición con sus referentes, con la existencia de esos referentes. De acuerdo con la noción pragmática de presuposición, que más adelante introdujo Strawson: son los hablantes quienes al utilizar las expresiones referidoras *presuponen* que los referentes que les corresponden existen.

4.6. La referencia de las distintas expresiones lingüísticas

En primer lugar, una advertencia terminológica: denominaremos expresiones referidoras tanto a las descripciones definidas y las indefinidas, como a los nombres propios. En ocasiones también utilizaremos el término más técnico ‘sintagma’ para traducir el inglés ‘*phrase*’ y así, hablaremos de los ‘sintagmas referidores o referenciales’. Siempre que se hable de que los términos, expresiones o ‘sintagmas se refieren a’, habrá que entender, y en esto seguimos a Strawson, que en definitiva son los hablantes quienes, al usar esas expresiones, ‘se refieren a’. Utilizaremos los dos modos de hablar de modo indistinto, en este buen entendimiento, con el objeto de evitar una cantidad excesiva de perífrasis. Al objeto, o clase, identificado mediante un uso adecuado de las expresiones referidoras le denominaremos ‘referente’ o ‘referencia’. Usaremos ahora este último término de un modo preferente para designar la relación que se da entre expresiones referidoras y sus correspondientes referentes. Son ejemplos de expresiones referidoras: ‘Jorge IV’, ‘el filósofo más leído del siglo XX’, ‘caballo’, ‘Pegaso’, ‘quien descubrió la forma elíptica de las órbitas planetarias’, etc. Otra alternativa terminológica, es la de reservar la denominación ‘referencia’ para la noción pragmática y utilizar el término más neutro ‘extensión’, introducido por R. Carnap, para la noción semántica. Sin embargo, los hechos de que en castellano exista el verbo ‘referir’ y que la terminología carnapiana no se haya extendido nos han llevado a adoptar la primera alternativa.

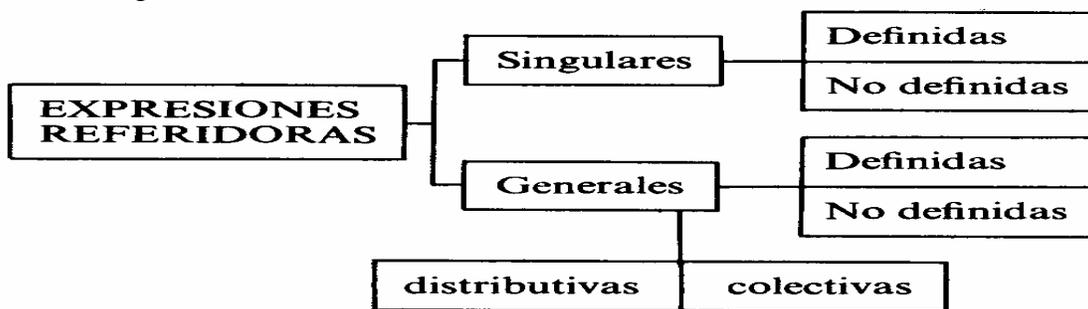
Hay muchas clases de referencia. Para empezar es natural suponer que las diferentes categorías gramaticales tradicionales se refieran a distintas cosas. Así, se suele afirmar que los nombres propios se refieren a individuos (‘particulares’, sería más correcto decir), los nombres comunes a conjuntos de individuos, los adjetivos a propiedades de individuos o de conjuntos de individuos, los verbos a acciones, los adverbios a propiedades de acciones, etc. Sólo de las partículas sincategoremáticas (preposiciones, conjunciones, artículos, etc.) se diría que no refieren a nada, que no les corresponde un referente. Son categorías gramaticales cuya única función sería servir de elementos de unión entre las otras categorías para formar categorías gramaticales de orden superior: sintagmas preposicionales, adverbiales,

nominales, verbales, oraciones, etc ; al menos según la semántica tradicional, muy puesta en cuestión en la actualidad.

Pero no solamente las categorías gramaticales básicas refieren, también lo hacen algunas de las categorías superiores. Así, las descripciones tienen referentes, su función es referir a referentes mediante expresiones complejas que bastan para individualizar su referente, en un determinado contexto comunicativo: ‘el actual rey de Francia’, ‘la persona que viene a las ocho’, etc. son expresiones referidoras, como ocurre con algunas cláusulas de relativo; uno de los ejemplos de Frege ya mencionado ilustra este extremo: ‘quien descubrió la forma elíptica de las órbitas planetarias’ es toda una oración que refiere al individuo Kepler.

A falta de una teoría general de la referencia en el lenguaje natural, las discusiones conceptuales se han centrado sobre todo en determinadas categorías que ponen de relieve de un modo especialmente claro el funcionamiento de la referencia. Pero antes de abordar el examen de estas categorías es necesario distinguir previamente entre lo que es una referencia correcta y una referencia que tiene éxito. La corrección de la referencia es una cuestión semántica, mientras que su éxito es una cuestión pragmática. Alguien puede utilizar una expresión referidora construida correctamente, suceder incluso que exista el referente al que se aplica esa descripción y, no obstante, no tener éxito al hacer tal referencia; por ejemplo, por ser incapaz el hablante de identificar el referente, etc. En este capítulo nos limitaremos al tratamiento semántico de la referencia dejando de lado sus aspectos pragmáticos.

Como ya hemos mencionado, entre los referentes se pueden distinguir dos clases de entidades, con larga tradición filosófica: los particulares y los universales, o dicho de otro modo, los individuos y las clases de individuos. Estos referentes son referidos respectivamente por expresiones que podemos calificar como singulares y generales. Además, se puede distinguir, por el modo en que se efectúa una referencia, entre referencias específicas y no específicas, entre referencias definidas y no definidas. Por otro lado, en lo que respecta a las clases de individuos, por el modo, las referencias pueden dividirse en distributivas y colectivas. De tal manera que, respecto a las expresiones que se refieren a individuos o clases, tenemos el siguiente cuadro clasificatorio:



Entre las expresiones que tienen una referencia singular definida se encuentran las descripciones definidas (o sintagmas nominales definidos), los nombres propios y los pronombres (en singular). Los nombres propios merecen un tratamiento aparte, aunque sólo sea por la especial atención que les han dedicado los lógicos y semánticos. Los pronombres, a su vez, son una categoría imposible de tratar si no es mediante un estudio sistemático, e integrado en la teoría lingüística, del contexto de comunicación. Por lo tanto es la categoría de las descripciones definidas la que será objeto de examen. Ya hemos mencionado ejemplos de esta categoría lógico-gramatical. Sin embargo, es preciso tener en cuenta que las fronteras entre descripciones definidas y nombres propios, por una parte, y entre aquéllas y las des-

cripciones no definidas no es fácil de trazar. Así, en el primer caso, ciertos títulos, honoríficos o no, oscilan entre las dos categorías; por ejemplo ‘el Papa’, ‘el Coco’, etc. Por otra parte, las oraciones introducidas por pronombres indefinidos a veces tienen referencia específica y a veces no. Por ejemplo, en ‘cualquiera lo puede hacer bastante mal’ la referencia del sintagma que desempeña la función de sujeto es habitualmente no definida, mientras que en ‘quienquiera que lo mató, sabía lo que se hacía’ se aproxima más a la específica. Además, expresiones que lingüísticamente pueden parecer como no referidoras de una manera específica, pueden operar en determinados contextos comunicativos como descripciones definidas. Así, expresiones como ‘un tal Martínez’ pueden ser utilizadas para hacer una referencia, si no correcta, por lo menos con éxito, a un referente específico, sobre el que no cabe ninguna duda ni al hablante ni al oyente. El uso del artículo determinado no debe engañar al respecto: en ‘el hombre es un animal vertebrado’, la expresión-sujeto no tiene una referencia definida, ni siquiera singular, mientras que ‘el hombre está en la puerta’ puede tenerla en determinadas circunstancias.

Que la referencia general se puede efectuar de una manera distributiva o colectiva se sigue de que cuando nos referimos a una clase podemos referirnos a ella como a un todo o referirnos a cada uno de sus miembros. En ‘los libros cuestan dos mil pesetas’ la expresión ‘los libros’ puede referir de cualquiera de los modos citados; la oración anterior puede indicar que *todos* los libros que considera el hablante cuestan en conjunto ese dinero, o que cada uno de los libros en cuestión tiene ese precio: la desambiguación depende del contexto. Sin embargo, hay casos más claros que ilustran uno y otro modo de referir. En

- (32) los ingleses son un pueblo simpático
- (33) me gusta el cordero asado
- (34) el hombre está condenado a desaparecer de la tierra

es altamente probable la referencia colectiva. En cambio en

- (35) los españoles tienen derecho a una vivienda digna
- (36) los corderos pasaron por el portillo
- (37) afíleme los lápices

la referencia es distributiva.

Los problemas de la referencia singular no definida están inmersos, por otra parte, en una zona de mayor vaguedad, que en este caso ha de entenderse como dependencia del contexto. De hecho, se puede afirmar como regla general que tanto más difícil es fijar el tipo de referencia de una expresión cuanto más apelación al contexto hay que hacer. Compárense las oraciones ‘me lo dijo un pajarito’, en que la referencia es no definida y no específica, con ‘un pájaro me despierta todas las mañanas’ en que la referencia se puede considerar como indefinida, pero en algún sentido específica. Una particular clase de expresiones de este tipo, de un gran interés lógico, es la de los pronombres indefinidos que sirven para cuantificar expresiones, por ejemplo ‘todos’, ‘alguien’, ‘ninguno’, etc. El estatus referencial de estos pronombres ha sido objeto de largas discusiones entre los lógicos y un actual problema de investigación lingüística es dotar a estos términos de una representación semántica adecuada.

4.7. Los problemas de la referencia

Entre los muchos problemas relacionados con la referencia se pueden enumerar los planteados con ocasión de los vacíos referenciales (*referential gaps*), la opacidad referencial, las cuestiones referentes al alcance (*scope*) de los cuantificadores, la denominada inescrutabilidad de la referencia -con sus repercusiones en el problema más general de la *traducción radical*- y, en general, todo el planteamiento del aprendizaje de la referencia. Los hay más o menos técnicos, que atañen directamente a la elaboración de una teoría semántica para el lenguaje natural, y más o menos filosóficos, que están vinculados a una concepción filosófica del lenguaje, de su aprendizaje y conocimiento, etc. Dentro de nuestras limitaciones, se puede intentar esbozar el planteamiento de algunos de estos problemas, su fisonomía conceptual.

A. Los vacíos referenciales

Es un tema que ya suscitó controversias en los orígenes de la teoría de la referencia. De hecho, las teorías de Russell y Strawson se diferencian fundamentalmente en el punto crítico de las repercusiones que tienen los vacíos referenciales de las descripciones definidas o denotativas. Un vacío referencial se produce cuando no hay una entidad que satisfaga una expresión. En la polémica Russell-Strawson, la discusión se ceñía a los vacíos referenciales de las descripciones definidas, a lo que en lógica se denominan descripciones impropias. Pero la discusión se puede ampliar, y de hecho se ha ampliado, en la investigación contemporánea, a otras categorías.

El problema que tanto Russell como Strawson se planteaban era el de las consecuencias que dicho vacío referencial tenía sobre las oraciones en que entraban a formar parte como sujetos descripciones impropias. Mientras Russell se dejaba llevar por sus intereses lógicos, Strawson parecía más apegado al funcionamiento del lenguaje natural. El primero afirmaba que las oraciones en que entraban a formar parte descripciones impropias como sujetos eran falsas, el segundo que carecían de valor de verdad.

Como hemos dicho, el problema tiene dimensiones más generales; en primer lugar, no sólo las descripciones definidas pueden ser sujetos de oraciones sino que también pueden serlo nombres propios o comunes, pronombres, incluso oraciones. En segundo lugar, tal como lo plantean Russell y Strawson, parece que el problema se ciñe exclusivamente a las expresiones-sujeto. Esto podría deberse a una influencia de la concepción tradicional de la oración, que divide a ésta en dos categorías lógico-gramaticales: sujeto y predicado. Una inspección más detallada de la semántica oracional pone de relieve que las cosas son más complicadas. Por citar un ejemplo, las oraciones de identidad no pueden ser reducidas al esquema sujeto-predicado, como por otro lado reconocía el propio Russell, sino que se conciben mejor como enunciados relacionales diádicos, es decir, como fórmulas predicativas con dos argumentos. De hecho, en la investigación semántica contemporánea, se suele asumir que la forma lógica de las oraciones es precisamente la de fórmulas predicativas con uno o más argumentos. Por ejemplo 'Juan es amigo de Pedro' puede formalizarse Aab . De acuerdo con ello, cualquier vacío referencial que se produzca en uno de los argumentos de la oración tiene repercusión sobre la asignación del valor de verdad a ésta. Claro está que parece que la aportación del sujeto parece decisiva y no puede calibrarse con el mismo rasero que la de los demás argumentos, pero, aun así, el problema no es tan simple como lo veían Russell y Strawson.

