

Juan José Lara Peñaranda

Indeterminación y Primera Persona

Director: Manuel Hernández Iglesias

Universidad de Murcia 2009

Departamento de Filosofía

ÍNDICE

Introducción	5
Parte 1. La tesis de la indeterminación	19
1.1. La base del sistema de Quine	21
1.2. Indeterminación e inescrutabilidad	28
1.3. Relatividad ontológica	51
1.4. La propuesta de Davidson	61
1.5. La indeterminación en Davidson	70
1.6. Indeterminación y lenguas naturales	77
1.7. Holismo y caridad	85
1.8. ¿In- o infradeterminación?	100
Parte 2. Indeterminación y primera persona	117
2.1. Indeterminación y reducción al absurdo	119
2.1.1. Relatividad ontológica y valor literal	126
2.1.2. Lenguaje de fondo y sistema coordinado	129
2.2. Primera persona y conductismo	133
2.3. Primera persona y explicitismo	142
2.4. Acceso privilegiado	154
2.5. Contenido intencional y mentalismo	163
2.5.1. Contenido intencional e internismo	165
2.5.2. Mentalismo y uso	178
2.6. La paradoja pragmática según Soames	189
2.6.1. Infradeterminación y referencia	189
2.6.2. Referencia tarskiana y referencia intuitiva	198
2.7. Nivel inmanente y nivel trascendente	206
Parte 3. Intuición lingüística y subjetivismo	223
3.1. Intuiciones lingüísticas	224
3.1.1. Intuiciones e infradeterminación	226
3.1.2. Nivel teórico y nivel básico	229
3.1.3. Nivel básico: infra- e indeterminación	233
3.2. La teoría «de la teoría» del lenguaje	238
3.3. Intuiciones y externismo	247
3.3.1. Triangulación y externismo	248
3.3.2. Interpretación e indeterminación	253
3.3.3. Interpretación sin convenciones	264

3.4. Antiintuicionismo lingüístico	269
3.4.1. Intuiciones, sinonimia y analiticidad	270
3.4.2. La labor de la semántica	274
3.4.3. Intuiciones y competencia	280
3.5. Davidson: interpretación y primera persona	287
3.5.1. Asimetría epistémica y simetría semántica	288
3.5.2. La autoridad como precondition	306
Conclusiones y Perspectivas	321
1. Conclusiones	321
2. Perspectivas	326
Referencias	331

INTRODUCCIÓN

Éste es un trabajo sobre la tesis de la indeterminación de la traducción, «el argumento filosófico más fascinante y el más discutido desde la Deducción Trascendental de las Categorías de Kant» según Putnam, «el más radical y original problema escéptico que ha visto la filosofía hasta la fecha» según Kripke.¹ La tesis fue formulada por primera vez por Quine, en la ya clásica obra de 1960 *Palabra y Objeto*, y ha sido defendida posteriormente por muchos y renombrados filósofos, especialmente por Davidson. Aparece allí, por primera vez en la literatura filosófica – aunque hallamos ya reflexiones en esta línea en Wittgenstein (v. L. Wittgenstein 1953, §206) –, el experimento mental de la *traducción radical*, que, más tarde, Davidson reconvertirá en la *interpretación radical*.² El objetivo del experimento es plantear la cuestión de cómo logramos comprender el lenguaje de otro hablante del que nada conocemos. Es decir, llegados a una remota tribu, de cuyos nativos todo nos es desconocido, ¿cómo lograr traducir su lengua? O, en la versión davidsoniana del experimento, ¿cómo entendemos – prescindamos de la traducción – a un hablante del que nada conocemos, y esto sin presuponer, tan siquiera, que dominamos ya una lengua? El resultado del experimento es que la traducción está indeterminada, en el sentido de que siempre será posible confeccionar manuales de traducción que ofrezcan traducciones diferentes de las oraciones del interpretado y que, sin embargo, resulten, todos, empíricamente correctos. Si una misma oración puede traducirse a otro lenguaje de diferentes – e incompatibles – formas, todas empíricamente correctas, pierde su sentido hablar de «la traducción

¹ La afirmación de Putnam en 1975, p. 159, la de Kripke en 1982, p. 60.

² En el párrafo apuntado, Wittgenstein recurre ya al mismo experimento mental que Quine: «Piensa que llegas como investigador a una tierra desconocida con una lengua que te es completamente extraña. ¿Bajo qué circunstancias dirías que la gente allí da órdenes, entiende órdenes, las siguen, se rebelan contra ellas, etc.?»

correcta» y, por ende, dado que la traducción empíricamente correcta preserva el significado, del «significado correcto» más allá de lo que se hace manifiesto en la conducta verbal de los hablantes. Lo mismo vale para los términos que forman parte de dichas oraciones, ya que es posible asignar diferentes referencias a un mismo término mientras que, de nuevo, las traducciones resultantes de las oraciones que los contienen resulten empíricamente correctas. Estas ideas permanecen en Davidson, sustituyendo los manuales de traducción quineanos por teorías interpretativas.

La tesis, de ser correcta, vendría a suponer la refutación definitiva de cualquier teoría semántica «realista»; es por esto que los críticos de la tesis ven en ella la disolución de dos nociones centrales de la semántica tradicional, como son la de «significado» y la de «referencia», y se ve la tesis como derivando, pues, en una especie de nihilismo semántico. Los defensores, por su parte, rechazan tal diagnóstico. Dado que el lenguaje constituye *esencialmente* una práctica pública, en el sentido de intersubjetiva, no parece que extraer consecuencias filosóficas de ello pueda derivar en nihilismo. Se trata, eso sí, del rechazo al significado o la referencia entendidos a la manera tradicional, es decir, con fuertes resabios platónicos y metafísicos. Davidson resultará especialmente expeditivo en esta cuestión: «No veo ninguna manera de evitar la indeterminación», dice en su artículo de 1997 precisamente llamado «Indeterminism and Antirealism», «pero creo que deja intacta la realidad de lo mental» (p. 73).

Pero la tesis ha sido también criticada por un número no menor de igualmente ilustres filósofos; estas críticas se han dirigido a diversos aspectos de la misma – p. ej., la

posibilidad de determinar la referencia de los términos a partir de su uso o lo errado de una de las nociones que sirven de pilares a la tesis en su formulación quineana, la de significado estimulativo – , pero este trabajo se centra en una crítica formulada por primera vez por J. Searle: la tesis es paradójica, constituye una reducción al absurdo de sí misma. Y percatarse de esta autorefutación supone, además, percatarse de que la asimetría entre la primera y la tercera persona resulta fundamental en cuestiones de semántica. Desde que J. Searle efectuara dicha crítica, en su ya clásico artículo de 1987 «Indeterminacy, Empiricism and the First Person», han sido muchos los filósofos que han visto en la tesis una amenaza inaceptable a la autoridad de la primera persona respecto al significado de sus palabras. Efectivamente, la tesis pretende refutar el bautizado por Quine como mito del museo por medio de la demostración de que el significado de cualquier oración puede ser caracterizado de diferentes maneras *por parte de un traductor o intérprete*, ahora bien, se nos dice ahora, esta perspectiva de tercera persona no hace justicia a la naturaleza del significado. El hablante, desde su perspectiva de primera persona, debe saber cuál de las diferentes traducciones/interpretaciones es la correcta. El problema fundamental de la tesis de la indeterminación o la referencia se hallaría en la metodología de tercera persona adoptada por Quine y Davidson, y su consiguiente exclusión de todo elemento subjetivo en la semántica – B. C. Smith llega a hablar de «la experiencia fenomenológica» al oír palabras cuyo significado conocemos de antemano, Searle apela a «significados intencionales» o al «contenido Intencional» que le damos a nuestras palabras – . Se reclama, desde esta crítica, un reconocimiento decidido de la asimetría entre la primera y la tercera persona en la reflexión acerca de la semántica.

Y es este rechazo de la asimetría el causante de que la tesis resulte según Searle, como dijimos, una reducción al absurdo de sí misma. Adoptemos el famoso ejemplo de Quine; llegados a una tribu que utiliza la palabra ‘gavagai’ en las mismas circunstancias en las que nosotros utilizamos la nuestra ‘conejo’, podemos elaborar un manual de traducción donde el término ‘gavagai’ se traduce por ‘conejo’, pero igualmente podríamos traducirlo por ‘parte de conejo no separada del resto del conejo’ o ‘instanciación de la conejeidad’, etc. (realizando también los pertinentes cambios en la traducción de otras partes del lenguaje para mantener invariables las condiciones de verdad de cada oración). Ahora bien, pregunta Searle, si nuestro término ‘conejo’ resulta empíricamente equivalente a ‘parte de conejo no separada del resto del conejo’ o a ‘instanciación de la conejeidad’, ¿por qué no entiendo, al leer a Quine, que lo que se me está diciendo es que ‘conejo’ tiene la misma referencia que ‘conejo’ o que ‘conejo’? No es esto lo que entiendo, concluirá Searle, porque cuando yo digo ‘conejo’ le doy a este término una referencia muy concreta – me refiero a conejos – y, si bien es cierto que otra persona podría interpretarme como refiriéndome con él a partes no separadas del conejo, yo sé que me refiero a conejos. Así, se trataría de una tesis que es formulable porque es falsa – comete una paradoja pragmática, en la expresión sugerida a Searle por R. Kirk –. La impresión de Searle es que la tesis parece una refutación de posturas que rechazan la asimetría entre la primera y la tercera persona en cuestiones de semántica mucho más que una defensa de dicho rechazo. Estaríamos ante una simpática ironía: Quine habría pergeñado su propia refutación sin percatarse de ello.

La primera respuesta contundente recibida por el artículo de Searle sería de manos de D. Føllesdal, en su artículo de 1990 «Indeterminacy and Mental States». La idea principal aquí es que la crítica de Searle se basa, simplemente, en una incomprensión de la tesis. Searle desarrolla su crítica tomando los términos por separado y afirmando entonces que ‘conejo’ se refiere a conejos y ‘parte no separada de conejo’ a partes no separadas de conejos – apelando, por cierto, a algo así como convenciones en la referencia de los términos en las lenguas naturales: en castellano, ‘conejo’ se refiere a conejos – . Sin embargo, la tesis resulta inseparable del «principio de oracionalidad»: los vehículos del significado son las oraciones, no los términos. Así, ‘conejo’ no puede referirse a los conejos, dado que su capacidad de referencia sólo tiene efecto, por decirlo así, en el contexto de una oración, y, una vez aquí, la tesis hace patente que el término es susceptible de ser entendido como refiriéndose a diferentes entidades – la tesis ni afirma ni implica la falsedad obvia de que un conejo sea una parte de conejo no separada del conejo – sin que por ello varíen las condiciones de emisión de cualquier oración que lo contenga. Además, continúa Føllesdal, insistir en que *yo sé* que cuando digo ‘conejo’ me refiero a los conejos, lleva a Searle a apelar a los «contenidos intencionales» que el hablante da a sus palabras, noción oscura y, como le recriminará Devitt, de un cartesianismo insoslayable.

Por otra parte, si bien Quine nunca desarrolló una defensa de sus ideas ante el tipo de crítica señalado, Davidson sí lo hizo – y gran parte de nuestro trabajo, claro, está dedicado a su exposición y análisis – . Resulta el gran mérito de Davidson aquí haber

mostrado cómo defender simultáneamente la indeterminación y la autoridad de la primera persona.

Así, la primera parte de este trabajo se dedica a exponer la tesis de la indeterminación, tal y como aparece en Quine primero, y en Davidson después. Se intenta subrayar, en especial, el marco conductista y antimentalista en el que la tesis surge en la obra de Quine, pero también cómo ésta resulta independiente, en gran medida, de dichos factores; este hecho se hará patente en la aceptación de la tesis por parte de Davidson, cuya teoría semántica no parte de los presupuestos quineanos citados. Resulta esencial en nuestra exposición de la tesis la diferenciación estricta entre la indeterminación del significado y la inescrutabilidad de la referencia. La primera de éstas era considerada por Quine una tesis de calado filosófico, al contrario que la última, vista, por un lado, definitivamente probada (por medio de las llamadas funciones vicarias) y, por otro, como una «idea nada sorprendente». Sin embargo, será ésta la que más centre nuestra atención, ya que la crítica searleana se dirige especialmente a ella. Por lo demás, se examinan en esta primera sección algunas cuestiones centrales para nuestro debate, como la diferencia entre in- e infradeterminación, la relevancia del holismo y la relación entre la indeterminación y la tesis llamada de la relatividad ontológica. El holismo es el elemento causante de la indeterminación del significado y se debe, por tanto, prestar atención a esta idea. No obstante, dado que, como hemos señalado, es la inescrutabilidad la idea que más parece preocupar a Searle, resulta aquí de vital importancia la tesis de la relatividad ontológica, pues Quine hace depender la una de la otra. La relatividad ontológica defiende que la ontología con la que un lenguaje o teoría

se compromete se halla indeterminada. Como se ve, la idea es hermana de la inescrutabilidad: tómense los términos de un lenguaje – nos vienen a decir ambas conjuntamente – y siempre será posible adscribir referencias incompatibles entre sí manteniendo la corrección empírica; estas diferentes referencias constituyen las diferentes ontologías o marcos ontológicos. Así, podemos elegir ver al nativo como comprometido con una ontología de partes de conejo no separadas o proyectar la nuestra y traducir su ‘gavagai’ por ‘conejo’.

A Davidson, sin embargo, no le satisfará esta carga ontológica que comporta dicha formulación de la tesis. Se trata en su visión, simplemente, de que una vez que se utiliza como metalenguaje – en una teoría interpretativa – un lenguaje con una cierta riqueza y complejidad, como son nuestras lenguas naturales, lo asombroso sería, más bien, que sólo pudiéramos interpretar una oración de una única manera. Una teoría semántica, según Davidson, trata de ofrecer las condiciones de verdad de cada oración del lenguaje objeto en el metalenguaje – y dar las condiciones de verdad es, según Davidson, una manera de dar el significado – y, como hemos dicho, la indeterminación/inescrutabilidad no supondría más que la posibilidad de hacerlo recurriendo a diferentes expresiones y conceptos. ¿Qué más da, se pregunta Davidson a este respecto, si al interpretar la oración «Sam is on Susan’s right» decimos que significa que Sam está a la derecha de Susan o que Susan está a la izquierda de Sam?

La segunda parte del trabajo comienza con una exposición de la crítica de J. Searle y la similar de P. Alston. Se analiza aquí, especialmente, el supuesto carácter paradójico de

la tesis de la indeterminación y la petición searleana de asumir una perspectiva de la primera persona en nuestro estudio de la semántica. Desde dicha perspectiva, afirmará Searle, se aprecia el carácter errado en su base del «conductismo semántico» y, una vez rechazado éste, desaparece la indeterminación. Se exponen las respuestas que ofreció Quine en un principio, apelando a un lenguaje de fondo y a un valor literal de las palabras, y en el resto de la sección se intenta mostrar cómo, aun estando Searle en lo cierto respecto a lo desacertado de estas respuestas, la tesis no comete la falacia de la que se la acusa. Esto es, una comprensión correcta de la tesis no la hace aparecer como autorrefutante. Nuestra argumentación se desarrolla en una doble perspectiva. Por un lado, se hace ver que la opción de Searle al rechazo de la tesis resulta sumamente insatisfactoria, apelando a un «contenido Intencional» de nuestras palabras del que poco o nada se nos dice, aparte de suponer un mentalismo rayano en lo mágico, como señalan autores como M. Devitt y D. Føllesdal. Fijar la referencia apelando a factores como el mencionado «Contenido Intencional» o la «referencia intencionada» (N. Georgalis) abren un camino que sólo puede conducir a un solipsismo y subjetivismo semántico poco acorde con el carácter público del lenguaje. Por otro lado, como se ha dicho, se intenta ofrecer una versión adecuada de la tesis de la indeterminación, de tal manera que ésta no cometa la falacia de la que Searle la acusa, y se utiliza para ello la diferenciación entre un nivel inmanente y un nivel trascendente respecto al significado. Esta diferenciación vendría a apoyar la concepción davidsoniana de la tesis. En el primero de los niveles, los hablantes explican la referencia de sus términos utilizando otros términos (o el mismo, en el caso homofónico); en el segundo, se trataría de fijar unívocamente una manera de dar la referencia de un término, y es aquí donde se aplica

la tesis de la indeterminación. Las críticas de corte searleano parecen, por un lado, confundir ambos niveles; es decir, del hecho de que, normalmente, explique la referencia de mi término ‘conejo’ diciendo que se refiere a conejos, se extrae la conclusión errónea de que la referencia de ‘conejo’ se halla unívocamente fijada, *al menos para mí*.

Con todo, no resulta difícil predecir la crítica que persistirá ante la versión de la tesis señalada. Los términos, aun en el metalenguaje, o los utilizados para dar la referencia en el nivel inmanente, deben tener una referencia determinada, so pena de abrir un regreso al infinito. Esto es, si explicamos la referencia de un término mediante otros que, a su vez, tampoco poseen una referencia determinada, nos veremos obligados a dar la referencia de éstos últimos, en un metalenguaje del metalenguaje, y así *ad infinitum*. Es ésta la misma crítica que, aquí justamente, dirigía Searle a la apelación quineana a un lenguaje de fondo. Por otro lado, dada la clara problemática de la apelación searleana al contenido Intencional, algunos autores han apelado a las *intuiciones* de los hablantes como el elemento determinante del significado o la referencia. Así pues, tenemos por un lado la labor de explicar cómo adquieren una referencia los términos utilizados en el llamado nivel inmanente y, por otro, la de aclarar en qué consisten y cómo consiguen fijar la referencia las llamadas intuiciones semánticas de los hablantes. Al análisis de estas ideas se dedica la tercera sección.

La tercera parte analiza, pues, esta forma actual que ha tomado la apelación a la primera persona respecto a la indeterminación: recurrir a esas llamadas «intuiciones

lingüísticas» de los hablantes. Se expone la idea, especialmente, tal y como aparece en la obra de D. Bar-On. La idea se sostiene sobre el bautizado como «verificacionismo semántico», a saber, que los hechos acerca del significado de nuestras palabras no pueden ser desconocidos para los hablantes competentes. De nuevo, nuestra argumentación tiene un doble objetivo aquí. Por un lado, se hace ver que la idea de que los hablantes competentes conocen todos los hechos acerca del lenguaje y, en especial, acerca de cómo funciona el mecanismo de la referencia, es, sencillamente, falsa. Ha sido M. Devitt quien con más contundencia ha subrayado este hecho. Los hablantes, viene a decir Devitt, para ser hablantes competentes necesitan conocer los mecanismos de la referencia tan poco como un buen ciclista necesita conocer el mecanismo de su bicicleta. Por otro lado, intentamos mostrar cómo la apelación a dichas intuiciones no resulta más satisfactoria que la apelación searleana al contenido Intencional.

Hay otra idea que ocupa, también, parte de nuestro análisis en este contexto, se trata de la teoría conocida como teoría «de la teoría» del lenguaje. Tanto las reflexiones de Quine y Davidson como de autores en una línea contraria, como Chomsky, se han referido al aprendiz de una lengua como un teorizador; parece desprenderse de aquí que aprender una lengua es confeccionar una teoría – una teoría del significado – . Ha sido también D. Bar-On quien más ha combatido esta idea, utilizando este rechazo de la teoría «de la teoría» contra la indeterminación. El argumento parece lógico: si la indeterminación surge, como ha remarcado especialmente Davidson, en el marco de la confección de una teoría semántica, debe quedar claro que aprender una lengua no es confeccionar o aprender teoría alguna, ni explícita ni implícitamente y, por tanto, la

indeterminación no tendría lugar. Volvemos la vista, entonces, al proceso mediante el que, según Davidson, se aprende el lenguaje, la llamada «triangulación». Se pretende mostrar aquí cómo las palabras adquieren el contenido que adquieren, y cómo, efectivamente, aprender una lengua no supone pergeñar una teoría semántica. Ésta constituye sólo una exposición hiperintelectualizada de la *habilidad* que adquiere el hablante. Ahora bien, no se ve cómo esto podría eliminar la indeterminación. Es el objetivo de esta tercera parte de nuestro trabajo mostrar cómo las intuiciones semánticas de los hablantes – juicios de nivel básico, se los llama – no tienen papel alguno en el proceso de aprendizaje de la lengua y, por tanto, no tiene sentido concebir la labor de la semántica como la sistematización de éstos. El hecho de que el aprendiz del lenguaje no confeccione una teoría semántica durante su aprendizaje debe hacer pensar, más bien, en el carácter inocuo de la indeterminación, como ha subrayado el propio Davidson; los hablantes utilizan sus palabras de la forma en que creen que serán entendidos como ellos desean por parte del oyente, el oyente intenta interpretarlas, y si éste pretende explicar qué quería decir el hablante, la indeterminación es inevitable.

Si bien el objetivo de nuestro trabajo, como hemos dicho, era, exclusivamente, analizar la crítica a la tesis basada en que ésta no toma en consideración la autoridad que la primera persona tiene sobre la semántica de sus palabras, y nuestra argumentación ha intentado demostrar que dicha crítica no es acertada, no parece conveniente dar por acabado el trabajo con esta consideración negativa – criticar la crítica – . Así pues, cerramos la tercera sección con unas consideraciones acerca de una adecuada concepción de la autoridad de la primera persona respecto a la semántica. Nuestras

consideraciones suponen, básicamente, una defensa de las ideas davidsonianas al respecto. Davidson ha señalado la manera de conjugar la indeterminación con dicha autoridad. La idea se resume así: dado que los hablantes se limitan a aplicar las palabras en las circunstancias en las que fueron entrenados para hacerlo, éstos saben sin necesidad de autointerpretarse el significado de sus palabras. Vista desde el punto de vista de la primera persona, la autoridad se basa, pues, en que, al contrario que a los demás hablantes, uno no necesita interpretarse a sí mismo para saber qué quiere decir, y en nada afecta a esta autoridad que otro hablante pueda traducir mis palabras a su idoloecto alejándose del modo homofónico. Prestamos especial importancia aquí a la analogía que Davidson establece entre la indeterminación y la posibilidad de utilizar diferentes escalas de medida para la temperatura o el peso. Con dicha analogía, Davidson pretende mostrar cómo en nada afecta la indeterminación a la «realidad» de los significados. Dicha «realidad» viene dada por lo que resulta invariante en todas las interpretaciones posibles: las oraciones que el hablante sostiene como verdaderas en determinadas circunstancias. Ésta es la materia objetiva, si se quiere decir así, que describen las diferentes teorías semánticas. Debemos a Davidson, mediante esta «teoría de las invariantes», el mérito de haber señalado cómo autoridad de primera persona e indeterminación no son incompatibles. Eso sí, se trata de una autoridad decididamente anticartesiana: no hay acceso a significados o conceptos en la mente, sino, sólo, como hemos dicho, la aplicación del patrón lingüístico aprendido.

Cerramos el presente trabajo con una sección de «conclusiones» donde resumimos la idea principal que hemos pretendido exponer a la vez que señalamos algunos puntos de máxima controversia filosófica en la actualidad en torno a nuestro problema.

1. LA TESIS DE LA INDETERMINACIÓN

Este capítulo, como hemos dicho ya, expone la tesis de la indeterminación del significado. La exposición es, digamos, cronológica, viendo primero la idea en la exposición quineana y después en la davidsoniana. Veremos cómo aparece la tesis en el marco de las doctrinas de Quine, cómo queda formulada y con qué argumentos es defendida. Se presta una especial importancia a las razones que apoyan el declarado extensionalismo quineano, es decir, su intento de ofrecer una explicación del lenguaje sin apelar a intensiones ni a entidades mentales, sino explicando nociones como verdad, significado o sinonimia a partir, exclusivamente, de las emisiones de los hablantes y de sus estimulaciones nerviosas. (Así, donde la totalidad del comportamiento verbal haga inexistente un «matiz semántico» que, intuitivamente, considerábamos relevante... el problema estaba en nuestra intuición, no en nuestra explicación. Ésta constituiría la idea subyacente de la tesis de la indeterminación).

Una vez vista la tesis según Quine, expondremos la formulación que de ésta ha realizado Davidson. Veremos cómo en su proyecto – hacer de una teoría tarskiana de la verdad una teoría del significado para las lenguas naturales – la indeterminación resulta inevitable.

La traducción radical al modo de Quine es un experimento mental (*Gedankenexperiment*, en el germanismo habitual en la jerga filosófica), algo así como el equivalente en filosofía del lenguaje al estado natural en filosofía política o en ética,

pues así como no se tiene constancia de la existencia histórica de tal estado, ninguno de nosotros suele verse durante su vida en el brete de realizar una traducción radical y, además, raras veces ha sucedido en la historia, pues pocas culturas se hallan completamente aisladas. No abundan los doctores Livingstone. El experimento, aun siendo mental, resulta útil cuando uno está interesado, como Quine, en aclarar nociones como sinonimia o analiticidad, o sea, en responder a las preguntas acerca de cuándo ha de considerarse que dos enunciados diferentes (compuestos por diferentes palabras), incluso de lenguas diferentes, *significan lo mismo*, y de cuándo ha de considerarse que un enunciado resulta verdadero – o falso – independientemente de la experiencia, por mera convención lingüística.

Mas si lo que nos interesa es dilucidar la noción misma de significado, podemos idear otro experimento que resultará de mayor utilidad: el de alguien que se ve inmerso en una comunidad sin conocer ni el lenguaje de dicha comunidad ni ningún otro. Ante las emisiones de los demás se verá obligado a interpretar directamente, sin verterlas antes a un lenguaje que él ya domina. Ya que aquí no le suponemos al sujeto del experimento dominio de lengua alguna, podemos decir que dicho experimento, al que llamaremos *interpretación radical*, es más «radical» que el de la traducción radical. Y, sin embargo, este trance sí que lo sufrimos todos y cada uno de nosotros, ya que la lengua materna se aprende siempre partiendo de cero. Para quienes, como Quine o Davidson, no comulgan con teorías innatistas del significado, si hay algo de lo que podamos estar seguros es de que el neonato no conoce lengua alguna.

1.1. La base del sistema de Quine

Conocer las ideas básicas de Quine acerca del lenguaje resulta, posiblemente, imprescindible para entender la tesis de la indeterminación. Saber cuál era el punto de partida nos ayudará a alcanzar el de llegada.