Al hilo de los problemas conexos con los vacíos referenciales se ha llegado a la formulación de un problema general de gran interés lógico, lingüístico y epistemológico incluso. ¿Es suficiente una lógica bivalente para describir la semántica de los lenguajes naturales?

Los partidarios de una respuesta afirmativa tienden en general a sostener que las oraciones en que se producen vacíos referenciales son anómalas desde el punto de vista de la semántica, y es precisamente esa anomalía la causa de que no se les pueda asignar uno de los dos valores de verdad. Por su parte, los que argumentan su respuesta negativa rechazan que a estas oraciones no se les asigne, *de hecho*, en el lenguaje natural, un valor de verdad. Este valor de verdad puede ser el de ‘indeterminado’ o ‘neutro’, en cuyo caso nos encontramos ante la necesidad de una lógica trivalente. O pueden ser valores intermedios entre 1 (verdad) y 0 (falsedad), en cuyo caso nos vemos en la necesidad de apelar a una lógica polivalente de grado superior a la trivalente.

B. La opacidad referencial

La opacidad referencial, aunque es un caso especial de vacío referencial, merece un tratamiento aparte por muchas razones. La primera indicación de este fenómeno fue hecha por Frege, pero el término técnico fue introducido por W. O. Quine. Como hemos visto en 4.3 aquél ya había advertido que en determinados contextos ciertas expresiones no podían ser sustituidas por otras que tuviesen el mismo referente (que fuesen correferenciales) sin alterar el valor de verdad de las oraciones en que entraban a formar parte. A estos contextos los denominó indirectos, pero más adelante se han dado en llamar de *actitud preposicional*, justamente porque los más típicos se producen en relación con verbos que expresan la *actitud*, del hablante o de otros individuos, respecto al objeto verbal. Quine denomina a estos contextos opacos por oposición a los contextos *transparentes*. Cuando el objeto verbal es toda una oración, los contextos opacos se caracterizan por impedir la sustitución de esa oración por otra con el mismo valor de verdad. En este sentido, como quedó claro en 4.3, los contextos opacos constituyen una excepción al principio leibniziano de sustituibilidad *salva veritate*.

El estudio de los contextos opacos ha producido en el campo filosófico progresos notables. Entre ellos se puede citar la comprensión de la semántica de las expresiones en que aparecen los verbos ‘conocer’ y ‘saber’, lo que técnicamente se denomina «discurso epistémico». Consecuentemente, se ha desarrollado una rama de la lógica modal, la lógica epistémica, que trata de describir el funcionamiento semántico de este subconjunto de verbos condicionales. Con ello se ha logrado penetrar más profundamente en el conocimiento de la semántica de los lenguajes naturales, pero aún queda mucho por hacer; en especial, es necesaria una teoría general sobre el funcionamiento de los verbos de actitud proposicional y el desarrollo de nuevas y prometedoras ramas de la lógica modal. Las investigaciones de los lingüistas que han trabajado en este campo tienen el interés de que integran estos problemas lógico-semánticos en el marco de una teoría específica sobre el lenguaje, la generativo-transformatoria, en algunas de sus versiones.

5. INTENSIÓN Y SIGNIFICADO

5.1. Concepciones clásicas sobre la intensión

Como hemos visto en el capítulo anterior, G. Frege tenía una teoría semántica del significado en la que jugaban un papel central las nociones de sentido y referencia. Otros pensadores han asignado a nociones semánticas similares a las fregeanas una función importante en la teoría del significado. Entre ellos podemos destacar a R. Carnap, quien introdujo los conceptos de *extensión* e *intensión*.

Hablando a grandes rasgos, la extensión de una expresión, tal como aquí utilizamos el concepto, es la referencia que tiene cuando se la pone en uso en condiciones y contextos «nor-

males»: para la estrecha conexión entre extensión y referencia no será necesario seguir hablando aquí de la extensión de diversas expresiones lingüísticas, que tratamos en el capítulo anterior.

El término que hace pareja con el de extensión en la terminología carnapiana es el de ‘intensión’. Esta noción tiene precedentes que se remontan a los filósofos griegos y, especialmente, a los medievales. En efecto, en la tradición filosófica encontramos las nociones análogas de *comprensión*, de *connotación* y *sentido* (Frege). La diferencia en el caso del término ‘comprensión’ estriba en que éste se aplicaba a *conceptos*, que no son entidades lingüísticas, mientras que ‘intensión’ se aplica a entidades lingüísticas, a expresiones.

En su caso más simple y de un modo más bien general y vago, la noción de intención se puede explicar de un modo muy sencillo: la intención de un término o expresión general es el atributo o propiedad que ese término expresa, donde ‘atributo’ o ‘propiedad’ no se entiende en el sentido extensional como equivalente a la clase de todas las cosas a las que el término se aplica, sino como serie de características. Así, la intención del término ‘oro’ es la serie de características que comparten los objetos o los trozos de oro: quizás ser de un determinado color, poseer una dureza perfectamente especificada, etc.

Junto con el concepto de intención de un término o expresión singular el de intención de una oración es el de mayor importancia filosófica. La intención de una oración (declarativa) es el pensamiento o *idea* que esa oración expresa. Aquí no hemos de tomar ‘pensamiento’ o ‘idea’ en un sentido subjetivo, sino objetivo o intersubjetivo. El problema filosóficamente acuciante a que responde la noción de intención de una oración es el de saber cuándo dos o más oraciones «dicen» lo mismo, lo cual está directamente emparentado con el problema de saber cuándo dos personas dicen lo mismo, no en el sentido de que emitan las mismas palabras (ejemplares de las mismas palabras-tipo), sino de que el «contenido» de esa serie de palabras sea el mismo.

Reservamos el término ‘idea’ para la intención de una oración; en el caso de la intención de una expresión general no hablaremos de ‘idea’ sino de *concepto* (y, a veces, de *noción*). Así, aunque a veces se dice ‘la idea de mesa’, nosotros diremos ‘el concepto de mesa’. (Recuérdese que, para Frege, el concepto era la referencia -la extensión- y no la intención de un término general.)

La aplicación de la noción de intención a otras expresiones lingüísticas es secundaria. De todas formas, debemos mencionar que se toma como la intención de una expresión o término singular la serie de características que convienen al individuo que es la extensión de ese término singular, teniendo en cuenta el modo de designarlo. A esta serie de características la denominó R. Carnap -y la terminología parece conveniente- *concepto individual*. Este concepto hace posible diferenciar en cuanto a contenido ciertas expresiones que tienen la misma extensión: así, el concepto individual de la *torre Eiffel* es diferente al concepto individual de la *construcción más alta de la capital de Francia*.

Hasta aquí la teoría clásica y «ortodoxa» de la intención que podemos resumir del siguiente modo:

Tipos de expresiones	Intensión correspondiente
Términos individuales Términos generales Oraciones	Conceptos individuales Conceptos Ideas

De todos modos, esta teoría es bastante oscura; se apoya en conceptos como el de serie de características y el de idea. Pero ¿qué es exactamente una característica? y ¿qué es una idea objetiva o intersubjetiva?

5.2. La precisión de los conceptos intensionales

Muchos han sido los filósofos y lógicos que, insatisfechos por la caracterización clásica de los diversos tipos de intención, han tratado de sustituirla por otra más precisa. Entre los pioneros de esta labor destacan A. Church y el propio R. Carnap. El gran crítico de los filósofos que piensan poder clarificar estas nociones intensionales es el lógico y filósofo norteamericano W. O. Quine. La labor de Church y de Carnap ha conducido a precisiones matemáticas de los conceptos intensionales que se desarrollan plenamente en la obra de R. Montague y de D. Lewis. En el corazón de estas precisiones está el concepto de *mundo posible* introducido en la moderna filosofía del lenguaje por S. Kripke. Según las precisiones mencionadas las intensiones tanto de una expresión singular como de una expresión general y de una oración son funciones (en el sentido matemático del término). La intención de una expresión singular es una función del conjunto de los mundos posibles en el conjunto de los individuos o particulares existentes en esos mundos, es decir, una función que asigna a cada mundo posible un individuo de ese mundo; la intención de una expresión general es una función que asigna a cada mundo posible un conjunto (o una relación) en ese mundo; por último, la intención de una oración (declarativa) es una función que asigna a esa oración un valor de verdad en cada mundo posible.

Estos conceptos intensionales así precisados producen al principio una notable sensación de extrañeza. ¿A qué viene toda esa chá-chara sobre mundos posibles? Vamos a intentar explicarla al menos para el caso de la intención de las oraciones y de los términos generales. El punto de vista intuitivo en el que se sitúan los autores citados es más o menos el siguiente: ¿qué quiere decir que sabemos cuál es el «contenido» de una oración (declarativa)? Quiere decir, ni más ni menos que, para cualquier situación que quepa imaginar sabríamos decir, en principio, si esa oración es verdadera o falsa. ¿Qué quiere decir que sabemos que el «contenido» de la expresión general ‘criatura con riñones’ es distinto del «contenido» de la expresión general ‘criatura con corazón’? Pues que, si bien ambas expresiones se aplican en el mundo real a las mismas cosas, cabe imaginar otras situaciones, otros «mundos alternativos» en que hay criaturas con riñones que no tienen corazón o a la inversa.

Pues bien, un mundo posible no es sino una de esas situaciones imaginables, o mundos alternativos imaginables (excepto en que, claro está, el mundo real se considera también un mundo posible). Ese saber decir, respecto a una oración, para cada situación imaginable (mundo posible), si es verdadera o falsa se capta matemáticamente con la idea de una función que asigna a cada mundo posible un valor de verdad. Y ese saber que el «contenido» de la expresión ‘criatura con riñones’ es distinto del de ‘criatura con corazón’ porque somos capaces de imaginar mundos alternativos (mundos posibles) en que hay criaturas que po-

seen una cosa pero no la otra se capta matemáticamente diciendo que a esas expresiones les corresponden dos funciones *distintas* tales que a cada mundo posible asignan un conjunto de criaturas no siempre idéntico.

Sin duda, el paso de las nociones intuitivas a las precisiones matemáticas supone una mayor abstracción, pero estas últimas tienen la ventaja de que tienen propiedades bien definidas y podemos manejarlas con exactitud. Sean bienvenidas en la medida en que nos sirvan para resolver problemas de más bajo nivel de abstracción interesantes para nosotros, problemas como ¿qué es lo que cree -cuál es el objeto de la creencia- una persona cuando decimos de ella con verdad que cree algo? ¿Cuándo dos personas creen o desean lo mismo? ¿Cuándo expresan la misma idea dos oraciones?, etcétera.

Si M es la clase de los mundos posibles, M_1, M_2, \dots son sus elementos, es decir, los diversos mundos posibles; D_{M_1}, D_{M_2}, \dots los conjuntos de individuos o particulares existentes en esos mundos; R_{M_1}, R_{M_2}, \dots los conjuntos de las relaciones (monádicas, diádicas, etc.) de individuos de los mundos M_1, M_2, \dots , respectivamente; y V, F , como de costumbre, los valores semánticos de verdad y falsedad, podemos resumir así la teoría matemática de la intensión:

Tipos de expresiones	Tipo de intensión correspondiente
Términos singulares	$f: M \rightarrow UD_{M_i}, \quad i \geq 1$ (para cada término singular t , $f_t(M_i) \in D_{M_i}$)
Términos generales	$g: M \rightarrow UR_{M_i}, \quad i \geq 1$ (para cada término general P , $g_p(M_i) \in R_{M_i}$)
Oraciones	$h: M \rightarrow \{V, F\}$

5.3. Semántica y teoría del significado

Las teorías de la extensión y de la intensión constituyen el campo de lo que en un sentido estricto se llama ‘semántica’, si les añadimos la denominada ‘semántica de lexemas’ o ‘lexicología’. A veces, sin embargo, el término ‘semántica’ tiene un significado más amplio, equiparable al de ‘teoría del significado’. La terminología más clara sigue siendo probablemente la de Ch. Morris que explicamos en el capítulo 2. Ampliaremos algunas de las observaciones que allí hacíamos.