Quine defiende una visión radicalmente naturalista y conductista del hombre en general y del lenguaje en particular. Esta visión³ puede analizarse en tres puntos principales:

1. *Conductismo semántico*. Para Quine «el lenguaje es un arte social que todos nosotros adquirimos a partir, únicamente, de la evidencia de la conducta manifiesta [*overt*] de otra gente bajo circunstancias públicamente reconocibles» (Quine 1969a, p. 26). Ésta es una idea repetida continuamente en toda la obra de Quine. Así, por ejemplo, el prefacio de *Palabra y Objeto* comienza: «El lenguaje es un arte social. Para adquirirlo hemos de depender enteramente de los datos intersubjetivamente accesibles, como qué decir y cuándo. De ahí que no haya justificación para añadir [*collate*] significados lingüísticos, excepto en términos de disposiciones a responder de forma observable [*overtly*] a las estimulaciones socialmente observables» (Quine 1960, p. ix).

³ Que Gibson ha bautizado como la tesis NB, de *naturalistic-behavioristic*, naturalista – conductista.

De hecho, es ya una de las tesis de Quine más citadas que, en psicología, se puede ser conductista o no serlo, pero en materia de lingüística no hay opción: hay que serlo.

Defiendo que... la postura conductista es obligatoria [*mandatory*]. En psicología uno puede ser conductista o no serlo, pero en lingüística uno no tiene elección. Cada uno de nosotros aprende su lengua observando la conducta verbal de otra gente y haciendo que los demás observen la conducta verbal vacilante de uno y la refuercen y corrijan. Dependemos estrictamente de la conducta manifiesta [*overt*] en situaciones observables. Mientras nuestro dominio de la lengua supere todos los controles externos [*fits all external checkpoints*], donde nuestra emisión o nuestra reacción a la emisión de otro puede ser evaluada a la luz de alguna situación compartida, todo es correcto. Nuestra vida mental entre los controles es indiferente a nuestra evaluación sobre el dominio del lenguaje. No hay nada en el significado lingüístico más allá de lo que se puede advertir a partir de la conducta manifiesta [*overt*] en circunstancias observables. (Quine 1999, pp. 37-38).

La idea es sumamente coherente: si aprendemos el lenguaje sin tener acceso a ningún tipo de entidad mental de nuestros mentores, entonces es que *nada* que sea relevante para el lenguaje – incluidos, por supuesto, los significados – es mental. Todos hemos aprendido a aplicar la palabra ‘rojo’ a «la sangre, los tomates, las manzanas maduras y las langostas hervidas» (Quine 1974, p. 51). La supuesta idea asociada, la supuesta sensación asociada, será lo que sea, pero «el lenguaje se salta la idea y se asienta en el objeto. Pocas cosas son menos útiles que la idea para el estudio del lenguaje.» (*Ibid.*).

2. *Naturalismo*. El naturalismo quineano se basa en la idea de que no existe más conocimiento que el científico: no hay verdades filosóficas, metafísicas o como queramos llamarlas que sirvan de base a la ciencia. No hay una verdad fundamental o «filosofía primera» a la que se llegue por pura reflexión y que no se obtenga de datos científicos, ya que todos los datos científicos se obtienen partiendo de ella. La epistemología, candidata principal a ocupar ese puesto de filosofía primera entre los empiristas clásicos, forma también parte de la ciencia empírica, en concreto, de la psicología:

La epistemología, o algo similar, encaja perfectamente como un capítulo de la psicología y, por tanto, de la ciencia natural. Estudia un fenómeno natural, a saber, un sujeto físico humano. Este sujeto humano recibe un cierto *input* experimentalmente controlado – ciertos patrones de irradiación en frecuencias ordenadas, por ejemplo – y con el paso del tiempo el sujeto ofrece como *output* una descripción del mundo externo tridimensional y de su historia. La relación entre el magro *input* y el *output* torrencial es una relación que nos vemos inclinados a estudiar por, más o menos, las mismas razones que siempre han llevado a estudiar la epistemología: a saber, para ver cómo se llega de la evidencia a la teoría, y de qué maneras la teoría de la naturaleza de uno trasciende toda la evidencia disponible... pero una notable diferencia entre la vieja epistemología y la tarea epistemológica en esta nueva posición psicológica es que ahora podemos hacer libre uso de la psicología empírica. (Quine 1969a, pp. 82-83).

Quine señala que los empiristas clásicos temieron este tipo de respuesta por miedo a la circularidad: si la ciencia investiga partiendo de unos datos – sobre

cómo se adquiere ese conocimiento – , esos datos no pueden ser, a su vez, objeto de la ciencia, pues entonces pediríamos un nuevo sustento para las verdades científicas, entrando en un regreso infinito. Para Quine no hay tal circularidad. Cuando afirma que no hay nada más profundo que la verdad ofrecida por la ciencia, lo que está diciendo se aplica particularmente a lo que está diciendo, esto es, tampoco la idea de que no haya verdades metafísicas es una verdad metafísica. Es la ciencia la encargada de decir si las hay o no. Por ahora, la ciencia niega su existencia, si se demostrara – de forma científica – dicha existencia, habríamos de aceptarla. La idea de que todo nuestro conocimiento de la realidad física se produce por la información transmitida a través de nuestros receptores nerviosos no es metafísica empirista, sino un hallazgo científico más, tan falible como cualquier otro.

El peligro de circularidad al defender que la ciencia se da su propia base desaparece cuando concebimos la ciencia a la manera de Neurath, como una actividad que conjetura de la mejor manera que puede en cada momento dado, sin nunca llegar a verdades irrefutables. De hecho, *Palabra y Objeto* lleva como lema la famosa cita de Neurath donde se establece un bonito símil entre la ciencia y un barco en alta mar al que, para arreglarlo, no se puede traer nunca a dique, siempre hay que hacerlo en travesía: «Wie Schiffer sind wir, die ihr Schiff auf offener See umbauen müssen, ohne es jemals in einem Dock zerlegen und aus besten Bestandteilen neu errichten zu können», «Somos como

navegantes que han de reparar su barco en alta mar, sin poder traerlo nunca a dique seco y reconstruirlo con mejores piezas».

Podemos denominar esta visión *antifundamentismo*: el rechazo de que la ciencia deba (o pueda) tener una fundamentación no científica (filosófica o, más concretamente, metafísica).

3. *Antimentalismo*. En estrecha conexión con el naturalismo y conductismo que acabamos de ver, Quine se muestra declaradamente antimentalista. Lo que podríamos llamar su antropología coloca al ser humano en un mismo plano que al resto de los animales.⁴ La diferencia entre el ser humano y el resto de animales es, en todo caso, cuantitativa (memoria, habilidad psicomotriz...), no cualitativa. La mente ha de ser estudiada como un fenómeno más de la realidad física; en palabras de Quine: «Con Dewey mantengo que el conocimiento, la mente y el significado son parte del mismo mundo con el que tienen que ver, y han de ser estudiados con el mismo espíritu empírico que mueve a la ciencia natural» (Quine 1969a, p. 26).

La idea de Quine es que «las entidades mentales son inobjetables si se conciben como mecanismos físicos hipotéticos y se postulan exclusivamente con la intención de sistematizar los fenómenos físicos. Hay que postularlas con la esperanza de que se puedan someter un día a una explicación física completa»

⁴ No es casualidad que una de las últimas obras de Quine (Quine 1995a) comience: «Nosotros y otros animales...».

(Quine 1974, p. 49). En esa misma vena, Quine nos avisa de que no debemos creer que «la familiaridad con que usamos el modo de hablar mentalista sea claridad, ni hay que dejarse tentar por el mundo soñado de la introspección» (Quine 1974, p. 49).

Hemos hablado en la introducción del *antirealismo semántico* quineano. Como dijimos, podemos definir éste como el rechazo de los significados entendidos como entidades perfectamente aislables y con claros criterios de identidad. La teoría del significado de Quine, siguiendo a Dewey, considera que con la noción de «significado» no nos referimos a otra cosa que a una propiedad de la conducta humana. Los seres humanos, adiestrados de la forma adecuada, desarrollamos la disposición a emitir ciertos sonidos cuando experimentamos determinadas estimulaciones. Eso es todo. Sin embargo hay quien ha considerado que el ataque quineano a lo que el propio Quine bautizó como «mito del museo» se dirigía a una consideración mentalista del significado, pero en realidad se dirige a cualquier reificación de éste. Tal y como él lo expresa:

La objeción primaria persiste aun cuando no tomemos los objetos expuestos como ideas mentales sino como ideas platónicas o incluso como los objetos concretos denotados. La semántica está viciada por un mentalismo pernicioso mientras consideremos la semántica de un hombre de alguna manera determinada en su mente más allá de lo que podría estar implícito en sus disposiciones de conducta observable [*overt*]. Son los hechos mismos sobre el significado, no las entidades significadas, lo que ha de construirse en términos de conducta... (Quine 1969a, p. 27).

Este párrafo nos indica que debemos entender la tesis de la indeterminación de la traducción de forma más literal de la habitual, o sea, no como atacando la idea de que el significado sea una entidad mental, sino atacando la mera idea de que éste se halle determinado. A este respecto comenta Gibson: «Es importante hacer notar explícitamente que lo que Quine dice *es la conclusión* cuando, con Dewey, adoptamos una visión naturalista del lenguaje y una visión conductista del significado: (1) abandonamos la figura del lenguaje del museo, (2) abandonamos la seguridad de la determinación y (3) reconocemos que no hay significados o similitudes o distinciones de significados más allá de aquellas que se hallan implícitas en las disposiciones verbales de la gente» (Gibson 1982, p. 65).

El mito del museo es la idea de que existen unas entidades, los significados, con sus respectivos rótulos, las oraciones o términos. Es importante, como digo, percatarse de que Quine se opone a esta explicación del lenguaje aunque no concibamos los objetos expuestos en el museo como entidades mentales. No se trata aquí del antimentalismo comentado arriba. Es la idea misma de los significados como entidades concretas y determinadas, sean del tipo que sean, la que constituye un mito; aunque bien es cierto que la tesis surge en un principio como reacción a teorías del lenguaje con residuos mentalistas – como afirma M. Garrido:

El descubrimiento de la tesis de la indeterminación parece estar genética y lógicamente conectado con el deseo por parte de Quine de prescindir de la teoría del significado como carga de filosofía platónica y mentalista que gravitaba innecesariamente sobre el pensamiento empirista de Carnap y el Círculo de Viena. (M. Garrido 1976, p. 50).

1.2. Indeterminación e inescrutabilidad

La tesis de la indeterminación de la traducción parte de un experimento mental: imaginemos un lingüista que se introduce en un pueblo lejano de lengua y creencias completamente desconocidas. Éste se ve obligado a realizar una *traducción radical*, es decir, a confeccionar un manual de traducción de una lengua de la que no se conoce nada, con la que no ha habido ningún contacto anterior. Quine propone este experimento para ilustrar cómo nuestras teorías del lenguaje disponen, como *única* evidencia, de las emisiones que los hablantes realizan – o estarían dispuestos a realizar – según los estímulos sensoriales que reciben; éstos son los únicos datos de los que dispondría el hipotético traductor radical. La conducta verbal, pues, no debería ser descrita haciendo uso de nociones como significado o referencia, sino en términos de regularidades estadísticas entre emisiones y estímulos. Este «conductismo reduccionista», como se lo ha llamado⁵, es apuntado en el capítulo primero de *Palabra y Objeto*, y es el que conduce a la tesis de la indeterminación, expuesta en el célebre capítulo segundo.

Así pues, tenemos que la única evidencia que Quine admite para un traductor radical es la de las emisiones que los nativos realizan o se hallan dispuestos a realizar en

⁵ Véase, por ejemplo, H-J. Glock, 1996, p. 145.

determinadas circunstancias: «Nos interesa aquí el lenguaje en cuanto el complejo de disposiciones actuales para la conducta verbal.» (Quine 1960, p. 27). (Una gran ventaja de que el traductor radical realice un trabajo de campo – y no aprenda la lengua mediante grabaciones, por ejemplo – es que puede indagar sobre las disposiciones, ya que puede mostrar un conejo, pongamos por caso, y emitir en forma interrogativa «¿gavagai?», y ver si los nativos estarían dispuestos a emitir dicha oración).

El traductor, pues, habrá de comenzar su tarea observando la conducta verbal e indagando sobre las disposiciones lingüísticas⁶. Esto obliga a que el proceso de traducción comience con las *oraciones ocasionales*, es decir, aquellas oraciones cuyo significado se determina exclusivamente a partir de la información sensorial en el momento de la emisión. Las oraciones más cercanas a la observacionalidad absoluta, oraciones como ‘hace frío’ o ‘esto es rojo’, serían *oraciones observacionales*; y aquellas que se alejan de la observacionalidad hasta el punto de que la situación perceptiva del momento deja de ser relevante, son llamadas *oraciones permanentes (standing sentences)*⁷. Excepto en el caso de oraciones absolutamente observacionales entra en

⁶ Aquí resultará fundamental el reconocimiento de los gestos y palabras que el nativo utiliza para mostrar asentimiento y disentimiento. Quine no ve en ello un especial problema (v. *op. cit.*, pp. 29-30), pero otros autores han subrayado el hecho de que el significado de estos gestos y palabras se halla también indeterminado.

⁷ Antes de seguir, un apunte sobre el peculiar uso de Quine del término ‘oración’, que le lleva a tildar de oraciones expresiones como ‘conejo’ o ‘rojo’, que nosotros calificaríamos de términos o meras «palabras». Para Quine expresiones como ‘rojo’, ‘cuadrado’ o ‘ay’ han de ser consideradas oraciones, pues normalmente se aprenden por separado – habitualmente por definición ostensiva –. Incluso oraciones más largas pueden ser aprendidas como ‘ay’ o ‘cuadrado’, o sea, por condicionamiento directo de toda la emisión a una estimulación sensorial.

Evidentemente, no todas las oraciones se aprenden como un todo, sino que la mayoría se construyen con partes, por analogía con la forma en que esas partes han aparecido en otras oraciones que pueden haber sido aprendidas como un todo o no. Qué oraciones se han formado a partir de esta síntesis analógica y cuáles se han obtenido directamente, es cuestión de la historia olvidada de cada individuo. Para identificar

juego lo que Quine denomina *información intrusiva* o *colateral*, es decir, la información de la que el hablante disponga acerca de los objetos referidos. Citando de nuevo sus palabras:

Las oraciones *ocasionales*, al contrario que las *permanentes*, son oraciones como ‘Gavagai’, ‘Rojo’, ‘Hace daño’, ‘Su cara está sucia’, que provocan asentimiento o disentimiento solamente si son emitidas tras una estimulación apropiada. (Quine 1960., pp. 35-36).

Con todo, la observacionalidad debe contemplarse como una gradación: hay oraciones más o menos observacionales que otras; dependiendo de cuánta información colateral haya de tenerse en cuenta. Una oración como ‘Rojo’ es, por ejemplo, menos susceptible que ‘Conejo’ a las influencias de información intrusiva («hay menos margen para la información colateral al decidir si lo que se ve es algo rojo que al decidir si es un conejo», Quine 1960, pág. 41). (Como ejemplos de oraciones observacionales, Quine ha hablado también de «los gritos de los monos y las llamadas de los pájaros», Quine

mejor las expresiones que según este criterio han de ser consideradas oraciones, Quine las escribe con mayúscula inicial.

Ahora bien, lo que en cualquier caso resulta patente para Quine es que los límites entre lo que ha de ser tenido por una palabra o por una oración no son susceptibles de una total precisión. Podemos llamar a esta tesis la *indeterminación de la gramática*: «Los principios tras el uso del impresor de los espacios son oscuros, y la relevancia de tales principios para cualquiera de nuestras consideraciones lo es doblemente» (Quine 1960, p. 13). Esta tesis se sustenta en un supuesto ya apuntado: si la única realidad objetiva es la disposición de los hablantes en la conducta lingüística, entonces, hemos de admitir que esa única evidencia disponible no determina una única gramática del idioma en cuestión. Esto es, habrá siempre varias gramáticas extensionalmente equivalentes aunque mutuamente incompatibles que den perfecta cuenta de las disposiciones de los hablantes y que, por tanto, serán igualmente válidas. Escogeremos una en vez de otra en virtud de su simplicidad, de la facilidad que ofrece para traducir de un idioma a otro o de su superior capacidad para reflejar la psicología del nativo, pero todas estas son consideraciones puramente pragmáticas, que no afectan a la corrección en sí de todas y cada una de las gramáticas en cuestión.

1995b, p. 251). Y una como ‘Soltero’ es aún más susceptible que ‘Conejo’. Se debe tener en cuenta, además, que no existen oraciones puramente observacionales:

Incluso oraciones altamente observacionales no derivan lo que consideramos ordinariamente sus significados sólo de los nexos entre patrones de estimulación sensorial y nuestras disposiciones a asentir a o disentir de ellas. Sus nexos con otras expresiones en el lenguaje, y por tanto con la teoría del hablante sobre el mundo, son también relevantes. (R. Kirk 2004, p. 167).

Ya que el traductor radical no conoce las creencias de los nativos, su labor de traducción debe comenzar por las oraciones libres de información intrusiva. Por tanto, el primer paso de la tarea del traductor radical consistirá en el emparejamiento de oraciones observacionales de la lengua nativa con la hipotética traducción a la suya. Las oraciones observacionales, al hallarse libres de teoría, ofrecen este «punto de arranque», se convierten en «la entrada del traductor radical al lenguaje» (Quine 1993, p. 42). Esta fase no ofrece, según Quine, más problemas que los derivados del proceso inductivo que lleva al emparejamiento. En palabras de Quine:

Las emisiones traducidas en primer lugar y de manera más segura en un caso semejante [una situación de traducción radical] son las unidas a situaciones presentes que resultan visibles al lingüista y a su informante. [*The utterances first and most surely translated in such a case are ones keyed to present events that are conspicuos to the linguist and his informant*]. (Quine 1960, p. 29).

Hagamos un inciso sobre la traducción de las conectivas lógicas. Ésta no ofrece demasiados problemas, según Quine. Así, si el nativo asiente a una oración del tipo ‘gavagai guk worfi’ sólo si asiente por separado a ‘gavagai’ y a ‘worfi’, ‘guk’ vendría a ser la conjunción del español, si asiente a una oración como ‘gavagai ot worfi’ sólo si está dispuesto a asentir a ‘gavagai’ o a ‘worfi’ por separado, ‘ot’ es la disyunción del nativo, y de manera similar el traductor puede hallar el resto de conectivas. No obstante, subrayemos que incluso para obtener estas modestas conclusiones respecto a las conectivas lógicas, el traductor ha de ir algo más allá del estricto marco conductista: ha de utilizar el *principio de caridad*. Según este principio, nuestros manuales de traducción han de minimizar la adscripción de creencias falsas, especialmente respecto a las oraciones observacionales y a las conectivas lógicas. Ya que, como dice Quine, es «menos probable» que el nativo tenga creencias manifiestamente erróneas que nuestra traducción esté equivocada (Quine 1960, p. 59), o en palabras de Føllesdal: «hay una máxima de la interpretación, enfatizada tanto por Quine como por Davidson: si al interpretar las emisiones de una persona le atribuyes una opinión que nadie, razonablemente, podría tener, entonces ello cuenta contra tu interpretación» (Føllesdal 1990, p. 99). (Habremos de volver sobre el principio de caridad, ya que éste tiene en Davidson un papel mucho más fundamental del que Quine le confirió).

Ahora bien, la entrada en juego de información colateral origina ya indeterminación en el significado. Para determinar que la oración nativa ‘Gavagai’ es sinónima de la nuestra ‘Conejo’ necesitamos conocer la información de la que el hablante dispone: «La dificultad es que el asentimiento o disentimiento de un informante ante ‘¿Gavagai?’

puede depender excesivamente de información colateral anterior añadida al estímulo actual» (Føllesdal 1990, p. 37). A esto se ha de añadir la idea holista, es decir, las creencias del hablante – la información intrusiva – se hallan interconectadas de tal manera que unas se construyen sobre otras. Esto causa que el significado de toda oración no observacional dependa del significado de muchas otras oraciones. (En este problema se apoya el llamado «argumento desde arriba», que analizamos más adelante.)⁸ Tal y como lo expone un comentador:

Son las oraciones observacionales las que se hallan unidas a las estimulaciones; son los nexos entre una *oración* y otras los que tienen un papel en el significado de oraciones no observacionales. Podríamos exponer la cuestión así: Una oración observacional se halla directamente unida a las estimulaciones, y esos nexos determinan su uso correcto y, así, su significado. Para una oración no observacional, los nexos con las estimulaciones son igualmente importantes, pero, en este caso, esos nexos son parcial o totalmente *indirectos*. La oración se halla unida a otras oraciones que se hallan, por su parte, unidas a otras oraciones, y así hasta llegar a las oraciones observacionales. Este final es el que provee el significado empírico para una oración, por muy indirectos y complejos que sean los nexos. (P. Hilton 2004, p. 136).

Así, el traductor se percatará pronto de que no le resulta posible traducir muchas de las oraciones de la lengua nativa siguiendo el método utilizado con las oraciones más observacionales. Cuando las oraciones se alejan de la observacionalidad estricta no es posible correlacionarlas directamente con estímulos sensoriales simultáneos a la

⁸ Resulta célebre a este respecto el ejemplo sugerido por Davidson a Quine acerca de la mosca del conejo en el país del «gavagaiense». V. Quine 1960, p. 37.

emisión. El traductor necesita ahora recurrir a *hipótesis analíticas*, es decir, dejará de tratar las oraciones holofrásticamente y comenzará a analizarlas, las segmentará e intentará identificar términos, pronombres, artículos, etc., traduciéndolos según el paralelismo que halle entre la función de estos segmentos en la lengua nativa y en la suya propia. El traductor va aquí, por primera vez, más allá de de la mera correlación entre emisiones y estimulación sensorial.

La relevancia de las hipótesis analíticas en el proceso de traducción radical procede de la consideración quineana de la referencia como una relación que mantienen las oraciones y no los términos con la realidad – o estimulación sensorial – .

La referencia es una relación entre palabras – partes significantes constituyentes de oraciones – y objetos. La relación fundamental del lenguaje con el mundo, sin embargo, no tiene lugar a ese nivel, sino que se da al nivel de la relación entre oraciones y estimulaciones. Nos vemos obligados a atribuir significación a las partes constituyentes de las oraciones porque es sólo así como podemos ver los patrones y analogías en el lenguaje que nos hacen posible usarlo. Esta situación deja abierto el que, de hecho, podría haber más de una forma de atribuir estructura a nuestras oraciones, más de una manera de atribuir significación a sus partes constituyentes – más de una manera de analizarlas para darle un sentido a nuestro lenguaje. (P. Hilton 2004, p. 137).

Pero es en esta fase, con las hipótesis analíticas, donde entra en juego la referencia: el traductor intenta identificar las secciones de las oraciones que funcionan como términos y les asigna una referencia. Para fijar la referencia de los términos se hace necesario también fijar el aparato de individuación, es decir, pronombres, artículos y

cuantificadores. Lo relevante aquí, desde el punto de vista quineano, es que las hipótesis analíticas se hallan infradeterminadas por la evidencia. El traductor radical puede llegar inductivamente a la conclusión de que la oración nativa ‘Gavagai’ y la castellana ‘Conejo’ son estimulativamente sinónimas, y concluir, a partir de la similitud con el castellano, que ‘gavagai’ es un término general que se refiere a los conejos. El problema radica en que al hacer esto se está suponiendo que el nativo se compromete con la misma ontología que nosotros, cuando no tenemos ninguna razón para hacerlo. Podría suceder que el término nativo ‘gavagai’ no fuera un término general sino un término de masa que hace referencia a la porción de la Tierra ocupada por carne de conejo o un término abstracto que se refiere a instanciaciones del concepto platónico de la conejeidad.

Se puede pensar que la solución a este problema le vendría dada al traductor cuando consiguiera traducir el aparato de individuación. Así, por ejemplo, nuestro cuantificador existencial significa ‘hay algún’ porque hablamos de objetos (así, ya que los conejos para nosotros son objetos, decimos que hay algún conejo), si nuestra ontología fuera, por ejemplo, de masas, significaría algo así como ‘hay algo de’ (diríamos que hay algo de conejo en el sentido en el que decimos que hay algo de agua). Pero, obviamente, la traducción del aparato de individuación halla el mismo problema que la de los términos: toda la evidencia no determina una única traducción. En palabras de Quine: «La dificultad es fundamental. La verdad de las oraciones categóricas depende de los objetos, externos o inferenciales, de los que los términos componentes son verdad⁹; y

⁹ En la siguiente sección se ofrece una breve caracterización de las oraciones categóricas.

qué objetos son esos no está unívocamente determinado por los significados estimulativos» (Quine 1960, p. 61). Así, no podemos esperar a fijar la referencia de los términos para fijar la interpretación del aparato de individuación, ya que la referencia que asignamos a los términos depende de cómo interpretemos dicho aparato, y tampoco podemos esperar a fijar la interpretación del aparato para dar por determinada la referencia de los términos, ya que la interpretación que le damos al aparato depende de la referencia que le asignamos a los términos; y la traducción de ambos queda infradeterminada por toda la evidencia posible.

Es así como obtenemos la *inescrutabilidad de la referencia*.¹⁰ Los términos de una lengua pueden ser interpretados de diferentes maneras incompatibles entre sí pero todas ellas compatibles con toda la evidencia posible, luego no estamos justificados para pensar que hay una asignación de referencia que sea *la* asignación correcta. Por tanto, la referencia o el significado se hallan indeterminados, no meramente infradeterminados (v. 1.8).

En su artículo de 1970, «On the Reasons for Indeterminacy of Translation», Quine distinguió dos formas de defender la indeterminación. La que llamó *argumento desde abajo* [*argument from below*] defiende que la evidencia disponible (ahora y siempre) no puede bastar para hacernos escoger racionalmente entre traducciones alternativas estimulativamente sinónimas de un término. El llamado *argumento desde arriba*

¹⁰ Aunque Quine denominó así la tesis en un principio, resultaría preferible hablar, simplemente, de indeterminación de la referencia en vez de inescrutabilidad, para evitar la sugerencia de una especie de inaccesibilidad epistémica.