Los estudios semánticos hacen abstracciones notables a la hora de estudiar de qué hablan las expresiones del lenguaje. En un sentido estricto, las expresiones lingüísticas no hablan de nada; son los hablantes los que hablan de algo al proferir en ciertos contextos tales expresiones. Sin embargo, la mayoría de los estudiosos están de acuerdo en la conveniencia de realizar esta abstracción y buena parte de ellos, fijándose quizás en el ejemplo de las ciencias naturales, donde nociones muy abstractas son totalmente fructíferas, son partidarios de una cierta preeminencia epistemológica de las nociones semánticas sobre las pragmáticas (también históricamente las primeras se desarrollaron antes que las últimas y ello explica en parte que las nociones semánticas estén más claramente delimitadas que las pragmáticas). En algunos autores, esto se llevó al extremo de querer *definir* las nociones pragmáticas a partir de las semánticas. Otros, sin embargo, piensan lo contrario, que hay algún tipo de prioridad epistemológica de las nociones pragmáticas sobre las semánticas y los más radicales de entre ellos creen posible introducir las nociones semánticas a partir de las pragmáticas (casualmente tenemos un ejemplo de esto en la caracterización que hemos

hecho en el primer apartado de la noción semántica de extensión a partir de una noción pragmática de referencia; sin embargo no pretendemos con esa caracterización sino intentar una pequeña clarificación terminológica y no adoptaremos en consecuencia el punto de vista general mencionado). Un tercer grupo de estudiosos del lenguaje no ve conflicto o preeminencia alguna ; para ellos se trata simplemente de armonizar los aspectos semánticos y los pragmáticos; ambos aspectos nos descubrirían facetas del *significado* de las distintas unidades lingüísticas. Los aspectos semánticos y pragmáticos serían pues partes de una teoría general del significado. Si se quiere se puede hablar de ‘teoría semántica del significado’ y de ‘teoría pragmática del significado’ para designar las partes semántica y pragmática del estudio del significado, aunque quizás sería más claro reservar esos términos para las teorías del significado de los filósofos, lógicos o lingüistas que confieren preeminencia a los aspectos semántico y pragmático, respectivamente.

Tras la discusión precedente podrá apreciarse lo confuso que resulta hacer equivaler sin más ‘teoría del significado’ a ‘teoría semántica’ y ‘significado’ a ‘semántica’. Ejemplos de esa confusión se encuentran en multitud de estudiosos del lenguaje. Sólo están autorizados a hacer esa equivalencia los que piensen que se puede explicar el significado sólo (o principalmente) a partir de nociones semánticas. Un caso extremo quizás sea el del primer B. Russell quien tenía una noción del significado totalmente equiparable a la de extensión (Russell utilizaba, como vimos, el término ‘referencia’, pero nosotros preferimos reservar ese término, siguiendo a Strawson, para una noción pragmática).

5.4. Sinonimia y congruencia

Es corriente que hablemos u oigamos hablar de *lo* que una expresión (o proferencia) significa o no significa, de cuál es *el* significado de una expresión (o proferencia). Como muy bien argumentó L. Wittgenstein el uso del artículo determinado (*el* significado) o de formas pronominales (*lo* que significa) nos induce a pensar que el significado de una expresión (o de una proferencia) es algo, por así decir, perfectamente especificable, con límites bien definidos y que posee, por tanto, algunas de las características que normalmente atribuimos a los objetos físicos. Contra esa «cosificación» de los significados -de la que podemos decir que participa Frege, padre de la moderna filosofía del lenguaje- han argumentado convincentemente Wittgenstein y Quine. Quine afirma, como conclusión de sus argumentos contra la concepción de un significado «cosificado», que el objetivo de una teoría del significado no puede ser el estudiar un cierto tipo de objeto peculiar al que llamamos ‘significado’, puesto que tales objetos son un mito, sino, en todo caso, el de las diversas *relaciones* entre los significados de las expresiones y, especialmente, la relación de *sinonimia* (o igualdad de significado) de las formas o expresiones lingüísticas.

La relación de sinonimia entre expresiones cobra gran importancia filosófica y extrafilosófica si se la considera como parte de una investigación más amplia: la de dar criterios para averiguar cuándo dos hablantes -o dos textos- dicen lo mismo (en cuanto a «contenido») independientemente de las palabras que emplean e incluso de la lengua que utilicen, o si lo que dice uno de ellos está implicado por lo que dice el otro.

Las propuestas contemporáneas al respecto de la sinonimia de expresiones lingüísticas se reducen a decir que dos expresiones son sinónimas (es decir, tienen el mismo significado) si y sólo si tienen la misma intensión. Esto tiene tres graves inconvenientes. .

El primer inconveniente es que el valor explicativo y epistemológico del concepto de intensión está en entredicho. Como vimos en el apartado anterior se dispone de una teoría matemáticamente impecable de las intensiones, pero su valor empírico depende de que consi-

gamos dar una interpretación aceptable del concepto básico de mundo posible o situación imaginable. Desde luego, nada es más fácil de concebir que *una situación* imaginable, pero el *concepto* mismo de situación imaginable es bastante oscuro. ¿Es imaginable que Reagan no fuese ahora (1981) el presidente de los Estados Unidos? Seguramente sí. ¿Es imaginable que fuese en cambio secretario del P.C.U.S.? ¿Es imaginable que hubiese nacido de otros padres -no sólo que se averiguase que otros son sus auténticos padres-? ¿Seguiría, en tales casos, siendo la *misma persona*? Hay multitud de preguntas embarazosas relativas a situaciones imaginables, es decir, a mundos posibles.

El principal crítico del concepto de mundo posible no podía ser otro que Quine quien, como hemos visto, ha rechazado siempre las nociones intensionales como demasiado oscuras para propósitos científicos. Kripke y otros filósofos del lenguaje tratan en la actualidad de salvar estas nociones perfilando un concepto «sensato» de mundo posible. Esa cuestión puede considerarse abierta en la actualidad.

El segundo inconveniente del criterio de sinonimia como identidad de intensión es un inconveniente técnico: según ese criterio todas las oraciones equivalentes lógicamente serían sinónimas; para salvar esta dificultad propuso Carnap reforzar el criterio y exigir para la sinonimia además de la igualdad de intensiones la isomorfia (misma forma) estructural. Este criterio ha sido recientemente refinado por D. Lewis, pero no podemos explicar técnicamente aquí este punto.

El tercer inconveniente es que el criterio identifica prácticamente ‘significado’ con ‘intensión’. Como veremos en el siguiente apartado hay poderosos argumentos en contra de tal identificación.

Un enfoque totalmente original relativo al problema de la sinonimia se debe al lógico y filósofo de la ciencia norteamericana P. Suppes. El enfoque consiste en considerar la noción de sinonimia como poco interesante o prometedora desde un punto de vista científico y en sustituirla por una serie de nociones de *parentesco de significado* totalmente precisas. Suppes puede hacer estas precisiones porque primeramente ha descrito un aparato semántico preciso para transformar los árboles de descripción sintáctica (que pueden verse en el capítulo 3. 3). No entraremos ahora en muchos detalles (véase el capítulo 7.2 si se desea mayor información sobre este punto), pero un árbol semántico se consigue asignando extensiones a las palabras que se encuentran en los nodos terminales (cada uno de los puntos de ramificación del árbol es un nodo; los nodos de los que no parte ninguna ramificación son nodos terminales) y combinando estas extensiones en los nodos de nivel superior hasta formar una extensión total en el nodo origen, es decir, la raíz del árbol (recuérdese que tal como se los representa normalmente se trata de árboles invertidos). Las extensiones que se asignan a las palabras terminales son relativas a un mundo posible (individuos de ese mundo, subconjuntos del dominio de individuos, relaciones entre individuos, etc.), de modo que variando el mundo posible se obtiene una denotación diferente para todo el árbol. Si consideramos las extensiones en todos los mundos posibles tenemos, de nuevo, la intensión. Sin embargo, Suppes utiliza un concepto restringido de mundo posible, pues se encuentra entre los que piensan que si el concepto de mundo posible ha de ser fructífero para el análisis semántico de las lenguas naturales ha de restringirse su alcance. Así, por ejemplo, opina que ha de mantener algunas teorías matemáticas «fijas», es decir, que sólo admitimos como posible un mundo en el que al menos algunas de las verdaderas matemáticas de nuestro mundo -por ejemplo, las de la aritmética- siguen siendo válidas.

Sobre la base de este aparato semántico preciso Suppes define varios conceptos de parentesco (o *congruencia*, como él los denomina siguiendo la terminología geométrica) en el

significado de las expresiones. Mediante estas definiciones es posible clasificar los diversos parentescos en el significado de las siguientes oraciones:

- (1) Algunos diputados han cambiado de partido político.
- (2) Alguns diputats han canviat de partit polític.
- (3) Más de un diputado ha cambiado de partido político.
- (4) Ha cambiado de partido político más de un diputado.
- (5) Hay al menos dos diputados que han cambiado de partido político.
- (6) Los diputados que han cambiado de partido político son por lo menos dos.
- (7) Han cambiado de partido político al menos dos diputados.
- (8) More than one deputy has changed of political party.

El parentesco que se da entre (1) y (2) por un lado y (3) y (8) por otro es el que Suppes llama parentesco ‘fuerte’. Al que existe entre (3) y (4) lo denomina ‘permutacional en forma y contenido’. Al existente entre (4) y (8) ‘permutacional en contenido’. Por último, todas esas oraciones (1)-(8) tienen un parentesco: son *paráfrasis* unas de las otras. Aunque no podemos dar aquí las definiciones exactas de estos conceptos esperamos que con los ejemplos sea suficiente para captar la idea fundamental que se esconde tras ellos.

Naturalmente se trata sólo de un comienzo. Se ha abierto una nueva línea de investigación y, en principio, parece posible delimitar nuevas nociones de parentesco entre expresiones (las definiciones existentes son también aplicables a expresiones distintas de oraciones).

Como el propio Suppes señala, lo más interesante sería definir rigurosamente conceptos de parentesco no entre expresiones, sino entre preferencias o entre actos de habla (veremos lo que son estos actos de habla en el capítulo 9). La mayor dificultad para la realización de esta idea estriba en que no existe un aparato conceptual tan desarrollado para las nociones pragmáticas como en el caso de las nociones semánticas (recuérdese que las definiciones de Suppes sólo son posibles porque dispone de nociones precisas como las de extensión, árbol semántico, etc.). Se trata pues de una interesante tarea a desarrollar en el futuro.

5.5. Intensión y significado

Como hemos visto en el apartado anterior la teoría de la intensión puede jugar un papel importante en una teoría del significado: puede desempeñar una función importante a la hora de hacer explícitas nuestras intuiciones sobre parentescos de significado, por ejemplo. Pero esto no quiere decir que debamos pensar que las nociones de intensión y significado son intercambiables o que la primera es un sustituto más preciso de la segunda.

Esta identificación entre intensión y significado es todavía hoy moneda corriente entre muchos filósofos del lenguaje, mucho más que la identificación de significado y extensión que veíamos en *Russell*. Sin embargo, hay varios argumentos poderosos contra ella. El más general es que la noción de significado tiene muchas veces un carácter pragmático más que semántico. Cuando preguntamos ¿qué quieres decir? o ¿qué significa lo que has dicho? estamos preguntando por el significado de una preferencia hecha por una determinada persona en un determinado contexto extralingüístico, no por el significado de una expresión.

Recientemente, el destacado filósofo de la ciencia y del lenguaje H. Putnam ha llamado la atención sobre ciertos extremos que están en contra de la identificación del significado con la intensión.

En primer lugar hay que fijarse en la gran fuerza del concepto de intensión, tanto en su caracterización tradicional como en su caracterización actual. Si la intensión de un término

general, por ejemplo, es la sede de las características comunes de las cosas a las que se aplica ese término, el conocimiento de esas características -de la intensión- posibilita el conocimiento de la aplicación o la no aplicación con propiedad de ese término a cualquier cosa. Por ejemplo, si sabemos cuáles son las características que comparten objetos a los que llamamos 'oro', cuando se nos presente un objeto sabremos decir si se trata o no de oro, es decir, si le «conviene» (como se dice tradicionalmente) o no el término 'oro'. Dicho de otro modo, si conocemos la intensión de un término somos capaces de fijar la extensión del mismo. La intensión determina la extensión (al menos en el caso de los términos generales). ¿Qué sucede en cambio con el significado? Sencillamente no digamos que una persona no sabe lo que significa 'oro' porque no sea capaz de identificar un objeto determinado; una persona puede confundir el oro con el llamado oro alemán o con el simple dorado de un objeto, y confundir la plata con la alpaca e incluso con el estaño, pero no por eso diríamos que no sabe lo que significan los términos 'oro' y 'plata'. Se puede, pues, saber el significado de una expresión sin saber su extensión. Por el mero hecho de conocer el significado de una expresión, una persona no es capaz normalmente de fijar su extensión, mientras que sí lo sería si conociese su intensión (al menos en el caso de los términos generales); esto habla indudablemente en contra de la identificación de las nociones de significado e intensión. ¿Cómo se fija en realidad la extensión de una expresión? Sencillamente no son los ciudadanos «normales» los que la fijan (por más que sepan lo que significa). La extensión de un término la fijan los expertos,- por ejemplo, la extensión de los términos 'oro' y 'plata' la fijan los joyeros o, en última instancia, los químicos. Tenemos aquí lo que Putnam ha denominado división del trabajo lingüístico. Como recalca Putnam, éste es un aspecto bastante obvio sobre el que las teorías del significado, la intensión y la extensión no habían llamado la atención hasta ahora.

El siguiente punto de Putnam se puede captar bastante fácilmente si nos preguntamos lo siguiente: cuando alguien nos pregunta por el significado de una expresión ¿cómo le respondemos? ¿cómo respondemos a los niños cuando nos preguntan qué significa tal o cual palabra, por ejemplo 'oro' o 'tigre'? Simplemente les damos unos cuantos rasgos o características que suelen tener los objetos de oro o los tigres. Por ejemplo: 'el oro es un metal blando, brillante, de color amarillo y muy caro'; 'el tigre es un animal del tamaño de un león, con rayas en el cuerpo y fuertes garras y colmillos'. Esas características que damos en la explicación del significado de un término constituyen lo que Putnam llama estereotipos. No hay que confundirlos con las intensiones pues éstas incluyen las características que todos los objetos que caen bajo un mismo término tienen, de tal modo que a un objeto se le aplica el término si y sólo si posee esas características. En nuestros ejemplos, las características mencionadas (metal brillante, muy caro, etc.; con fuertes garras, rayado, etcétera) no son de este tipo: un trozo de metal puede no ser brillante (por estar sucio) o ser barato y ser de oro; un animal puede no tener garras fuertes y no estar rayado y ser un tigre (un tigre albino con una enfermedad en las extremidades sería un tigre). En resumen: cuando explicamos el significado de una expresión (pensamos aquí, sobre todo, en los términos generales) no damos la intensión sino un estereotipo.

En lo anterior hemos tenido presente las versiones tradicionales de la teoría de la intensión; pero lo mismo valdría con las versiones actuales. Vamos a utilizar estas versiones para explicar la última crítica de Putnam a la identificación de significado con intensión. Si la intensión de una expresión es una función de mundos posibles a extensiones y la intensión se identifica con el significado lo mismo dará decir que se ha captado el significado de una expresión que decir que se ha *captado la intensión*. Pero ¿qué quiere decir 'captar una in-

tensión’, o lo que es lo mismo, ‘captar una función de (todos) los mundos posibles a extensiones’? Desde un punto de vista psicológico resulta un misterio la explicación de un proceso «mental» descrito así: ‘una persona capta una determinada función de todos los mundos posibles en extensiones’. No parece haber muchas perspectivas de que esa explicación pueda ser dada. De todos modos, estamos aquí ante problemas sobre los que, evidentemente, aún no se ha dicho, ni mucho menos, la última palabra. En el capítulo 7 presentamos un tipo de semántica que tiene en cuenta algunos de los aspectos de ‘proceso mental’ que se hayan involucrados en la comprensión y uso del lenguaje.

5.6. La relevancia filosófica de la noción de intensión

Aparte de la trascendencia que pueda tener la noción de intensión para una teoría del significado y por tanto, indirectamente, en muchas cuestiones filosóficas ligadas a los problemas del significado, ese concepto está involucrado de un modo más directo en problemas epistemológicos y ontológicos. En efecto, la noción de intensión de un término es la contrapartida en el lenguaje de la importante noción tradicional de *esencia*. Veamos qué se quiere decir con esto.

La esencia de una cosa o un conjunto de cosas es el conjunto de propiedades o atributos que hacen a esa o esas cosas ser lo que son, es decir, que cualquier cosa que carezca al menos de una de esas propiedades o tenga al menos una más es una cosa diferente o un tipo diferente de cosa. Los filósofos que creen posible clarificar y utilizar esta noción la han utilizado con diferentes propósitos. Aristóteles, a quien debemos la noción más o menos en la forma en que se ha presentado (un tanto simplificada), la utilizaba para explicar el cambio. Las cosas cambian continuamente de muchas maneras (de posición y lugar, de color, menguan y aumentan, etcétera) pero no pasan a ser otras cosas *diferentes* hasta que no pierden (o adquieren) alguna propiedad esencial.

Según esta doctrina tradicional de la esencia son las cosas las que tienen esencia. Según la moderna teoría de la intensión son las *expresiones lingüísticas* que se refieren a, y cuya extensión es, esas cosas las que tienen intensión. Por lo demás, las nociones de esencia e intensión parecen dos caras de la misma moneda. Veámoslo.

Parece que cada pregunta que hagamos sobre la esencia de una cosa o cosas puede transformarse en una pregunta sobre el término cuya extensión es esa cosa o cosas. Por ejemplo, la caracterización que hemos hecho de la noción de esencia se transforma fácilmente en una «definición» de la de intensión. La intensión de un término es el conjunto de propiedades o atributos que una cosa o un conjunto de cosas deben poseer para que se les aplique con propiedad ese término. Así, la esencia de los leones es el conjunto de propiedades o atributos que hacen que algo sea un león y no, por ejemplo, un tigre, un jaguar o un elefante, y la intensión del término general ‘león’ es el conjunto de propiedades o atributos que hacen que a algo se le llame con propiedad ‘león’.

Se puede establecer, pues, una estrecha analogía entre las nociones de esencia de una cosa o cosas y de intensión de un término. Quizás el primer filósofo que se dio cuenta de esto fue John Locke. Por supuesto Locke no utilizaba el término ‘intensión’ sino el de ‘esencia nominal’; tampoco corresponden exactamente nuestras intensiones con las esencias nominales de Locke: esto es debido a la teoría de las ideas imperante en los empiristas británicos de los siglos XVII y XVIII (y no sólo entre ellos); pero éste es un extremo sobre el que no vamos a entrar aquí. Después de Locke, un gran número de filósofos empiristas y pragmatistas (por ejemplo, J. Dewey, C. I. Lewis y W. O. Quine) se han apresurado no sólo a señalar la analogía sino a afirmar que en realidad la noción de esencia de una cosa o cosas se

deriva totalmente de la de intensión de un término, de modo que la primera depende lógica y epistemológicamente de la última. En otras palabras, los filósofos esencialistas atribuyen a las cosas propiedades que en realidad sólo tienen sentido -si es que lo tienen- en relación con los términos que a ellas se refieren: las supuestas propiedades esenciales de las cosas sólo revelan peculiaridades de nuestro vocabulario. Y éstas, a su vez, intereses de comunidades de hablantes.

La respuesta de un filósofo esencialista sería más o menos la siguiente: hay ciertamente relatividad a peculiaridades del vocabulario y, en último término, a intereses sociales, lo cual puede admitirse en el caso de las *esencias nominales* de Locke (u otro concepto similar). Pero el mismo Locke las distinguía de las esencias reales; éstas no revelan peculiaridades del vocabulario sino la verdadera naturaleza de las cosas y, al contrario de lo que Locke pensaba, son cognoscibles: el conocimiento científico es precisamente un intento de conocer esas esencias reales, como ya sostenía Aristóteles. Por supuesto que el hecho de atribuir las propiedades *amarillo* o *caro* al oro revela peculiaridades e intereses humanos relacionados con la vida cotidiana, pero no sucede lo mismo si decimos que las propiedades esenciales del oro son tener tal y cual estructura atómica, el número atómico 79, un peso atómico de 196,967, etc.

Esta es una buena jugada del esencialista; el esencialista no está necesariamente en contra de la identificación de esencia con intensión; sólo distingue entre dos tipos de lenguaje: el lenguaje cotidiano en el que los términos tienen su intensión ordinaria (equivalente si se quiere a «esencias nominales») y el lenguaje científico en que los términos tienen una nueva intensión relacionada con su esencia real. El esencialista redondea así una postura epistemológica realista.

Sin embargo, el antiesencialista no carece de posible réplica. Puede alegar que las nuevas esencias, las «esencias reales» no están menos ligadas a intereses humanos, sólo que esta vez se trata de intereses peculiares, inherentes a la empresa científica. A ésta pertenece la búsqueda de leyes, las más generales y simples posibles, y a esta búsqueda está supeditada la agrupación de objetos bajo un mismo término del lenguaje científico. La ciencia es un instrumento de explicación, predicción y elaboración de aplicaciones prácticas y con estos objetivos se formulan las teorías científicas y los conceptos que éstas utilizan. La posición final del antiesencialista es, a menudo, pragmatista, convencionalista e instrumentalista.

Con esto queda desplazado el campo de la disputa a un lugar nuevo, donde lo que se enfrentan son diversas concepciones epistemológicas, distintas filosofías de la ciencia. Este desplazamiento de la cuestión es frecuente en filosofía del lenguaje: muchas veces lo que empieza siendo una cuestión relacionada directamente con el lenguaje se muestra, al ser tratado cuidadosamente, como una cuestión estrechamente ligada a concepciones de otras áreas de investigación filosófica.

6. VERDAD Y SEMÁNTICA

6.1. Introducción

Hay muchos sentidos en los que se dice de algo que es verdad o verdadero. Sólo alguno de estos sentidos interesa a la filosofía del lenguaje, principalmente el sentido en el que se asevera de una entidad lingüística (un objeto que la teoría del lenguaje estudia) que es verdadera. La mayor parte de nuestro discurso habitual se refiere a hechos, exteriores o interiores a nosotros mismos, y en relación con la forma en que nos referimos a estos hechos utilizamos la expresión ‘... es verdad’. Ahora bien, hay varios modos de entender esta relación entre

entidades lingüísticas y hechos, cada una de las cuales ha configurado un tipo de teoría de la verdad. Pero, antes de pasar a exponer brevemente los principales tipos de teoría de la verdad mantenidos por los filósofos del lenguaje, es preciso aclarar unas cuantas cuestiones previas.

En primer lugar, antes de examinar los diferentes modos en que se concibe el predicado 'es verdad' hay que determinar a qué tipo de entidades lingüísticas se refiere este predicado. En filosofía del lenguaje se han barajado fundamentalmente tres alternativas:

(1) 'es verdad' se aplica a oraciones declarativas o, más precisamente, a preferencias de oraciones declarativas (o a sus resultados) ; (2) se adjunta a enunciados. Dada la caracterización variable de la noción de enunciado, dependiendo del nivel gramatical en que se esté formulando su definición, esta postura nunca se ha distinguido por su claridad. No obstante, lo que parecen tener en mente los que sostienen que la verdad es una propiedad de los enunciados, es una caracterización pragmática de esta noción. Lo que es verdadero o falso es lo que *se hace* al proferir una oración declarativa, es decir, por lo general una afirmación o aserción. (3) La verdad es una propiedad de las ideas o proposiciones, esto es, de lo significado por las oraciones declarativas (por sus preferencias). Según esta postura, dos oraciones sinónimas, por el hecho de expresar la misma idea, han de tener el mismo valor de verdad si se profieren en un mismo momento y con parecidas intenciones por parte de sus proferentes. Dos preferencias de un mismo tipo de oración, si tienen valores de verdad distintos es porque expresan ideas diferentes.

Se ha argumentado largamente en favor de una u otra postura en filosofía del lenguaje. No siempre se ha llevado la polémica de modo muy preciso, pues en ocasiones no se acertaba a distinguir las nociones de oración-tipo, preferencia de oración y enunciado, de modo adecuado. El problema de fondo, como tantos otros en la teoría filosófica del lenguaje, parece ser el de la admisión o rechazo dentro de la teoría de entidades conceptuales independientes del lenguaje, como las ideas. En general, quienes admiten la legitimidad y necesidad de las ideas dentro de la teoría lingüística sostienen que son estas entidades las portadoras de la verdad. Quienes, por el contrario, rechazan la independencia de esas entidades abstractas respecto a las concretas, las lingüísticas, tratan de retrotraer la noción de verdad a las oraciones o a las preferencias de oraciones.

Por el momento, y sin que esto constituya una toma de postura definitiva, vamos a considerar que la verdad es una propiedad que se predica de hechos lingüísticos concretos, de preferencias de oraciones declarativas efectuadas en un determinado contexto, y más adelante podremos averiguar si esta posición teórica es suficiente o no para nuestros propósitos.