[*argument from above*] defiende que toda construcción teórica (y, por tanto, un manual de traducción) se halla infradeterminada por todos los datos disponibles (ahora y siempre). Algunos autores han considerado que la tesis es la misma bajo dos puntos de vista:

Si ese argumento funciona, [Quine] ha establecido lo que llama la «inescrutabilidad de la referencia» (o «de términos»). Es exactamente lo mismo que la tesis de la indeterminación de la traducción excepto que se aplica a términos y no a oraciones enteras[.] (R. Kirk 2004, p. 164).

Sin embargo, el propio Quine se adhirió a quienes ven en ambos argumentos dos tipos diferentes de indeterminación. A pesar de la controversia filosófica que ha suscitado, Quine llegó a considerar la inescrutabilidad de la referencia como una tesis de escaso alcance, demostrada inequívocamente por el argumento de las funciones vicarias (v. 1.3):

Se trata de la nada sorprendente idea [*It is the unsurprising reflection*] de que interpretaciones divergentes de las palabras en una oración pueden compensarse unas con otras de tal manera que mantengan una traducción idéntica de la oración como un todo. (Quine 1999, p. 50)

Así, la tesis realmente relevante para Quine sería la indeterminación del significado, que, no obstante, sólo sería una conjetura:

La indeterminación de la traducción que conjeturé hace mucho tiempo, y la indeterminación de la referencia que probé, son indeterminaciones en sentidos diferentes. Mi anterior uso de palabras diferentes, «indeterminación» para una e «inescrutabilidad» para la otra, habría sido más inteligente. (Quine 1997, p. 573).

En esta misma línea, B. Dreben llega a afirmar: «Al contrario que la indeterminación de la referencia, que tiene su prueba simple y conclusiva en las funciones vicarias, la indeterminación de la traducción fue siempre una conjetura, aunque una conjetura plausible.» (B. Dreben 2004, p. 289). En todo caso, el punto crucial aquí estaría en que «la indeterminación puramente holofrástica se dirige a y se limita (al menos al principio) a oraciones y el significado de oraciones, y no penetra en el nivel suboracional de los términos y sus significados o referentes.» (A. Orenstein 2002, p. 146).

Estructuremos, pues, nuestra exposición distinguiendo tres tipos de indeterminación: 1) indeterminación de la traducción, 2) de la referencia y 3) de comprensión. Expongamos brevemente cada una de estas ideas.

1. *Indeterminación de la traducción.* El argumento desde arriba defiende que es posible obtener manuales de traducción correctos que, sin embargo, no sean compatibles entre sí (que incluso asignen diferentes condiciones de verdad a una misma oración). Este argumento se basa, pues, en el de la infradeterminación de toda construcción teórica. Quine no ha ofrecido ejemplos a este respecto, ya que, según él esta indeterminación «rige de manera demasiado amplia en un

lenguaje como para permitir una ilustración factual [*draws too broadly on a language to admit of factual illustration*]]» (1999, p. 50), pero otros han intentado hacerlo por él. H. Field (1973), por ejemplo, afirma que podemos traducir teorías físicas antiguas a la teoría actual de diversas maneras, así, el término ‘masa’ de la física de Newton puede ser traducido tanto como el de ‘masa relativista’ como el de ‘masa en reposo’. Igualmente sucede con el término newtoniano ‘gravitación’, que podría ser traducido como ‘campo de fuerza’ o como ‘acción a distancia’ sin que haya ninguna diferencia empírica entre ambas opciones – es decir, con ambas traducciones se deducen las mismas oraciones observacionales –. Algún otro ejemplo que ilustre la tesis:

Para ofrecer un ejemplo, incluso bajo la guía del principio de caridad no está completamente determinado si debemos interpretar el antiguo término griego «areté» como expresando el mismo concepto que nuestro término «virtud» y atribuir diferentes creencias a los griegos o si, más bien, debemos decir que los griegos poseían un concepto para el que nosotros no tenemos una expresión lingüística, ya que nosotros, al contrario que ellos, no atribuiríamos virtud a los caballos. En cualquiera de las maneras que decidamos interpretar este término particular, ambas opciones interpretativas explicarán las diferencias en la conducta lingüística y se hallan igualmente apoyadas por la evidencia. (K. Stüber 1996, p. 341).

O un muy sencillo ejemplo propuesto por el propio Davidson:

Tú y yo, mirando con perfecta visión al mismo mantel, no nos ponemos de acuerdo acerca de si es violeta o azul. O mejor, uno de nosotros dice «es violeta», y el otro dice

«es azul». ¿Estamos en desacuerdo sobre el color o sobre el uso correcto de las palabras? Continuar hablando y hacer observaciones mutuas de ejemplares podría resolver la cuestión o podría no hacerlo. Podría no haber nada que resolver; podría no tener sentido resolverlo [*there may be no point in deciding*]. Cualquier tipo de evidencia, disponible y contrafáctica, podría ser consistente con ambas respuestas. Podemos decidir arbitrariamente el significado y hacer el ajuste obvio en la creencia, o hacerlo de la otra manera. (Davidson 1999a, p. 529).

Sobre cómo la infradeterminación genera indeterminación en el caso de los manuales de traducción, v. 1.8.¹¹

2. *Indeterminación de la referencia*. El argumento desde abajo defiende que podemos asignar diferentes referencias a los términos que constituyen las oraciones manteniendo constante el valor de verdad de éstas. (Obviamente, no podemos simplemente cambiar la referencia de los términos, cada cambio nos obligará a cambiar también la interpretación del aparato de individuación. Éste es un punto esencial sobre el que volveremos). Esto es, un manual podría traducir el término ‘gavagai’ – una vez analizada la oración pertinente y considerada dicha palabra como término – como ‘conejo’, mientras otro manual

¹¹ En una teoría del lenguaje como la quineana, que rechaza los significados más allá de los patrones de a/disentimiento, hay problemas para aclarar cómo dos teorías pueden resultar incompatibles entre sí y, sin embargo, tener las mismas conclusiones observacionales (deducirse de ellas las mismas categóricas observacionales). Este problema motivó que Quine fuera dando más peso en sus escritos al argumento desde abajo, llegando incluso a considerar que la tesis de la infradeterminación debía ser abandonada. Con todo, el argumento desde arriba puede sobrevivir sin infradeterminación, basta la conjunción de la tesis verificacionista – queriendo decir aquí que el significado de las oraciones teóricas se determina por las oraciones observacionales que se extraen de ellas – más el holismo – significando aquí que nuestras creencias forman una compleja red de interdependencias – . Para seguir las rectificaciones de Quine al respecto, v. D. Bar-on 1986.

podría traducirlo como ‘epifanía de la conejeidad’ (ya que no todos los manuales tienen por qué compartir nuestro marco ontológico), y los datos disponibles (de nuevo, ahora y siempre) harán tan correcta una traducción como la otra. Quine lo expresa así:

Pues, considera ‘gavagai’. ¿Quién sabe que los objetos a los que este término se aplica no son conejos después de todo, sino meros estados o breves segmentos temporales de conejos? En cualquiera de ambos casos las situaciones estimulativas que causarían asentimiento a ‘¿Gavagai?’ serían las mismas que para ‘Conejo’. O quizá los objetos a los que se aplica ‘gavagai’ son siempre partes no separadas de conejo; de nuevo el significado estimulativo no experimentaría diferencia. Cuando el lingüista pasa de la igualdad de significado estimulativo de ‘Gavagai’ y ‘Conejo’ a la conclusión de que un gavagai es un conejo entero duradero, simplemente está dando por sentado que el nativo es lo suficientemente similar a nosotros como para tener un breve término general para conejos y no un breve término general para estados o partes de conejo. (Quine 1960, pp. 51, 52.)¹²

En todo caso, tanto según un argumento como otro, tengamos en cuenta que cuando el traductor comienza a tratar con oraciones ocasionales, necesitará asignar creencias a los nativos para poder interpretar sus palabras; es decir, para saber qué quieren decir las

¹² Encontramos aquí el concepto de *significado estimulativo*, fundamental en Quine. Podemos definir el significado estimulativo, *grosso modo*, como los receptores nerviosos activados en las ocasiones en las que el hablante está dispuesto a emitir o asentir ante una oración. La noción presenta diversos problemas y el propio Quine se vio obligado a variarla repetidamente. (Sobre dichos problemas merece la pena el artículo de López Losada, 2002). Aquí nos interesa subrayar la idea de Davidson (en una reunión mantenida entre él, Quine, Dreben y Føllesdal en Stanford en 1986) de que *Palabra y Objeto* podría reescribirse sin hacer uso de la noción de significado estimulativo. La indeterminación sigue en pie si cambiamos el estímulo próximo, es decir, el nervioso, por el estímulo distal, es decir, la causa física causante de nuestra emisión.

palabras del nativo, necesita saber qué cree el nativo. Así, por ejemplo, el término ‘gavagai’ podría no referirse a conejos, sino a conejos con una mosca invisible en la oreja, si los nativos creyeran que todos los conejos van acompañados de una mosca invisible en la oreja (v. nota 8). Es más, no queda claro el límite entre aquello que forma parte del significado y lo que ha de contar como creencias sobre aquello a lo que las palabras se refieren. Este nexo entre significado y creencias se conoce como *tesis de la inextricabilidad*¹³. Si a esto añadimos un cierto holismo en nuestras creencias, es decir, la idea de que nuestras creencias se hallan estrechamente unidas, obtenemos la indeterminación. Por poner un ejemplo, si el nativo cree que los conejos van siempre acompañados de una mosca invisible en su oreja, y cree, por otro lado, que la Tierra es redonda y gira, obtenemos de inmediato varias posibles traducciones del término ‘gavagai’: ‘conejo y la Tierra es redonda’, ‘conejo con mosca invisible en la oreja’, ‘conejo con mosca invisible en la oreja y la Tierra gira’... ya que no sabemos si los nativos seguirían dispuestos a emitir ‘gavagai’ ante los conejos si ya no creyeran que van acompañados de una mosca invisible o que la Tierra gira. Así, inextricabilidad más holismo, da como resultado indeterminación. (Trataremos esta cuestión en detalle al analizar el holismo como fundamento de la indeterminación).

Así, no se ve claro si estos dos argumentos deben entenderse como dos líneas de defensa de una misma tesis – el argumento desde arriba basado en el holismo, el argumento desde abajo en la relatividad de la ontología (v. 1.3) – , o como dos

¹³ Bautizada así por Michael Dummett en *Truth and Other Enigmas*, Harvard University Press, Cambridge, 1978, pp. 377-378.

indeterminaciones diferentes que afectan al lenguaje. Algunos autores que se suman a esta última idea hablan de una indeterminación general del significado, por un lado, y la inescrutabilidad de la referencia, por otro:

[L]o que llamo la indeterminación general del significado, que debe distinguirse estrictamente de la tesis de la inescrutabilidad. La tesis de la indeterminación del significado es una consecuencia de abandonar la distinción analítico/sintético y del holismo presente en la construcción de una teoría interpretativa de la verdad. Los únicos criterios para la interpretación correcta de una oración específica son, al final, la concordancia total entre la interpretación de una oración específica y la interpretación de todas las demás oraciones. En este contexto, el intérprete tiene un cierto espacio para construir interpretaciones diferentes que están igualmente justificadas, debido a que puede cambiar pragmáticamente la atribución de creencia y concepto. (K. Stüber 1996, p. 341).

Por otra parte, es importante percibir aquí que el recurso a una lengua lejana no era necesario: la indeterminación comienza en casa. El argumento se desarrolla así: 1) la traducción de cada oración no puramente observacional está indeterminada, pues traducciones no equivalentes son igualmente corroboradas por la evidencia disponible; 2) si cada una de estas oraciones puede ser traducida de diferentes maneras incompatibles entre sí, todas estas traducciones pueden ser consideradas como el significado de la oración en cuestión; 3) por tanto, se podría interpretar a otro hablante de nuestra propia lengua de acuerdo con cualquiera de esos significados; 4) y ya que nosotros hablamos esa misma lengua cuyas oraciones pueden ser interpretadas de diferentes maneras incompatibles entre sí, podemos interpretarnos a nosotros mismos de cualquiera de esas

maneras. Resumiendo, no se trata sólo de que ‘gavagai’ pueda significar tanto ‘conejo’ como ‘parte no separada de conejo’ o ‘conejea’, sino que ‘conejo’, tanto en boca de alguno de nuestros interlocutores como en la nuestra propia, puede igualmente ser interpretada de cualquiera de esas maneras, pues nada en la evidencia disponible (ni posible) nos obliga a interpretarlo de una manera y no de otra. En efecto, ésta es una conclusión ante la que Quine no se arredra:

Nuestra ventaja con un compatriota es que, con una pequeña desviación, la hipótesis automática u homofónica de la traducción cubre el expediente. Si fuéramos perversos e ingeniosos podríamos abandonar esa hipótesis y pergeñar otras hipótesis analíticas que atribuyeran creencias no imaginadas a nuestro compatriota, y que estuvieran conformes con todas sus disposiciones de respuesta verbal en todas las estimulaciones posibles. Pensar en términos de traducción radical de lenguas exóticas ha ayudado a hacer vívidos los factores, pero la lección principal que hemos de extraer tiene que ver con la laxitud [*slack*] empírica en nuestras creencias. Ya que nuestras propias creencias podrían ser sustituidas por esas otras atribuidas a nuestro compatriota en el poco práctico juego imaginado, no podrían sobrevenir nunca conflictos con la experiencia, excepto aquellos que también aparecerían en nuestras razonables creencias actuales. La traducción radical de oraciones está infradeterminada por la totalidad de las disposiciones de conducta verbal en el mismo grado en el que nuestras propias teorías y creencias en general están infradeterminadas por la totalidad de la evidencia sensorial posible de todos los tiempos (Quine 1960, p. 78).