Finalmente, se han de plantear de antemano las características que ha de reunir una teoría para que se pueda afirmar de ella que proporciona una *definición* del predicado 'es verdad'. Es decir, se ha de poner sobre el tapete la distinción entre lo que es una *definición de verdad* y lo que en cambio constituye un *criterio para la verdad*. En algunas teorías que tratan de este concepto la confusión o identificación intencionada entre definición y criterio tiene importantes consecuencias. Así sucede, por ejemplo, en las teorías pragmatistas sobre la verdad o en las teorías verificacionistas. Según las primeras (mantenidas entre otros por Ch. Peirce, W. James y J. Dewey) la propiedad de ser verdadero se puede reducir a otras ('es útil', 'es conveniente', etc.) para cuya aplicación (a hechos o creencias) se da un conjunto de criterios prácticos. De acuerdo con las segundas, 'es verdad' equivale a 'es verificable' y esta propiedad se define también en términos de criterios empíricos.

El contraste entre la noción de definición y la de criterio se suele establecer como sigue: mientras que la definición de 'es verdad' especifica el significado de esta expresión predi-

cativa, el criterio únicamente proporciona una regla o conjunto de reglas que nos permiten decidir cuándo una oración es verdadera o no, cuando se le puede aplicar con corrección el predicado en cuestión. Un ejemplo aplicado a un predicado distinto del de ‘verdadero’ aclarará la cuestión: se puede conocer el significado, la definición, de la expresión predicativa ‘es de 18 quilates’ sin conocer los medios que se han de utilizar para comprobar si ese predicado se aplica a cualquier objeto individual x con corrección, esto es, se puede no saber cuáles son los *criterios* que permiten decidir si algo es de 18 quilates o no, qué tecnología particular se emplea en la joyería para ello, cómo se utilizan los aparatos de medición de la pureza de determinados metales, etc.

En teoría filosófica del lenguaje la distinción entre definición y criterio de verdad es un trasunto de la separación entre ontología y epistemología, entre lo que objetivamente acontece, lo que constituye los hechos, y los modos en que los conocemos. Como es bien sabido, para cualquier estudioso de la filosofía, tal separación no es siempre fácil de trazar, e incluso hay tesis filosóficas que mantienen la imposibilidad de tal división. Sin embargo, admitiremos, aunque sólo sea a título provisional, la existencia de tal división admitiendo que se puedan hacer preferencias de oraciones declarativas verdaderas sin que el hablante conozca los criterios de verdad de sus afirmaciones.

Aunque en ocasiones se ha considerado que la teoría de la verdad de Alfred Tarski es una teoría criteriológica, el propio Tarski y otros importantes autores la consideran como genuinamente definicional. Como veremos, la teoría de Tarski proporciona un conjunto de reglas recursivas que definen para un lenguaje formal el predicado ‘es verdad’. Esas reglas no son criterios para la aplicación del predicado en cuestión, sino que constituyen una auténtica definición del significado de ‘es verdad’ para un lenguaje formalizado. En cambio, las teorías pragmatistas son explícitamente criteriológicas, en la medida en que todas ellas, de un modo u otro, identifican el significado de un predicado con los criterios que determinan su correcta aplicación.

6.2. Teorías de la verdad como coherencia

La teoría de la verdad como coherencia fue una vez característica de los sistemas filosóficos idealistas y está resumida en el dictum de Hegel: la verdad es el todo. Pero, aun siendo importantes las conexiones de las teorías de la verdad como coherencia con tesis metafísicas como las de Hegel, Bradley o Lotze, más importantes para nosotros son sus aspectos lógicos y lingüísticos. Desde este punto de vista, es interesante observar que la teoría de la verdad como coherencia tuvo su origen en algunos teóricos de una escuela filosófica opuesta al idealismo, el positivismo lógico. En los primeros momentos de su constitución, el positivismo lógico mantuvo una teoría de la verdad como correspondencia con los hechos, en algunos puntos similar a la expuesta por L. Wittgenstein en su *Tractatus Logicus Philosophicus*. Esta teoría era, o quería ser, específicamente criteriológica, esto es, pretendía hallar y formular los criterios epistemológicos y metodológicos no sólo de los enunciados verdaderos, sino también de los enunciados con significado.

El criterio de verificabilidad empírica propuesto por los positivistas dividía por lo tanto a los enunciados en significativos y asignificativos. Entre los primeros, algunos positivistas, como Carnap y Schlick, consideraron que, en cuanto a su modo de confrontación con la experiencia, se dividían a su vez en dos clases: enunciados directamente contrastables con la experiencia inmediata, porque expresan datos de nuestros sentidos, y enunciados cuya relación con la experiencia es más compleja, pues son enunciados pertenecientes a una teoría que expresa un conocimiento abstracto. Estos últimos se relacionan con la experiencia

mediante las conexiones lógicas que mantienen con los enunciados de percepción, protocolarios.

Posteriormente, a medida que el positivismo fue madurando sus posiciones, se puso en cuestión la existencia de auténticos enunciados protocolarios. En particular fue Otto Neurath junto con K. Popper) quien cuestionó la incorregibilidad de los enunciados protocolarios: todo enunciado es portador de una cierta carga teórica y, por tanto, no puede confrontarse directamente con la experiencia, sino que en realidad depende de las relaciones lógicas que mantiene con otros enunciados.

O. Neurath fue el primer positivista que defendió una teoría de la verdad como coherencia y quien primero formuló los dos requisitos básicos de una teoría así: la consistencia y la comprensión. Posteriormente, la teoría de la verdad como coherencia ha sido mantenida por metodólogos influidos especialmente por teorías formales o formalizadas. Según éstos, la verdad o falsedad de un enunciado depende de la relación que ese enunciado mantiene con otros enunciados, perteneciendo todos a una teoría. En general, se denomina a esa relación coherencia o consistencia, pero si se está tratando de teorías científicas, especialmente de las axiomatizadas, es más frecuente denominarla deducibilidad o incluso implicación.

En este tipo de teorías, un enunciado es verdadero si se encuentra en la adecuada relación de implicación con otros enunciados. Cuando este sistema de enunciados pertenece a una ciencia empírica, su contrastación con la realidad se realiza en bloque, pues el fallo de cualquier enunciado lleva consigo el de todos los demás, o al menos de una buena parte. Cuando el enunciado pertenece a una ciencia formal, su verdad o falsedad viene dada por la relación de deducibilidad que mantiene con un cuerpo de afirmaciones asentadas pertenecientes a una teoría en vigor; si la teoría en cuestión está axiomatizada, la relación tiene en el conjunto de esos axiomas uno de sus elementos, siendo el otro el enunciado a aceptar o rechazar. En estos sistemas, son las relaciones internas las que constituyen la teoría, puesto que cualquier alteración en ellas supone una alteración esencial en la teoría misma. Por ejemplo, el rechazo del principio del tercio excluido por parte de la escuela intuicionista impuso el rechazo de la teoría «ortodoxa» de la lógica y de la matemática, puesto que entrañó el rechazo de todos los enunciados deducibles a partir de, o con ayuda de, este principio.

La especial aplicación de esta teoría de la verdad a sistemas formales ha sido la causa de que se considere como una especie de teoría de las *verdades analíticas*, esto es, de las verdades que lo son en virtud de relaciones internas. Esta imputación está de acuerdo con el pensamiento idealista que constituyó en otro tiempo su trasfondo filosófico, pues el paradigma de conocimiento para este tipo de filosofía era el conocimiento *a priori* el conocimiento que se obtiene de un modo independiente de la observación de los hechos.

No obstante, la teoría de la verdad como coherencia se suele valorar habitualmente como una teoría complementada, y no sustitutiva, de la teoría de la verdad como correspondencia.

6.3. Teorías de la verdad como correspondencia

Históricamente la teoría de la verdad como correspondencia está relacionada con las metafísicas realistas. De hecho, se puede encontrar una de sus primeras y más claras formulaciones en Aristóteles:

Decir de lo que es que no es o de lo que no es que es, es falso; mientras que decir de lo que es que es y de lo que no es que no es, es verdadero.

Contra esta tesis, tan aparentemente evidente como trivial, formuló Eubúlides la paradoja del mentiroso, tradicional dificultad que ha de resolver toda teoría que considere la verdad como una relación entre oraciones y hechos. No obstante, la teoría de la verdad como correspondencia siempre ha tenido partidarios en la historia de la filosofía, desde los estoicos hasta los empiristas. A principios de siglo el interés por ella se renovó, debido principalmente al auge de los movimientos filosóficos anti-idealistas (positivismo, fenomenología, neokantismo, etc.). Esta reacción anti-idealista tuvo cualificados exponentes en Gran Bretaña, como G. E. Moore y B. Russell, filósofos que desarrollaron teorías paralelas sobre la verdad.

Tanto según G. E. Moore como B. Russell, y algunos otros autores posteriores, la verdad es primariamente una propiedad de «proposiciones», de lo significado por las oraciones, es decir, de lo que se corresponde de modo aproximado con lo que hemos denominado 'ideas'. El significado es la entidad conceptual de carácter complejo que entendemos -aprehendemos, comprendemos- cuando alguien profiere una oración. Por lo tanto, en la situación comunicativa se dan dos factores plenamente distinguibles: el *acto* de descifrar o entender un mensaje y *lo que* se descifra o entiende en ese mensaje, el *objeto* de nuestro acto de descodificación lingüística.

Estos objetos son los que tenemos presentes en el pensamiento, antes incluso de encontrarles una adecuada expresión verbal. Pensamos mediante proposiciones o ideas y son éstos los objetos a los que atribuimos las propiedades de verdad o falsedad. Igualmente, las ideas pueden ser consideradas como los objetos de nuestras creencias, en vez de los enunciados que las expresan. Decir que alguien cree algo que es verdad, o que mantiene una creencia verdadera no es ni más ni menos que decir que cree una idea verdadera, no que cree en una oración verdadera. Pensar o creer que una idea es verdadera es pensar o creer que esa idea es un hecho, mientras que pensar o creer que una idea es falsa es pensar que algo no es un hecho. Así pues, entre las ideas y los hechos se dan para estos filósofos las relaciones de verdad y falsedad.

Moore, Russell y Wittgenstein expusieron diferentes concepciones sobre estas relaciones entre las ideas y los hechos. Incluso alguno de ellos cambió de concepción a lo largo de su desarrollo filosófico (como Wittgenstein). En un principio, la teoría de la verdad como correspondencia estuvo ligada al atomismo lógico que sostuvieron Russell y Wittgenstein durante una época. Tanto uno como otro, pero especialmente éste último, concebían la correspondencia entre el pensamiento (la idea) y la realidad como un *isomorfismo estructural*, esto es, como una relación que asignaba a cada componente de la realidad un componente en el pensamiento y a la estructura relacional del hecho la de la idea.

En el *Tractatus*, Wittgenstein postuló que la relación abarcaba también al lenguaje: tal como lo expresaba Wittgenstein, la oración era una *proyección* de la idea, proyección a través de la cual el lenguaje se relacionaba con la realidad. Es decir, no sólo a cada elemento del hecho le correspondía un elemento de la idea (verdadera), sino que también le correspondía uno en la oración (verdadera). La oración era concebida como un agregado de elementos simples (los nombres) cuya ordenación debía reflejar (al menos en un lenguaje ideal) la estructura del pensamiento y, por lo tanto, la de la realidad.

El concepto de realidad que Wittgenstein manejaba en el *Tractatus* era el de un *espacio lógico* constituido por el conjunto de los hechos posibles, reales y no reales ; luego lo que Wittgenstein denominaba 'realidad' equivaldría a 'realidad posible'. El conjunto de los hechos reales es lo que Wittgenstein llamaba *mundo*; el mundo es descomponible en última instancia en los hechos (realmente existentes); los hechos forman la substancia del mundo,

son sus átomos. La oración o enunciado con sentido es el aspecto perceptible de una figura lógica, representación o modelo (idea) de un hecho, el cual forma o no parte del mundo: para saber si es verdadera hay que compararla con el mundo, esto es, ver si hay un hecho real que le corresponde. Esta comparación consiste fundamentalmente en un examen de estructuras lógicas: por una parte, la estructura lógica de la oración (a veces oculta por su forma gramatical) y, por otra, la estructura lógica de los hechos reales. Si esta comparación entre el elemento lingüístico y el real tiene como resultado el establecimiento de una identidad, entonces se puede aplicar a aquél el predicado ‘verdadero’; si no, el ‘falso’. No hay ninguna oración de la que se pueda decir que es verdadera o falsa si no se efectúa esta comparación, lo cual significa que no hay verdades *a priori* (las tautologías y las contradicciones no son consideradas auténticas oraciones por Wittgenstein) o, lo que es lo mismo, que no hay hechos necesarios en el mundo.