Como hemos señalado, Quine no ha sido pródigo en ejemplos respecto a la indeterminación del significado; respecto a la indeterminación de la referencia, las funciones vicarias ofrecen innumerables ejemplos. Con todo, Quine ha ofrecido un

ejemplo en una lengua natural real, relativo a los numerales japoneses. Existen en japonés unas partículas llamadas «clasificadores» de las que Quine propone dos interpretaciones. Normalmente son interpretados como conectados con los numerales para formar numerales compuestos. Si se le añade un cierto clasificador, el numeral resulta adecuado para contar objetos grandes, si se le conecta otro determinado clasificador, el numeral se hará servir ahora para contar objetos finos, por ejemplo, palillos o lápices. Pero podemos ofrecer interpretaciones alternativas de los clasificadores. Podemos considerar que éstos no forman parte del numeral sino del término mismo al que acompañan. Así, el clasificador haría la misma función que realizan en castellano expresiones como «vara de» aplicada a términos como «madera» o «hierro».

Así pues, la oración castellana ‘5 bueyes’ se expresaría en japonés mediante tres palabras: la primera, el numeral ‘5’; la segunda, el clasificador correspondiente para contar objetos grandes y la tercera equivaldría a ‘buey’. Así, según esta interpretación, el clasificador acompaña al numeral para hacerlo apto para contar bueyes. La oración vendría, entonces, a significar ‘5 bueyes’. Según la otra interpretación, el clasificador estaría acompañando no al numeral, sino al término, con lo que éste pasaría de ser el término individuativo ‘buey’ a corresponder más bien al término de masa ‘ganado’. El clasificador habría de considerarse ahora como una expresión del tipo ‘cabeza de’. Ahora obtenemos una oración que viene a significar ‘5 cabezas de ganado’ y, por tanto, el término individuativo referente a bueyes ha pasado a ser un término de masa que se refiere al ganado. Quine dice a este respecto:

La simplicidad de este ejemplo, cuyo propósito es puramente aclaratorio, necesita la connivencia del lector. Ha de entenderse ‘ganado’ como un término de masa que incluye sólo bovinos, y ‘buey’ como aplicable a todos los bovinos. Que estas usanzas no son invariables no se cuestiona. La cuestión es que la frase japonesa resulta ser ‘cinco bovinos’, que era lo deseado en cualquiera de los dos modos en que esté explicada. Un modo trata la tercera palabra japonesa como un término individuativo verdadero de cada bovino, y el otro modo la trata como un término de masa que abarca la totalidad no individualizada de carne bovina sobre pezuñas. Son dos modos muy diferentes de tratar la tercera palabra japonesa y la frase de tres palabras entera resulta correcta en ambos casos solamente a causa de las diferencias compensatorias en nuestra versión de la segunda palabra, el clasificador. (Quine 1969a, pp. 55-56).¹⁴

¹⁴ Putnam ha desarrollado otro ejemplo de indeterminación que resulta sumamente interesante. Quine, al hacer hincapié en los diferentes marcos ontológicos del tipo de conejos o partes no separadas de conejos y Davidson al hacerlo en la aplicación continua del principio de caridad y en la apelación al estímulo distal, pueden dejarnos la impresión de que la indeterminación, si se acepta la tesis, no llega nunca a extremos tan dramáticos como en realidad puede llegar. Putnam, efectivamente, defiende la inescrutabilidad de la referencia y la indeterminación del significado e incluso las radicaliza más:

Quine dice simplemente lo que yo dije, que *las condiciones de verdad para oraciones enteras* infradeterminan la referencia. Ya que ‘fragmentos de conejo’, ‘conejeidad’ y ‘partes no separadas de conejo’ tienen todas una estrecha conexión con los conejos, uno puede salir de *Word and Object* con la impresión de que todas las reinterpretaciones que dejan el valor de verdad de una oración invariable están al menos íntimamente conectadas con la interpretación estándar (de la forma que partes de conejo y conejeidad están conectadas con los conejos). El argumento expuesto en el apéndice e ilustrado en este capítulo muestra que las condiciones de verdad para ‘Un gato está en la alfombra’ ni tan siquiera excluyen la posibilidad de que ‘gato’ se refiera a *cerezas*. (Putnam 1981, p. 35).

Véase, pues, el ejemplo propuesto por Putnam en Putnam 1981.

Se puede percibir el teorema de Löwenheim-Skolem como el antecesor de la indeterminación de la referencia. Este teorema defiende que para cualquier sistema formal expresado en el cálculo funcional de primer orden se pueden hallar diversos modelos (básicamente, interpretaciones); por tanto, cualquier lenguaje formal de este tipo, independientemente de nuestras intuiciones acerca de los objetos sobre los que habla, puede ser interpretado de diferentes maneras que dejan intactos los valores de verdad de todas las posibles fórmulas. El propio Quine ha expresado su idea de una forma similar:

Podemos concebir el vocabulario de una teoría como conteniendo signos lógicos, como por ejemplo los cuantificadores, y los signos para las funciones de verdad y la identidad, y, además, signos descriptivos o no lógicos (...) Supón ahora que en las oraciones que contiene la teoría (...) extraemos los significados del vocabulario no lógico y de las variables. Nos quedamos con la forma lógica de la teoría o, como yo la llamo, la *forma de la teoría* [*theory form*]. Ahora podemos interpretar esta forma de la teoría escogiendo un nuevo universo para sus variables de cuantificación y asignando objetos de este universo para los nombres, y escogiendo subclases de este universo como extensiones de los predicados monádicos, etcétera. Cada interpretación tal de la forma de la teoría es denominada un modelo para ella y si sirve, es verdadero. (Quine 1969b, pp. 53-54).

3. *Indeterminación de la comprensión.* Podemos dar este nombre (no lo hace Quine ni la bibliografía sobre la cuestión, ya que no se trata de una tesis nueva y distinta de las anteriores, pero puede resultar útil a efectos hermenéuticos) a la forma netamente conductista de entender la comunicación humana que nos

queda si aceptamos la indeterminación del significado y la referencia. Podríamos formular este argumento de la siguiente manera: ya que ni la referencia de nuestros términos ni el significado de nuestras oraciones es *algo* que se halle en el cerebro de los hablantes – ni en ningún otro lugar –, no cabe hablar de cuándo un hablante dado «posee» o «ha captado» dicha referencia o dicho significado. La comprensión de la lengua viene a ser una cuestión de grado, y ese grado se mide según el acoplamiento de la conducta personal y la del resto de la comunidad lingüística, y cuándo ese acoplamiento es total y cuándo no, es algo imposible de precisar. En palabras del propio Quine:

En la práctica, consideramos que alguien entiende una oración si no nos sorprendemos por su reacción al oírla – a condición, además, de que su reacción no sea de asombro visible. [...] Con todo, no hay una frontera evidente, un criterio general para decidir si realmente malinterpreta la oración o simplemente mantiene alguna teoría inusual sobre la materia en cuestión.

Entender, visto de forma conductista, es, pues, un efecto estadístico: consiste en multiplicidades. El núcleo es la palabra, y la masa circundante está constituida por las incontables oraciones en las que aparece la palabra. Una coloración predominantemente saludable o insana es lo que cuenta como entender o no entender la palabra y las oraciones; y no se ha de buscar una frontera nítida. [...] La lexicografía no tiene necesidad de la sinonimia, como vimos, y tampoco tiene necesidad de una distinción nítida entre entender y no entender. La tarea del lexicógrafo es mejorar la comprensión del lector de las expresiones, pero puede conseguirlo sin trazar una frontera. Hace lo que puede, dentro de un estrecho margen, para ajustar la conducta verbal del lector a la de la comunidad, o a la de algún sector escogido de ella. El ajuste es una cuestión de grado, y

éste es vago: una cuestión de fluidez y efectividad en el diálogo. (Quine 1999, pp. 58, 59).

Como vemos, la tesis quineana es, sin duda, la que con mayor justicia puede ser calificada de «escepticismo semántico». Sólo en el caso de las conectivas de la lógica proposicional y de las oraciones puramente observacionales (esto es, oraciones del tipo de ‘Rojo’ o ‘Frío’) acepta Quine la determinación del significado. Según las oraciones se van distanciando de la pura observacionalidad y van ganando carga teórica, su significado pierde concreción y va quedando cada vez más indeterminado. (Además, recordemos, no existen oraciones cien por cien observacionales).

Hemos de subrayar, para acabar nuestra exposición, que lo que afirma la tesis de la indeterminación del significado no es que los hechos empíricos no sean suficientes para averiguar el significado de las oraciones. Lo que sucede es que no hay nada que averiguar, no hay materia objetiva (*fact of the matter*) cuando hablamos de significados. Pensar que, de los varios significados aparentemente dispares que podemos atribuirle a cada oración, sólo uno puede ser el correcto sería, para Quine, mera superstición. Superstición apoyada por la ciega creencia de que un bilingüe perfecto sí podría decir cuál es la traducción correcta – lo que serviría como prueba indirecta de que el significado sí está determinado, pues ya hemos visto de qué forma se hallan interrelacionadas las dos ideas – , ya que está a su alcance el trascender la evidencia empírica y comparar las *ideas* que expresan las diferentes oraciones. La respuesta de Quine suena así:

Una (...) causa de la incapacidad de percatarse de este punto es un tozudo sentimiento de que un auténtico bilingüe está con toda seguridad en condiciones de hacer correlaciones de oraciones entre sus lenguas unívocamente correctas. Este sentimiento está alimentado por una acrítica teoría mentalista de las ideas: cada oración y sus traducciones admisibles expresan una idea idéntica en la mente del bilingüe. El sentimiento también puede sobrevivir tras la repulsa de las ideas: uno aún puede argüir que la oración y sus traducciones corresponden a alguna condición neuronal idéntica aunque desconocida en el bilingüe. Admitamos eso ahora; sería afirmar, simplemente, que el bilingüe tiene su propia correlación semántica – que se halla, de alguna manera, en sus nervios. Mi idea se mantiene, ya que mi idea es que, entonces, otro bilingüe podría tener una correlación semántica incompatible con la del primer bilingüe sin desviarse del primer bilingüe en sus disposiciones verbales dentro de cualquiera de las lenguas, excepto en sus disposiciones de traducción. (Quine 1960, p. 74).

Así pues, rechazar la indeterminación de la traducción o la inescrutabilidad de la referencia es algo que, según Quine, sólo hará quien se comprometa con una «acrítica teoría mentalista de las ideas». Y es que la tesis de la indeterminación ataca, en realidad, tanto a los significados como a las actitudes proposicionales, ya que si la tesis es cierta, nuestras oraciones no expresan un contenido proposicional concreto, con lo que también deseos, temores o creencias se ven afectados de indeterminación. La tesis de la indeterminación rechaza, pues, cualquier tipo de entidad tradicionalmente considerada intencional. Unger expresa esta idea así:

Supón que quieres un conejo de mascota. Ya que tus pensamientos están conectados con tu lenguaje, tus deseos han de estar expresados claramente en tus palabras. Si traduzco tu ‘gavagai’ como se indicó, describiré tu deseo como queriendo tener unas partes de conejo no separadas para acariciarlas. (P. Unger 1984, p. 18).

Esto implica que la indeterminación afecta también a nuestros estados mentales. Por tanto, aun cuando la tesis haya nacido en el marco antimentalista/fisicalista de la filosofía quineana, no va irremediabilmente unida a tales ideas:

La indeterminación se aplica también a estados mentales. Esta extensión de la tesis de la indeterminación de las expresiones lingüísticas a estados mentales no hace asunciones sobre el fisicalismo. Podemos admitir que hay estados mentales y aún obtenemos indeterminación. (...) [Este argumento] pone de manifiesto lo que considero la clave de la indeterminación de la traducción: no proviene primariamente del fisicalismo, sino de la naturaleza pública del lenguaje. (Føllesdal 1990, p. 107).

1.3. Relatividad ontológica

La tesis de la relatividad ontológica constituye un pilar fundamental en Quine para su defensa de la indeterminación. Es más, el propio Quine comenta en *La búsqueda de la verdad* que la tesis de la inescrutabilidad y la de la relatividad ontológica nunca estuvieron claramente separadas en su cabeza; la sutil distinción fue labor de lectores avezados.