Uno de los problemas que plantea el tipo de correspondencia que el *Tractatus* establecía entre el lenguaje y la realidad, el isomorfismo estructural, es que requería (al menos) dos supuestos metafísicos: (1) la descomponibilidad de la realidad (y, *a posteriori*, del mundo) en elementos últimos, esto es, el atomismo, y (2) la posibilidad de construir un lenguaje lógicamente perfecto (o de convertir, mediante el análisis, el lenguaje natural en un lenguaje semejante), un lenguaje en el que se reflejara de un modo claro la estructura del pensamiento y de la realidad. Una vez que Bertrand Russell y especialmente Ludwig Wittgenstein abandonaron estos supuestos, la identidad estructural entre lenguaje y realidad se vino abajo: era necesario concebir la correspondencia entre oración y hecho de un modo distinto. Sobre esta nueva concepción del lenguaje hablaremos en el capítulo 9.

6.4. La teoría semántica de la verdad: Alfred Tarski

La teoría de A. Tarski constituye la expresión más acabada y correcta de la teoría de la verdad como correspondencia. Al estar formulada en términos semánticos y dirigida a la semántica de los lenguajes formalizados, evita las posibles confusiones que se derivan de intrincados análisis filosóficos sobre los hechos, análisis en los que hasta cierto punto se perdieron G. E. Moore, B. Russell y el propio Wittgenstein.

Como prólogo a su definición de verdad para un lenguaje formal, A. Tarski formuló las condiciones que debe cumplir cualquier teoría de la verdad, que pretenda definir ese predicado semántico, para que sea considerada una teoría *adecuada*, con probabilidades de éxito. Esas condiciones son dos: 1) *adecuación material* y 2) *corrección formal*. La primera es una condición sobre el contenido de las posibles teorías, la segunda, sobre la forma que dichas teorías deben tener.

En cuanto a la adecuación material, el requisito fundamental que *cualquier* presunta teoría de la verdad debe satisfacer es que de ella se puedan seguir enunciados o teoremas que tengan la forma de (1)

(1) O es verdadera si y sólo si p

es decir, de la definición de verdad que establezca la teoría se ha de poder deducir para cualquier oración O, perteneciente al lenguaje L *para* el que se pretende definir el predicado ‘es verdadera’, una oración que se ajuste al esquema de (1), o como se le suele denominar esquema T (de *truth*, la palabra inglesa que significa *verdad*).

En este esquema, O se ha de sustituir, para cada caso concreto, por el nombre de una oración del lenguaje para el que se define el predicado semántico de verdad, generalmente por

la oración en cuestión entre comillas (aunque hay otros medios de construir nombres para las oraciones). De tal modo que, si queremos definir ‘es verdad’ para la lengua inglesa, ocuparemos el lugar de O con los nombres de las oraciones pertenecientes al inglés; por ejemplo:

(2) ‘snow is white’ es verdadera si y sólo si p .

El lugar que en el esquema T está ocupado por p ha de reemplazarse por una oración del lenguaje en el que se está definiendo el predicado ‘es verdad’. Esta oración ha de representar (por lo menos parcialmente) las condiciones de verdad de la oración que ocupa el lugar de O, esto es, lo que se tiene que dar o tiene que suceder para que la oración que está en el lugar de O sea verdadera según la teoría semántica en cuestión.

Si la teoría es materialmente adecuada, y dado que la lengua en la que estamos escribiendo sobre el predicado semántico es el español, hemos de sustituir en el esquema T la oración p por una oración que sea una traducción adecuada de la oración cuyo nombre ocupa el lugar de O en ese mismo esquema:

(3) ‘snow is white’ es verdadera si y sólo si la nieve es blanca.

Otra posibilidad, que se da muy frecuentemente, es que una teoría de la verdad para un lenguaje o lengua se formule en ese mismo lenguaje o lengua, por ejemplo, que se proponga en castellano una teoría de la verdad *para* el castellano. Como el lector advertirá rápidamente, si ha comprendido 2.1, esto significa utilizar una lengua como metalenguaje de (parte de) sí misma, esto es, que en el lugar que ocupa O estarán los nombres de las oraciones del castellano que ocupan el lugar de p

(4) ‘la nieve es blanca’ es verdadera si y sólo si la nieve es blanca.

El esquema T, sin embargo, no es una definición, ni siquiera parcial, de la noción de verdad (sería trivial en este caso); la definición de verdad para un lenguaje ha de rellenar ese esquema con las oraciones de ese lenguaje y sus correspondientes condiciones de verdad. El esquema T sirve únicamente para descartar ciertas teorías de la verdad como no adecuadas desde un punto de vista material. En efecto, supóngase que se dispone de dos teorías alternativas sobre la verdad para la lengua castellana. Una de esas teorías, T_1 , permite deducir para cada oración perteneciente al castellano, un teorema o enunciado que se ajusta al esquema T, mientras que la otra, T_2 , no (hay algunas oraciones del castellano que quedan fuera de la teoría₂, por ejemplo la oración que ocupa el lugar de O en (4)). El criterio de adecuación material nos permite prescindir de T_2 , por incompleta, y admitir en cambio a T_1 como posible teoría de la verdad para el castellano.

La corrección formal es el otro requisito que debe cumplir una candidata a teoría de la verdad. El requisito atañe al lenguaje en que dicha teoría se formula y a los predicados o conceptos que se utilizan en ella. La observación de este requisito permite evitar paradojas como la del mentiroso, al tener en cuenta la distinción entre lenguaje-objeto y metalenguaje y los predicados que pertenecen a cada uno de estos dos niveles. El predicado ‘es verdad’, en particular, es un predicado metalingüístico, un predicado que pertenece al metalenguaje (de parte) del castellano. En general, cualquier teoría sobre la verdad para un lenguaje L, se formulará en un metalenguaje de L, M_L . Una consecuencia de esta distinción de niveles y

de la condición de adecuación material es que el M_L ha de contener, por una parte, nombres para cada uno de los elementos de L y, por otra, oraciones de M que sean una traducción adecuada de las de L . Si L y M_L son, como en el ejemplo que hemos venido manejando, partes del castellano, M_L ha de contener como subconjunto a L .

Aparte de esta neta distinción entre lenguaje objeto y metalenguaje, otro aspecto del requisito de corrección formal es que el lenguaje para el que se define el concepto ‘es verdad’ sea *especificable* o determinable. Esto significa que se ha de disponer de un procedimiento para, dada una particular oración, O , determinar si esa oración pertenece o no a dicho lenguaje-objeto. Este aspecto del requisito es ineludible, pues sin él ni siquiera se puede averiguar cuándo una teoría de la verdad es materialmente adecuada.

La teoría de la verdad diseñada por Tarski fue concebida para los lenguajes de la matemática, esto es, para lenguajes en que es relativamente sencillo cumplir este último requisito formal (especificando un conjunto de reglas de formación, por ejemplo). La determinación de los límites del lenguaje para el que se define el predicado ‘es verdad’ es pues el primer paso de una definición de ese predicado semántico.

Como resulta obvio, el segundo ha de ser la delimitación del metalenguaje a que pertenece el predicado ‘es verdad’. En el caso de la teoría construida por A. Tarski, el metalenguaje era a su vez un lenguaje semiformalizado en el que se definía ‘es verdad’ para el cálculo de clases, que era el lenguaje objeto. En general, la especificación del metalenguaje consiste en una enumeración de sus elementos. Esos elementos deben incluir, dadas las condiciones formales y materiales de adecuación que debe cumplir toda teoría de la verdad, 1) al lenguaje objeto, L_0 , si la definición del predicado semántico se hace en un M_L que es (parte de) el mismo lenguaje al que pertenece L_0 ; si L_0 y M_L son lenguajes diferentes, M_L debe contener traducciones adecuadas de todos los elementos de L_0 ; 2) además, el metalenguaje debe contener variables, predicados y conectores metalingüísticos, esto es, debe disponer de nombres para (cualesquiera) tipos de elementos de L_0 , de predicados que se aplican a esos elementos, como ‘es un nombre de L_0 ’, ‘es verdadero’, etc. y de conectores como ‘y’, ‘si y sólo si’, etc.

El objetivo de Tarski, y gran parte de la razón del éxito que ha obtenido su teoría, era definir el predicado semántico de verdad en términos no semánticos, esto es, no utilizar en esa definición ningún término semántico primitivo, sino tan solamente los recursos sintácticos habituales de los lenguajes formales. Pero para obtener este resultado tuvo que dar un pequeño rodeo a través de la noción semántica de *satisfacción*.

En el L_0 para el que Tarski definía el predicado de verdad, y en general en todos los lenguajes formales de orden superior a cero, que admiten expresiones cuantificacionales, se pueden formar *fórmulas abiertas*, esto es, fórmulas en que hay variables libres que no están ligadas por ninguna expresión cuantificacional. A estas fórmulas no se les pueden aplicar directamente los predicados de verdad o falsedad, sino que únicamente son satisfechas por determinadas secuencias. Como se recordará (sección 4.3), ejemplos de fórmulas abiertas en el castellano son las expresiones ‘ x es el rey de España’ o ‘ x está entre y y z ’, que son satisfacibles por determinadas secuencias de individuos. En la mayoría de las ocasiones, se pueden formar fórmulas cerradas o enunciados ligando las expresiones libres con una expresión cuantificacional, como por ejemplo “hay un x , un y y un z tal que x está entre y y z ’. La verdad de este enunciado depende entonces de si hay una secuencia que satisface la fórmula abierta que se cuantifica o, dicho de otro modo, la verdad de este enunciado se puede *definir* en términos de la *satisfacción* de la correspondiente fórmula abierta.

Esta es la razón de que el tercer paso en la construcción de una teoría de la verdad suele ser la definición de la noción semántica de *satisfacción*. Esto es lo que hizo Tarski en su caso particular, especificando tal definición en términos no semánticos.

Disponiendo de la definición de la noción de satisfacción el cuarto paso es finalmente la propia definición de ‘es verdadero’.

A continuación expondremos un ejemplo muy simple de construcción de una teoría de la verdad al modo tarskiano. Desde el punto de vista formal, el ejemplo es trivial, y solamente se propone con finalidades pedagógicas. Además, tiene la peculiaridad de que como el lenguaje que utilizaremos para el ejemplo -al que llamaremos L_s - no tiene variables, no será necesario el rodeo por la definición de satisfacción para obtener la de verdad.

Sintaxis de L_s . Presentamos primero el «vocabulario» de este lenguaje artificial, escribiendo entre paréntesis la equivalencia aproximada de cada categoría si se tratara de una gramática de una lengua natural.

- *Constantes* («nombres propios») : c_1, c_2, \dots, c_n ; con ‘ c_k ’ aludiremos a una constante cualquiera de éstas.

- *Símbolos predicativos monádicos* («términos generales») : P_1, P_2, \dots, P_m ; con ‘ P_j ’ aludiremos a un símbolo cualquiera de éstos.

- *Conectar monádico* («adverbio de negación»).

- *Conector diádico* («conjunción copulativa»).

Estos son los símbolos terminales. Los no terminales o auxiliares son: O (para ‘oración’), Pred (para ‘símbolo predicativo monádico’) y C (para ‘constante’).

Podemos ahora definir el conjunto de las oraciones de L_s mediante las siguientes reglas de una gramática generativa no contextual:

$$\begin{aligned} O &\rightarrow \neg O \\ O &\rightarrow O \wedge O \\ O &\rightarrow \text{Pred } C \\ \text{Pred} &\rightarrow \{p_1, p_2, \dots, p_m\} \\ C &\rightarrow \{c_1, c_2, \dots, c_n\} \end{aligned}$$

En lógica, donde se utilizan lenguajes artificiales como el que estamos describiendo, se suele definir el conjunto de oraciones de éstos de un modo distinto: mediante una «gramática» recursiva o -en la terminología usual- definición recursiva. Si utilizamos α y β como símbolos del metalenguaje podríamos dar esta definición del siguiente modo:

(i) Para cualquier constante ‘ c_k ’ y símbolo predicativo monádico ‘ P_j ’: ‘ $P_j c_k$ ’ es una oración.

(ii) Si α es una oración, $\neg \alpha$ es una oración.

(iii) Si α, β son oraciones, entonces $\alpha \wedge \beta$ es una oración.

Hasta aquí la sintaxis de L_s formulada de dos modos distintos pero equivalentes en el sentido de que nos proporcionan el mismo conjunto de oraciones. Vayamos con la definición de ‘es verdadera’ (es decir la definición de ‘oración verdadera’). Hacemos la definición relativa al concepto de *extensión*, *denotación* o *referencia*.

(i’) ‘ $P_j c_k$ ’ es verdadera si y sólo si

(a) Existe un cierto e tal que ‘ c_k ’ denota e (o e es la extensión de ‘ c_k ’).

(b) Existe un cierto R_j tal que ‘ P_j ’ denota R_j (R_j es la extensión de ‘ P_j ’).

- (c) e es (un) R_j .
- (ii') $\neg\alpha$ es verdadera si y sólo si α no es verdadera.
- (iii') $\alpha \wedge \beta$ es verdadera si y sólo si α es verdadera y β es verdadera.

Esta definición de verdad reúne las dos características de adecuación material y corrección formal, puesto que para cualquier oración de L_s la definición puede generar, a partir de las reglas (i')-(iii') el correspondiente enunciado T. Por ejemplo, considérese la expresión, perteneciente a L_s : $(P_1c_2 \wedge P_2c_1) \wedge \neg P_2c_1$. La derivación de ese enunciado sería como sigue (prescindimos de hacer explícito el papel que desempeñan los paréntesis, fácilmente comprensible en el ejemplo):

- (5) $(P_1c_2 \wedge P_2c_1) \wedge \neg P_2c_1$ es verdadera si y sólo si:
- 1) $P_1c_2 \wedge P_2c_1$ es verdadera y $\neg P_2c_1$ es verdadera.
 - 2) P_1c_2 es verdadera y P_2c_1 es verdadera.
 - 3) P_1c_2 es verdadera y P_2c_1 es verdadera y P_2c_1 no es verdadera
 - 4) Existe R_1, R_2 y e_1, e_2 tales que: ' P_1 ' denota R_1 ; ' P_2 ' denota R_2 ; e_1 es (un) R_2 y e_1 no es un R_2 .

Tal vez entenderemos esto mejor si por un momento nos imaginamos a L_s no como un lenguaje artificial, sino como un esquema de una lengua natural, o de una parte suya. Imaginemos, por ejemplo, que dicha lengua es la hablada por una «nación» de indios norteamericanos. Prosiguiendo el ejemplo puramente imaginario, pensemos que ' c_2 ' es el símbolo esquemático del nombre de un miembro de esa nación al que en nuestra lengua llamamos Nube negra y ' c_1 ' el de Flecha veloz. ' P_1 ' y ' P_2 ' denotan lo que en nuestra lengua denotamos mediante 'hechicero' y 'jefe de tribu', respectivamente. Entonces el enunciado T para el ejemplo anterior $(P_1c_2 \wedge P_2c_1) \wedge \neg P_2c_1$ (que consideramos ahora más bien como el esquema de una oración de la lengua de esa nación india) sería:

' $(P_1c_2 \wedge P_2c_1) \wedge \neg P_2c_1$ ' es verdadera si y sólo si Nube negra es un hechicero y Flecha veloz es un jefe de tribu y Flecha veloz no es un jefe de tribu.

Naturalmente, nuestra teoría de la verdad para la lengua en cuestión nos dice que esa oración es una contradicción. Hasta aquí el análisis del ejemplo.

Originalmente, la teoría tarskiana de la verdad tenía un carácter absoluto, esto es, definía el predicado 'es verdad' (o 'es verdadero(a) en términos de satisfacción por el mundo o la realidad. Posteriormente, la teoría sufrió modificaciones o reformulaciones que le hicieron perder ese carácter. En particular, en las ciencias formales (y en algunas naturales) se usa una concepción de la verdad relativa, que hace depender esta noción de un modelo con el que se pone en relación el lenguaje formal para el que se define el predicado. Así, la noción de verdad queda convertida en verdad en el modelo M_i .

Esta teoría modelista de la verdad elimina una de las características más sobresalientes de la teoría original de Tarski, su independencia de nociones semánticas primitivas. Al ser relativa la verdad de la oración de un lenguaje al modelo que se utilice para interpretarlo, es necesario introducir como noción primitiva en la definición la de interpretación en M_i , de tal modo que el esquema T tiene la forma:

(6) O es verdadera en M_i , si y sólo si p .

Sin embargo, la concepción modelista de la verdad tiene otros méritos que, para ciertos fines, la hacen preferible a la primitiva teoría de Tarski. Uno de ellos es que permite tratar a los lenguajes, y en particular a las teorías científicas, como estructuras formales que pueden tener diversas realizaciones o modelos matemáticos o empíricos.

Ilustraremos esta concepción modelista de la verdad mediante un ejemplo. Consideremos de nuevo el lenguaje L_s tal como lo habíamos pensado al principio, como un lenguaje artificial y no como una versión esquemática de (parte de) una lengua natural. Consideremos los dos modelos siguientes (definidos parcialmente pues sólo damos la interpretación de dos constantes y dos predicados para simplificar):

M_1 . Dominio de objetos: los números naturales.

Interpretación de c_1 : 2.

Interpretación de c_2 : 3.

Interpretación de P_1 : *número par*.

Interpretación de P_2 : *número impar*.

M_2 . Dominio de objetos: los números enteros.

Interpretación de c_1 : 4.

Interpretación de c_2 : 3.

Interpretación de P_1 : *número primo*.

Interpretación de P_2 : *número impar*.

Sería fácil modificar la definición absoluta de verdad para obtener las versiones relativas, en este caso, *verdad en M_1* y *verdad en M_2* . Sin necesidad de hacerlo se aclararán estas nociones pensando en algunas oraciones concretas de L_s y su respectivo valor veritativo relativamente a M_1 o M_2 . Consideremos, por ejemplo, las oraciones :

$P_1c_1 \wedge P_2c_2$; $P_2c_1 \wedge P_1c_2$; $\neg P_1c_1 \wedge P_2c_2$. Es bastante obvio que la primera, $P_1c_1 \wedge P_2c_2$ es verdadera en M_1 , pero no es verdadera en M_2 . En efecto, los correspondientes enunciados T modificados serían, respectivamente:

$P_1c_1 \wedge P_2c_2$ es verdadera en M_1 si y sólo si 2 es (un) *número par* y 3 es (un) *número impar*.

$P_1c_1 \wedge P_2c_2$ es verdadera en M_2 si y sólo si 4 es (un) *número primo* y 3 es (un) *número impar*.

El lector podrá comprobar, similarmente, que ' $P_2c_1 \wedge P_1c_2$ ' no es verdadera ni en M_1 ni en M_2 y que $\neg P_1c_1 \wedge P_2c_2$ no es verdadera en M_1 pero sí lo es en M_2 .

A. Tarski concebía su teoría semántica de la verdad como filosóficamente neutral: «... podemos aceptar la concepción semántica de la verdad sin abandonar ninguna actitud gnoseológica que podamos haber tenido; podemos seguir siendo realistas ingenuos o idealistas, empiristas o metafísicos, lo que hayamos sido antes. La concepción semántica es completamente neutral respecto a todas esas posiciones» (*La concepción semántica de la verdad...*, pág. 55). No obstante, H. Field (1972) ha subrayado que las exigencias metodológicas de Tarski estaban en la línea del fisicalismo. En particular, el interés tarskiano en la no utilización de nociones semánticas primitivas respondía a la necesidad fisicalista de definir nocio-

nes abstractas (como *verdad* y *satisfacción*) en términos de relaciones entre objetos físicos tanto como a la deseabilidad de la utilización de nociones sintácticas precisas. Este mismo autor mantiene la tesis de que, en realidad, Tarski no consiguió una reducción auténtica de las nociones semánticas, ya que en su definición aparecen de un modo implícito otras nociones semánticas primitivas. Sea esto como sea, lo que es cierto es que Tarski acertó a dotar al término ‘correspondencia’, clave en la teoría de la verdad, de un significado mucho más preciso y filosóficamente aceptable de lo que se había logrado hasta entonces. Con ello, contribuyó de forma decisiva a la aclaración de uno de los grandes problemas de la filosofía del lenguaje: ¿cuál es la naturaleza de la relación del lenguaje con la realidad?

6.5. La teoría pragmática de la verdad

En 2.6, se caracterizó a la pragmática como la disciplina que estudia los signos teniendo en cuenta los aspectos contextuales de su empleo. Una teoría pragmática de la verdad se distinguirá por tanto en que hace intervenir algunos de esos aspectos contextuales en su definición del concepto. La teoría semántica de Tarski no tiene en cuenta los elementos del contexto en que se pueden utilizar las oraciones: trata de la relación entre éstas (en definitiva, entidades abstractas ; cfr. 2.5) y la realidad de la que se supone que hablan.

Ahora bien, no es difícil advertir que, en un sentido estricto, la mayoría de las oraciones no son ni verdaderas ni falsas, sino que sólo se les pueden adscribir esas propiedades cuando son proferidas por hablantes concretos en circunstancias determinadas. Así, la oración ‘hoy está lloviendo’ no es ni verdadera ni falsa sino cuando es emitida por un hablante en un cierto momento de tiempo, t_k , y en un punto del espacio, s_j ; lo que son verdaderas o falsas son las diferentes *preferencias* que se pueden hacer de esa misma oración.

No obstante, la anterior observación no supone un gran problema para la teoría semántica de la verdad. Aunque se supone que ésta no tiene que hablar más que de determinado tipo de relaciones entre oraciones y elementos de la realidad, no es difícil ampliarla para dar cabida a elementos contextuales como el tiempo y el lugar de la preferencia, el sujeto que la hace, etc. De este modo se conviene la teoría en una teoría sobre la verdad de preferencias de oraciones, en vez de las oraciones mismas, pero esto no supone en principio ningún cambio fundamental: la verdad se sigue concibiendo como un tipo especial de correspondencia entre entidades lingüísticas y la realidad de la que hablan.

Cuando en la moderna filosofía del lenguaje se habla de teoría pragmática de la verdad, se quiere decir generalmente otra cosa; se quiere designar con esa expresión una teoría o grupo de teorías que mantienen, con diferentes matices, que ‘es verdad’ no es una propiedad que tenga que ver con la correspondencia entre el lenguaje y la realidad, e incluso que no se trata en absoluto de una propiedad de entidades lingüísticas, sean éstas oraciones o preferencias.

El origen de estas teorías pragmáticas de la verdad, entre las que destaca la del filósofo británico P. F. Strawson, está en la obra de F. P. Ramsey. Según este filósofo, el predicado ‘es verdad’ se puede eliminar de todos los contextos en que aparece sin que ello tenga como resultado una pérdida de contenido semántico o informativo. La tesis de Ramsey, conocida también como tesis o teoría de la *redundancia*, es que el predicado veritativo no añade nada al significado de una oración cuando se le aplica, de tal modo que ‘ α es verdadera’ significa lo mismo que ‘ α ’ y ‘ α es falsa’ significa lo mismo que ‘no α ’.

La teoría que defendió P. F. Strawson es una modificación de la teoría de la redundancia de F. P. Ramsey. A la teoría de P. F. Strawson se la suele calificar de teoría *realizativa* de la

verdad. Dicho brevemente, porque este extremo se aclara en el capítulo 9, una expresión realizativa es aquella que se emplea para *hacer* algo, y su simple emisión constituye, con la satisfacción de determinados requisitos, la realización del acto en cuestión. La pronunciación de las palabras ‘yo juro’ no sólo es una preferencia de una expresión lingüística con significado, sino que además constituye por sí misma un acto, el acto de jurar. Lo mismo sucede con expresiones del tipo ‘os declaro marido o mujer’, pronunciada por el sacerdote o el juez, ‘queda inaugurado este congreso’, en boca del político.

El punto básico de la teoría de Strawson es la afirmación de que la expresión ‘es verdad’ es una expresión *realizativa*. Según Strawson, cuando alguien afirma ‘ α es verdad’, donde α sustituye a un enunciado, está haciendo algo que se puede describir como el acto de *respal-dar, aceptar o estar de acuerdo con α* . Cuando alguien afirma ‘es verdad que está nevando’ no añade ningún contenido informativo nuevo, ninguna propiedad especial al enunciado ‘está nevando’. Lo que está sobreañadido en el primer enunciado es, en todo caso, la expresión de la actitud del hablante hacia ese enunciado.

El propio Ramsey había puesto de relieve este aspecto de la utilización de la expresión ‘es verdad’, que él identificaba con un uso típicamente adverbial, como el de las expresiones ‘realmente’, ‘ciertamente’, etc. El análisis de Strawson constituye pues un complemento a la teoría de Ramsey: afirma, como la teoría de la redundancia, que ‘es verdad’ no designa una propiedad lingüística, aunque es una propiedad del *uso* que se hace de expresiones lingüísticas.

Mientras que los partidarios de la aplicación al lenguaje natural de la teoría semántica afirman que el predicado ‘es verdad’ se utiliza en el metalenguaje para hacer afirmaciones sobre enunciados de un lenguaje, Strawson acude a argumentos que hacen apelación al uso del lenguaje natural, a las prácticas que corrientemente se llevan a cabo con él. Entre éstas se encuentra la de traducción; según Strawson ningún traductor viene el enunciado inglés ‘it’s true that’s snowing’ por el castellano «‘that’s snowing’ es verdad», por lo que el enunciado inglés no se utiliza como un enunciado acerca de otro. Aunque ‘es verdad’ se trata, gramaticalmente hablando, de un predicado, su uso no se corresponde con el de los predicados corrientes. En este aspecto es comparable al uso del adverbio ‘sí’. Cuando alguien afirma ‘sí, está nevando’ está hasta cierto punto *reforzando* su afirmación de que está nevando, pero no está añadiendo nada nuevo al enunciado mediante el cual comunica la citada información: simplemente está manifestando una actitud hacia tal enunciado.