La tesis de la relatividad de la ontología defiende la (desde Quine casi obvia) idea de que existen marcos ontológicos diferentes. Como vimos, es posible interpretar una

lengua extraña según diferentes marcos ontológicos sin que esto repercuta en la fluidez de nuestra comunicación con los nativos. Así, el hablante del «gavagaiense» puede ser visto como refiriéndose con sus términos a objetos físicos tanto como a epifanías de conceptos platónicos, partes de objetos no separados del resto del objeto, estadios espacio-temporales de objetos, etc. Y lo que resulta aún más curioso: podemos dejar de interpretar a los hablantes de nuestra comunidad lingüística como lo hacemos (esto es, de forma homofónica: donde lenguaje objeto y lenguaje sujeto coinciden) y comenzar a imputarles marcos ontológicos alternativos. Y la conclusión que Quine extrae de este hecho es que si podemos interpretar a los hablantes imputándoles diversas referencias incompatibles entre sí, y cada una de estas interpretaciones nos sigue ofreciendo un diccionario compatible con todos los datos empíricos disponibles (y posibles), la referencia pierde su aparente univocidad. Normalmente damos por sentado un marco ontológico determinado, pero lo hacemos de forma injustificada: muchos otros resultarían igualmente válidos. No disponemos de ningún dato empírico (al menos en el sentido quineano de empírico: conducta verbal y no verbal observable de los hablantes) que nos haga imputar a los hablantes de una lengua – o incluso a nosotros mismos – un marco ontológico y no otro. Respecto a leer en nuestros propios términos otra referencia a la habitual y a la disolución de la misma, dice Quine: «Ésta era la inescrutabilidad de la referencia aplicada a nosotros mismos; y convertía la referencia en un absurdo. Es justo; la referencia *es* un absurdo excepto relativa a un sistema de coordenadas.» (Quine 1969a, p. 48). Es absurdo, según Quine, preguntar si nuestros términos ‘conejo’, ‘parte no separada de conejo’, etc. *realmente* se refieren, respectivamente, a conejos, partes de

conejo, etc., ya que los términos no se refieren *realmente* – de forma absoluta – a nada. Además, la propia pregunta nos lleva a un regreso infinito:

Es absurdo preguntar esto de forma absoluta; podemos preguntarlo con sentido sólo relativamente a algún lenguaje de fondo. Cuando preguntamos “¿‘conejo’ se refiere realmente a conejos?” alguno puede contestar con la pregunta: “¿Se refiere a conejos en qué sentido de ‘conejos’?”, comenzando así un regreso; y necesitamos el lenguaje de fondo para regresar; el lenguaje de fondo le da sentido a la pregunta, aunque sólo sea un sentido relativo; un sentido relativo a éste, al lenguaje de fondo; preguntar por la referencia en algún sentido más absoluto sería como preguntar por la posición absoluta, o por la velocidad absoluta, en vez de por la posición o la velocidad relativas a un marco de referencia dado. Asimismo, es muy similar a preguntar si nuestro vecino no puede estar sistemáticamente viéndolo todo boca a bajo, o en un color complementario, y por siempre indetectable. (Quine 1969a, pp. 48-49)

Las que se conocen como *funciones vicarias* nos permiten pasar de un marco ontológico a otro. Esto es, estas funciones realizan las permutaciones necesarias para asegurar que el cambio de marco mantiene la nueva interpretación tan acorde con los datos empíricos como la anterior. Así, si pasamos de una ontología reísta a una de instanciaciones de conceptos, el cuantificador universal habrá de ser leído ‘toda instanciación de’ o ‘cualquier instanciación de’, en vez de ‘todo’ o ‘cualquier’, y el cuantificador existencial ‘algún’ pasará a leerse ‘alguna instanciación de’. Por ello, para Quine, ya que «la referencia es un sinsentido» – al menos en un sentido absoluto, no relativizado – , lo que sí tiene sentido es preguntar cómo se pasa de un marco ontológico a otro – las funciones vicarias que se deben utilizar – , no cuál es exactamente la

referencia de los términos de una lengua o una teoría dada: «Lo que tiene sentido no es decir cuáles son los objetos de una teoría, hablando de manera absoluta, sino cómo una teoría de objetos es interpretable o reinterpretada en otra.» (Quine 1969a, p. 50). De este fragmento dice Gibson que

expresa la esencia de la doctrina de Quine de la *relatividad ontológica*, que descansa, naturalmente, en la concepción NC [naturalista-conductista, v. nota 3] del lenguaje: ya que la doctrina de la relatividad no es más que una generalización de la doctrina de la inescrutabilidad de la referencia, la que, a su vez, es una consecuencia de la concepción NC del lenguaje. Si la tesis NC es verdadera, la referencia es inescrutable; y la referencia será escrutable sólo relativamente a una traducción que no es la única posible. Pero la referencia de los términos de un lenguaje o una teoría no es otra cosa que la ontología del lenguaje o teoría. Así, si la referencia es relativa, también lo es la ontología. (Gibson 1982, p. 76).

Se entiende, pues, por qué la posibilidad de hallar funciones vicarias es el argumento decisivo a favor de la tesis de la indeterminación.

Se genera aquí un problema que trataremos con más detenimiento en secciones ulteriores: Quine da por fijada la referencia de los términos del metalenguaje para evitar un regreso al infinito de lenguajes de fondo (v. 2.1.2), pero salta a la vista que resulta una contradicción mantener que la referencia sólo cobra sentido relativamente a una ontología de fondo y que la referencia de los términos del metalenguaje sí poseen una referencia concreta determinada, es decir, sin ser relativa a otro metalenguaje.

Efectivamente, la pregunta que hemos de hacer ante la tesis de la relatividad de la ontología es a qué es relativa. La respuesta es: al manual de traducción. Cada manual de traducción elige un marco ontológico de entre muchos posibles y esta elección se halla infradeterminada por todos los datos empíricos posibles. La ontología, pues, es relativa al manual de traducción que hayamos escogido. Hemos de subrayar que Quine habla de una doble indeterminación. La traducción de nuestras oraciones es relativa al manual de traducción – toda la evidencia empírica no determina unívocamente un manual – y la ontología de fondo del lenguaje al que traducimos – toda la evidencia empírica no determina unívocamente una ontología – .

Por otro lado, Quine ha realizado interesantes especulaciones acerca de nuestra ontología y la forma de adquirirla.

En *Las Raíces de la Referencia*, Quine explica cómo, en un principio, la ontología del niño debe de componerse exclusivamente de masas. Es decir, así como ‘Rojo’ para él indica una propiedad que a veces percibe, lo mismo sucede con ‘Agua’ o incluso ‘Mamá’. De estas palabras dice Quine que son términos observacionales, susceptibles de aprendizaje por ostensión; y en todos esos casos, haber aprendido el término es haber aprendido cuándo hay que asentir a él o disentir de él utilizado como sentencia ocasional. Tanto mamá como agua o rojo son presencias recurrentes reconocibles. Así, aunque mamá es un cuerpo y rojo o agua no lo son, las oraciones de una palabra ‘Mamá’, ‘Rojo’ o ‘Agua’ son aprendidas de forma similar. Pronto vendrá la consideración de mamá como un cuerpo y no como una masa, con lo que ‘Mamá’, que

fue considerada holofrásticamente como una oración de una palabra, se convertirá en un término singular concreto.

Viene ahora el aprendizaje de ‘perro’. Para ello no basta, como en el caso de ‘Mamá’ o ‘Agua’, la simple percatación de la presencia de mamá o de agua. Ahora necesitamos dividir la referencia. Hemos de aprender qué hemos de considerar *un* perro, cuando dos apariciones de perro son apariciones del *mismo* perro y cuando se trata de *otro* perro.

‘Mamá’ es un término singular, y su referencia es un objeto, un cuerpo físico; ‘agua’ o ‘rojo’ son términos de masa y su referencia es una masa dispersa por el mundo no uniformemente; ‘perro’ es un término general, y su referencia es un conjunto de objetos, a cada uno de los cuales se le aplica el término.

El aprendizaje de términos generales es de una complejidad mucho mayor que el de los otros tipos, ya que el niño no sólo ha de apreciar cuándo dos apariciones de un objeto son apariciones del mismo objeto, como en el caso de ‘mamá’, sino también cuando dos objetos se parecen lo suficiente para ser los dos apariciones de perros. Es, por decirlo de alguna manera, una semejanza de segundo grado, una semejanza de semejanzas. De esta manera aprendemos también predicaciones como ‘la sangre es roja’, donde subsumimos una masa, la sangre, en otra, el color rojo. A partir de este tipo de construcciones, formamos otras como ‘rojo es un color’ o ‘el cuadrado es una forma’, donde el término que se halla en posición de sujeto funciona como término singular abstracto. ‘Rojo’ no es aquí un término de masa, ya que no denota una masa esparcida por el mundo, sino un

concepto que engloba todas las superficies rojas. Así, podemos ver que los términos singulares abstractos y los términos singulares concretos resultan difíciles de discernir, ya que podemos considerar ‘rojo’ aquí como un término que se aplica a un objeto: el constituido por la unión de todas las superficies rojas. Andando el tiempo, el niño aprenderá también a utilizar términos generales abstractos, como ‘cuadrado’.

La composición atributiva será también la que permita al niño la formación de oraciones eternas, aquellas cuyos valores veritativos están fijados definitivamente, con independencia del hablante y la ocasión. Así, el infante puede formar ‘nieve blanca’ o ‘la nieve es blanca’ mediante el sencillo mecanismo de unir dos términos de masa aprendidos con anterioridad, recordando que siempre han aparecido solapándose. Éste sería el tipo más sencillo de sentencia permanente, de mayor complejidad sería ‘Donald es un pato’, donde el sujeto es un término singular y el predicado un término general.

Respecto a los términos sincategoremáticos, el aprendizaje es sencillo. Así, por ejemplo, el niño aprende que se asiente a una construcción con ‘y’ cuando se asiente a cada una de las oraciones por separado. También se aprenderá a asentir a una construcción con ‘en’, como ‘mamá en el jardín’ cuando el objeto denotado por el segundo miembro de la expresión engloba o contiene el denotado por el primero.

Así pues, según el propio Quine, al aprender los términos aprendemos la ontología a ellos asociada. Términos singulares, términos de masa, términos abstractos... cada tipo de término sirve para expresar una determinada ontología. Cuando ante una nueva

superficie roja, el niño emite la palabra ‘rojo’, ¿podemos decir que el niño ha aprendido ya a referirse al color rojo? ¿Por qué no interpretar su respuesta, en vez de refiriéndose al nombre del color, como utilizando un término general con el que denota cada porción roja de la superficie? ¿O un término general con el que se refiere a cualquier mancha roja entera, pero no a partes menores de tal mancha? ¿O un término general que denota todo episodio completo que exhiba llamativamente el rojo? Decir que el niño se refiere al color, en este estado prematuro, sería imputarle a él nuestra propia ontología sin justificación suficiente.

¿Cómo saber, pues, cuándo el niño domina nuestro aparato de referencia? Como hemos visto, la referencia implica más que la mera capacidad de percatarse de una presencia. Es necesario dominar lo que Quine llama el *aparato referencial* del lenguaje: pronombres, cópulas y terminaciones del plural.

Hemos visto que en el aprendizaje del lenguaje no hallamos diferencias relevantes entre los términos singulares, cuya referencia es un objeto, y los términos de masa, cuya referencia es una sustancia dispersa de forma no uniforme. Vimos también que es con los términos generales con los que el niño aprende a dividir la referencia. El ser humano, siempre según Quine, tiene una mentalidad innatamente somática («la inmemorial doctrina de objetos duraderos de tamaño medio»); la reificación nos ha dado grandes ventajas en la lucha por la existencia. De ahí que dominemos con relativa rapidez un mecanismo bastante complejo. «Los cuerpos son la realidad primaria, los objetos *par excellence*».

El pronombre relativo tiene un papel primordial a la hora de aprender nuestro aparato referencial. Así, en la oración ‘vi al hombre que te regaló una flor’ el pronombre relativo ‘que’ o (‘el cual’) nos permite sustituir en la oración el nombre del objeto. La cláusula relativa sirve de término general que, cuando se predica del objeto (el hombre), reafirma la sentencia originaria. Lo importante aquí es que la construcción de relativo es esencial para que el niño interiorice nuestra ontología de objetos. El niño ha de entender que ‘Veo la luna’ equivale a ‘la luna es *una cosa que yo veo*’.

A la hora de interiorizar nuestra ontología reista cobra especial importancia el dominio de los cuantificadores. No es difícil, sin embargo, imaginar cómo se aprende el mecanismo de la cuantificación universal y existencial. El aprendiz observará que el resto de hablantes sólo asienten a ‘Todo x es F ’ si para cualquier sustitución de x asienten a Fx . Y asimismo verá que los hablantes que lo rodean sólo disienten de ‘hay un x tal que Fx ’ si para ninguna sustitución de x asienten a Fx . A partir de este momento, el niño ya está preparado para hablar de *todos* los perros, *algunos* animales, *ningún* pájaro...