Por otro lado, Strawson insiste en un uso ligeramente diferente de la expresión ‘es verdad’. Se trata de lo que denomina *uso expresivo*, en el que la expresión es equiparable a ‘es cierto’, ‘de verdad que...’ o incluso a su forma interrogativa ‘¿es verdad que...?’. En estas ocasiones ‘es verdad’ se emplea como *medio* para que el hablante exprese actitudes, emocionales o no, pero específicas, tales como la duda, el escepticismo, el asombro, etc.

Las críticas a la teoría de Strawson comparten la característica de resaltar el hecho de que hay ciertos usos de ‘es verdad’ que no son reducibles ni al propiamente *realizativo* ni al *expresivo*. Así, P. T. Geach ha criticado los puntos de vista de Strawson con el siguiente argumento:

(7) Si α es verdad, entonces β
 α es verdad

luego β

Este argumento constituye una simple y válida aplicación de la regla lógica denominada *modus ponens*. Ahora bien, razonó Geach, si ‘ α es verdad’ se entiende en el sentido de Strawson, realizativo, no sucede lo mismo con el antecedente ‘si α es verdad’. No obstante, para que valga el razonamiento, los dos enunciados han de entenderse del mismo modo, luego ‘ α es verdad’ no es un enunciado realizativo.

La objeción de Geach se puede resolver fácilmente diciendo que, si ‘ α es verdad’ no es realizativo, es justamente porque significa lo mismo que ‘ α ’ (como Ramsey afirmó), de tal modo que el razonamiento anterior equivale ni más ni menos que a la regla deductiva escrita en su forma habitual:

$$(8) \begin{array}{l} \alpha \rightarrow \beta \\ \alpha \\ \hline \beta \end{array}$$

Por otro lado, también se ha aducido que hay usos de ‘es verdad’ que son *ciegos*, esto es, en que la expresión se aplica a enunciados no existentes, como cuando se afirma

(9) Todo lo que dice el secretario general de mi partido es verdad.

En estos casos no se puede apelar, por un lado, al uso descriptivo de la expresión en cuestión, pues no hay enunciados de los que se pueda afirmar que poseen propiedades y, por otra parte, no se puede invocar tampoco al uso realizativo.

El medio de tratar este problema ya fue avanzado por Ramsey: la cuantificación sobre enunciados. Así, la oración (9) queda convertida en

(10) $\forall \alpha$ (α es dicho por el enunciado general de mi partido $\rightarrow \alpha$).

No obstante, esta solución plantea a su vez los problemas de la cuantificación sobre enunciados, problemas sobre si los enunciados se pueden considerar como términos, si se puede cuantificar sobre ellos, etc.

6.6. La aplicación de la teoría semántica de la verdad al lenguaje natural. El programa de D. Davidson.

A. Tarski, que había diseñado la teoría de la verdad pensando en los lenguajes formalizados de la matemática, dudaba de que aquella se pudiera aplicar a las lenguas naturales. Fundamentaba su escepticismo en dos razones: 1) las lenguas naturales pueden funcionar como metalenguajes de sí mismas, y 2) las lenguas naturales son lenguajes que (presuntamente) no se pueden definir o especificar de un modo formal. La primera dificultad reside en el carácter *cerrado* de las lenguas naturales, puesto que éstas contienen no sólo predicados lingüísticos, sino también los predicados metalingüísticos, las expresiones y los medios para formar nombres de expresiones, etc. Esto tiene como consecuencia que una teoría de la verdad para una lengua natural, por ejemplo el español, sólo puede tener un carácter *parcial*, esto es, no puede abarcar a todo el español ya que una parte, la que contiene entre otras cosas el predicado ‘es verdad’, ha de quedar necesariamente fuera para que la teoría cumpla el requisito de corrección formal. Por lo tanto, en español se puede formar una teo-

ría de la verdad sólo de un *fragmento* del español, el que oficia de lenguaje objeto. Esta dificultad pierde parte de su importancia cuando se utiliza una lengua natural como meta-lenguaje de otra lengua natural diferente.

La segunda dificultad hace referencia a la imposibilidad de cumplir, en el caso de las lenguas naturales, el primer paso de la construcción de una teoría de la verdad para un lenguaje, el de especificar de modo adecuado ese lenguaje. Esta imposibilidad tenía más visos de verosimilitud en la época en que Tarski enunció su teoría que en la segunda mitad del siglo XX, en que la lingüística (especialmente la generativa) ha dado grandes pasos en las investigaciones gramaticales. En particular, cualquier paradigma de investigación lingüística considera hoy en día la sintaxis como una disciplina que provee justamente la especificación formal de un lenguaje. Del mismo modo que las reglas de formación de un lenguaje formal determinan el concepto ‘es una fórmula’ para ese lenguaje, las gramáticas de las lenguas naturales determinan el predicado ‘es una oración bien formada’ para cada una de esas lenguas. No obstante, esta determinación lo es de una *estructura sintáctica* y no de una *forma lógica*, por lo que en este punto se presenta un interesante problema a la aplicación de la teoría semántica de la verdad a las lenguas naturales, problema que los filósofos del lenguaje han tratado de resolver en los últimos años.

A pesar del pesimismo de A. Tarski, ha habido filósofos del lenguaje que han intentado extender la aplicación de su teoría a las lenguas naturales, convirtiendo tal extensión en el núcleo de sus teorías semánticas. Tal es el caso de D. Davidson, el principal representante de lo que se ha dado en llamar *semántica de las condiciones de verdad* (*truth conditional semantics*).

Según Davidson, la teoría semántica ha de dar cuenta de la *productividad* (competencia) semántica de los hablantes de una lengua natural, es decir, de la capacidad que éstos poseen para dotar de significado a un número (potencialmente) infinito de oraciones pertenecientes a su lengua. Esto significa que los hablantes de esa lengua manejan un conjunto de reglas finito y recursivo como base de esa capacidad; el objetivo de la semántica ha de ser por tanto describir ese conjunto de reglas. La teoría semántica en cuestión ha de poder producir, para cualquier oración, O_i , perteneciente a una lengua, un enunciado que tenga la forma:

(11) O_i significa que p

donde O_i es un nombre de la *i-ésima* oración de la lengua y p un equivalente metalingüístico de O_i . Si la teoría es adecuada y O_i y p pertenecen a la misma lengua, O_i ha de ser un nombre de p . Esto es lo que sucedía también en la teoría de la verdad. Pues bien, ¿por qué no pensar que la teoría de Tarski proporciona ese conjunto de reglas recursivas? Basta con sustituir ‘significa’ por ‘es verdad’ en el esquema (11), ya que la definición de ambas nociones ha de estar sujeta a las mismas condiciones. Esa es precisamente la conclusión de Davidson: dar una teoría del significado para una lengua natural es proporcionar una teoría de la verdad para ese lenguaje.

La tesis de Davidson se apoya en una intuición sólida: conocer el significado de una oración equivale a saber todas las (posibles) ocasiones en que dicha oración es verdadera. Y esas ocasiones están (supuestamente) determinadas por la oración p . Si fuera cierta, la tesis de Davidson tendría a su favor el hecho de que permite desembarazarse en teoría semántica de una noción bastante oscura, como la de significado, en favor de otra mucho más precisa y manejable, la de verdad. En efecto, cuando se afirma que dos enunciados son sinónimos, se piensa en una identidad de entidades denominadas *significados*. Pero, con la teoría de

Davidson, afirmar que dos enunciados son sinónimos no es ni más ni menos que afirmar que son verdaderos en las mismas ocasiones o bajo las mismas condiciones.

Así, las oraciones

(11) La neige est blanche

(13) Snow is white

son sinónimas porque tienen una misma condición de verdad, que la nieve es blanca; ambas *significan* que la nieve es blanca. Sin embargo, la tesis de Davidson ha de enfrentarse con algunas dificultades. La primera de ellas se presenta cuando la lengua en que se está formulando una teoría del significado contiene como parte el lenguaje objeto. Las condiciones de adecuación de Tarski establecen que la teoría de la verdad ha de generar enunciados como:

(14) 'la nieve es blanca' es verdadera si y sólo si la nieve es blanca.

Si se hace equivaler el predicado 'es verdadera' a 'significa', la teoría del significado genera entonces

(15) 'la nieve es blanca' significa que la nieve es blanca.

Ahora bien, como tanto el lenguaje objeto como el metalenguaje son (parte de) el español, la oración (15) carece de contenido empírico, no nos aclara nada acerca del significado de 'la nieve es blanca'. Esta dificultad se puede solucionar exigiendo que el lenguaje objeto para el que se define el predicado 'significa' no sea un subconjunto del metalenguaje en que se define tal predicado. Si se mira bien, esto es además lo que sucede cuando un hablante pregunta por el significado de una oración que pertenece a la lengua que él habla. Por ejemplo, supóngase que un hablante del castellano, *h*, pregunta por el significado de la oración

(16) Se puso las almadreñas

y se le responde con la siguiente equivalencia

(17) 'Se puso las almadreñas' significa que se puso los zuecos.

Si la equivalencia establece para *h* el significado de (16) es porque la oración a la derecha de la equivalencia pertenece al español que *h* habla y comprende, a lo que se denomina su *idiolecto* del español. El idiolecto de *h* opera entonces como metalenguaje y (16) como parte del lenguaje objeto: la equivalencia se ajusta entonces a las exigencias formales y sustantivas de una teoría de la verdad y del significado. Dadas estas observaciones, se comprenderá cuál es la dificultad de formular una teoría total del significado para una lengua natural cuando se está utilizando esa lengua como metalenguaje.

Supóngase en cambio que, para construir una teoría del significado en una lengua natural se emplea otra lengua natural, por ejemplo, que se emplea el castellano como metalenguaje del inglés. En este caso, O_i estará ocupado por un nombre de una oración del inglés y *p* por un equivalente de esa oración en el castellano, como es

(18) 'snow is white' significa que la nieve es blanca.

Ahora bien, ¿qué es lo que asegura que la oración que ocupa el lugar de p es una traducción correcta de la que está sustituyendo a O_i ? No se puede responder diciendo que es la identidad de significados, pues eso es justamente lo que postula la oración (18). La única respuesta posible es la equivalencia de condiciones de verdad: siempre que O_i es verdadera también lo es p .

En este punto se han señalado dos objeciones a Davidson. La primera es de tipo formal: puede ser que la equivalencia veritativa entre O_i y p no sea un requisito lo suficientemente fuerte como para asegurar su identidad de significado. Por ejemplo, considérese la oración

(19) 'l'eau bouille' es verdadera si y sólo si el agua está a 100°.

En esta oración la equivalencia veritativa puede ser cierta, es decir, generada por una teoría correcta de la verdad para el francés, y sin embargo no serlo.

(20) 'l'eau bouille' significa que el agua está a 100°.

Se han hecho algunas propuestas para solucionar esta dificultad, como por ejemplo exigir que la equivalencia veritativa sea *necesaria*. Pero debido a la poca claridad de la noción de *necesidad*, su introducción en el esquema T plantea más problemas de los que resuelve. La cuestión de fondo es sin duda la de si es adecuada y completa la reducción de un término intensional, como el de significado, a otro extensional, como el de verdad.

La segunda objeción a que aludíamos es de tipo epistemológico: ¿cómo se comprueba la verdad de las oraciones del tipo de (18)? La teoría del significado proporciona una definición recursiva de esta noción, pero esa definición parte de una *lista de oraciones simples* del lenguaje objeto y de sus equivalencias metalingüísticas. ¿Cómo se llega a esas equivalencias y qué garantiza su corrección? La respuesta de Davidson se halla una vez más en la *teoría de la verdad*. La oración (18) es verdadera si lo es la correspondiente equivalencia veritativa. Esta es una afirmación muy importante, porque muestra cómo se puede utilizar la teoría de la verdad para construir una *teoría de la interpretación*.

En efecto, supóngase que el investigador se encuentra con una lengua primitiva, de la que sólo conoce (parte de) su sintaxis. ¿Cómo puede resolver el problema de asignar un significado a las oraciones de ese lenguaje? ¿cómo puede establecer equivalencias entre las oraciones de la lengua que investiga y la propia? De acuerdo con la 'tesis de Davidson lo que tiene que hacer es construir una teoría de la verdad de esa lengua (utilizando la suya como metalenguaje, por supuesto).