Tengamos en cuenta que ‘negro’ y ‘cuervo’ no implican ontología reista. Puede ser que el negro se halle tras o junto al cuervo. Es sólo cuando queremos decir que la región ocupada por el cuervo y la ocupada por el negro son exactamente la misma cuando nos vemos obligados a reificar: ‘ $(\exists x)(x \text{ es negro})$ y $(\exists x)(x \text{ es cuervo})$ ’ no equivale a: ‘ $(\exists x)(x \text{ es negro y es un cuervo})$ ’. Reificamos (en nuestro ejemplo, el cuervo) para poder

predicar (el atributo negro). A estos primeros balbuceos de la reificación se los denomina oraciones *categóricas observacionales*. Adviértase, en todo caso, que la reificación, para Quine, supone un primer y fundamental paso en nuestra teoría sobre el mundo. Así, Quine afirma que «la reificación de cuerpos a través del tiempo se halla más allá del alcance de las oraciones observacionales y las categóricas. La reificación sustancial es teórica.» (1992, p. 25).

Si nuestra ontología no fuera de cuerpos, tendríamos serias dificultades para predicar. Como veíamos, percatarse de la presencia de negro y la presencia de cuervo no equivale a percatarse de la presencia de un cuervo negro. Si utilizáramos siempre una ontología de fenómenos atmosféricos, no podríamos expresar el hecho de un cuervo negro, ya que ‘negrea y cuervea’ no implica que ambos fenómenos se solapen. Son las categóricas observacionales las que ofrecen los *nudos* necesarios.

En todo caso, no nos interesa aquí exponer las conjeturas quineanas acerca de la forma en que interiorizamos la ontología, sino subrayar que el propio Quine da por hecho que nuestras palabras expresan un compromiso ontológico concreto. Es más, como hemos visto, ha teorizado acerca de la manera exacta en que dicho compromiso es adquirido por los nuevos hablantes. ¿No contradice esto la tesis de la indeterminación? ¿No mantiene Quine, a la vez, que la referencia se halla indeterminada y teoriza acerca de cuál es, exactamente, nuestra ontología y los mecanismos mediante los cuales la interiorizamos? La respuesta es que las funciones vicarias permiten siempre el mismo

tipo de predicaciones – reificaciones – , independientemente de las entidades postuladas:

Esta conclusión [la de las funciones vicarias] apenas resulta sorprendente en el caso de las partículas elementales y de los objetos abstractos. Éstos, como se sabe, aparecen sólo por los papeles que tienen en nuestras teorías, no por la ostensión. La conclusión parece extraña en el caso de los cuerpos ordinarios, que fueron reificados perceptualmente. Pero nótese que las funciones vicarias suplantarían éstos sólo mediante interpretaciones con la capacidad de reificaciones plenamente sofisticadas [*with proxies only in the capacity of full sophisticated reifications*] a través del espacio y el tiempo. Las oraciones observacionales correspondientes permanecen unidas a los rangos originales de entradas neuronales. La oración ‘Gavagai’, o ‘Ahí va un conejo’, permanece asociada con la tan esponjosa intrusión en la escena, y las interpretaciones son irrelevantes [*and proxies to the winds*]. Estaría, incluso, injustificado desmontar la plena reificación, tanto en nuestra propia lengua como al traducir otra. La cuestión de las funciones vicarias es sólo que la interpretación ontológica es indiferente a la relación evidencial entre la observación y la teoría, e indiferente a la conformidad de un manual de traducción con la conducta.

(...) Una lección de las funciones vicarias es que nuestra ontología, como la misma gramática, es parte de nuestra contribución conceptual a nuestra teoría del mundo. El hombre propone; el mundo dispone, pero sólo por medio de veredictos holofrásticos de sí o no a oraciones observacionales que expresan las predicciones del hombre. (Quine 1995c, p. 351).

1.4. La propuesta de Davidson

La filosofía del lenguaje de Davidson parte de la pregunta: ¿qué aprendemos al aprender una lengua? O formulada de otra forma: ¿qué debemos saber para poder

entender – interpretar – todos los posibles enunciados de una lengua? La respuesta, como veremos, será una teoría de corte netamente extensionalista.

Si conseguimos especificar qué es exactamente lo que aprendemos al aprender una lengua, habremos especificado en buena medida en qué consiste saber el significado de los términos y expresiones de una lengua (ya que saber una lengua es saber el significado de sus expresiones, aunque sólo sea en el sentido de ser capaz de interpretar y producir correctamente cada una de las potencialmente infinitas expresiones), y, por tanto, qué clase de conocimiento es el conocimiento del significado de esos términos y expresiones. Más concisamente, habremos explicado qué es el significado de términos y expresiones y cómo conseguimos aprenderlos; luego la importancia filosófica de tal tarea parece difícil de exagerar.

En la *interpretación radical* hallamos, pues, un sujeto que trata de comprender las emisiones particulares de los demás hablantes sin traducirlas previamente a ninguna otra lengua. Lo que nosotros deseamos saber es qué clase de conocimiento le permitiría realizar esto, qué «teoría» le proporcionaría una comprensión de emisiones particulares que no se da de antemano. En la traducción radical buscamos un método de traducción del lenguaje que se va a interpretar al lenguaje del intérprete, pero en la interpretación radical no podemos suponerle al intérprete conocimiento de idioma alguno, y mucho menos del que se pretende interpretar. Ponernos en el lugar del intérprete radical ha de ayudarnos, pues, a entender el proceso de confección de una teoría del significado: «El análisis de la interpretación radical es el intento de hacer explícitos los criterios y

principios en los que confiamos implícitamente para justificar interpretaciones particulares y decidir si una conducta particular constituye conducta lingüística» (Stüber 1996, p. 336); y otra explicación en la misma línea:

Situarse el punto de partida del análisis de la interpretación en esta situación extrema es un artificio metodológico (cuyas virtudes, sin embargo, no me parecen del todo claras) destinado a poner de manifiesto los aspectos implicados en la comunicación normal entre los seres humanos. La ventaja de este punto de partida consiste en que nos permite evitar que nos pasen inadvertidos presupuestos importantes de la comunicación, cosa que puede fácilmente suceder si analizamos la comunicación en el caso de sujetos que comparten una lengua y una cultura. (C. Moya 1992, p. 29).

La conclusión hasta aquí, pues, es que cualquier teoría cuyo conocimiento nos pueda capacitar para interpretar todos y cada uno de los enunciados de una lengua cumplirá dos condiciones: a) partiendo de un conjunto finito de axiomas, b) produce de forma recursiva una interpretación para cada enunciado.

Respecto al punto primero, conviene recordar el principio de composicionalidad, según el cual el significado de una oración depende del significado de sus partes componentes (aproximadamente, palabras). Luego el conjunto finito de reglas nos ofrecería el conocimiento acerca de cómo los significados de las palabras se combinan para formar los más complejos significados de las oraciones. Tendríamos pues un conjunto finito de axiomas, compuesto por el significado de las palabras – en el sentido que a continuación

se especifica – y las reglas antes mencionadas, más los infinitos teoremas que de aquí se pueden derivar y que más abajo caracterizamos. En palabras del propio Davidson:

Si las oraciones dependen de su estructura para su significado, y nosotros entendemos el significado de cada elemento de la estructura sólo como una abstracción a partir de la totalidad de oraciones en las que aparece, entonces podemos dar el significado de cualquier oración (o palabra) sólo dando el significado de cada oración (o palabra) en el lenguaje. Frege dijo que sólo en el contexto de una oración tiene significado una palabra; en la misma vena podría haber añadido que solo en el contexto del lenguaje tiene significado una oración (y por tanto una palabra). (Davidson 1967, p. 22).

De esta forma, la teoría puede proporcionar una interpretación para toda oración del lenguaje, ya que a partir de ese *stock* finito de significados (entendidos de la forma antes mencionada) y de las reglas indicadas sería capaz de ofrecer esquemas del tipo:

(V) *o* significa que *p*

donde *o* ha de sustituirse por un nombre metalingüístico de la lengua que se pretende interpretar y *p* por una oración que «dé el significado» – a continuación aclararemos de qué manera – de *o*. Dicho de otro modo, *o* ocupa el lugar de un nombre de una oración (o una oración mencionada) y *p* de una oración usada. (Una forma de mencionar un enunciado es entrecomillándolo, otra describiéndolo estructuralmente). La oración que se usa no hace otra cosa que describir las condiciones bajo las cuales *o* resultaría verdadero, ya que dar las condiciones bajo las que una oración resulta verdadera es,

cuando menos, *una manera de* dar el significado.¹⁵ Preguntémosnos, si no, si sería plausible decir con sentido que alguien sabe cuándo una oración es verdadera y cuándo falsa, pero que, sin embargo, no sabe lo que la oración en cuestión significa. Luego podemos formular el esquema anterior de modo que prescindamos del controvertido «significa que»:

(V) *o* es V si y sólo si *p*.

Como se observa, estamos ante la Convención V de Tarski. Vale la pena volver a citar a Davidson aquí:

No hay necesidad de suprimir, está claro, la conexión obvia entre una definición de la verdad del tipo que Tarski ha mostrado cómo construir, y el concepto de significado. Es ésta: la definición funciona dando condiciones necesarias y suficientes para la verdad de cualquier oración, y dar condiciones de verdad es una manera de dar el significado de una oración. Saber el concepto semántico de verdad para un lenguaje es saber qué es para una oración – cualquier oración – ser verdadera, y esto viene a ser [*amounts to*], en el buen sentido de la expresión, entender el lenguaje. Esto, de todas maneras, es mi excusa para un rasgo de la presente discusión que puede asustar a la vieja guardia [*shock old hands*]; mi libre uso de la palabra «significado», ya que lo que yo llamo una teoría del significado resulta que, después de todo, no hace uso de significados, ni de oraciones ni de palabras. De hecho, ya que una definición tarskiana de la verdad ofrece todo lo que hemos pedido de una teoría del significado, está claro que una teoría semejante cae

¹⁵ Resulta llamativo que Davidson utilice esta expresión para referirse a la relación entre dar el significado y dar las condiciones de verdad. El párrafo concreto será citado más adelante en toda su extensión.

cómodamente bajo lo que Quine denomina una «teoría de la referencia» como diferenciada de lo que llama una «teoría del significado». (Davidson 1967, p. 24).

La teoría que buscábamos puede quedar más o menos así caracterizada: descubrimos el «significado» de las palabras basándonos en lo que éstas aportan a las condiciones veritativas de *todas* las oraciones en los que podrían aparecer y una vez asimilados estos significados, somos capaces de precisar las condiciones de verdad de cualquier oración (o sea, somos capaces de producir una oración V, donde la oración se halla mencionada a la izquierda del «es V si y sólo si» y a la derecha se halla una oración que describe las situaciones en las que la oración anterior resultaría verdadera), por novedosa que la oración en cuestión pueda resultar para nosotros.

Una teoría de la verdad semejante habría de cumplir algunas condiciones, a saber:

- 1) Producir para toda oración su correspondiente oración V, donde la verdad de la oración usada *implicaría* la verdad de la oración mencionada.
- 2) Salir airosa del test empírico. Es decir, la teoría ha de ofrecernos una interpretación *racional* de la conducta verbal de los hablantes.
- 3) *Ser aprendible*. Es obvio que no podemos describir la forma en la que se aprende el lenguaje ofreciendo la teoría que debería dominar todo hablante competente y que dicha teoría resultara, en la práctica, imposible de aprender; y

esto sucedería, por ejemplo, si se defendiera que el número de axiomas o de reglas es infinito. La teoría de Davidson resulta aprendible en este sentido, ya que los axiomas son finitos, pero no lo resultaría, por ejemplo, la propuesta por Frege. Esto sucede por la concepción fregeana de que a cada expresión ha de corresponderle una referencia y un sentido, siendo éste algo así como el contenido asociado a la expresión, «la forma de llegar» a la referencia, y que en los contextos oblicuos la referencia pasa a ser el que era el sentido habitual. Esta teoría incumple el requisito de aprendibilidad porque los contextos oblicuos se pueden iterar. Así, la expresión ‘la Tierra se mueve’ tiene un sentido «normal» en la oración ‘La Tierra se mueve’, otro diferente (su sentido sería ahora algo así como el sentido de un sentido, ya que la oración ha pasado a tener como referencia el que era su sentido habitual) en ‘Galileo creía que la Tierra se mueve’, otro nuevo sentido en ‘el alumno sabe que Galileo creía que la Tierra se mueve’ y así indefinidamente, con lo que para entender la expresión ‘la Tierra se mueve’ habríamos tenido que aprender una serie infinita de sentidos.

El proceso de diseño de una teoría de la verdad – o sea, del significado – para un lenguaje natural transcurriría más o menos así: En primer lugar asimilamos la estructura lógica del nuevo lenguaje, normalmente éste viene a consistir en la teoría de la cuatificación de primer orden más identidad. La evidencia consiste aquí en clases de oraciones mantenidas siempre como verdaderas o como falsas (verdades lógicas potenciales). Una vez identificados los predicados, los términos singulares, los cuantificadores, las conectivas y la identidad, nos concentramos en las expresiones

indéxicas: aquellas oraciones mantenidas algunas veces como verdaderas y algunas veces como falsas de acuerdo con cambios observables en el mundo. El último paso se ocupa de las oraciones respecto de las cuales no hay acuerdo uniforme, pues el valor de verdad mantenido no depende sistemáticamente de cambios en el entorno.

Tras haber visto a Quine, esta reconstrucción resultará tan familiar que parecerá calcada. Como el propio Davidson apunta en una nota al pie tras detallar el proceso descrito, la explicación es paralela a la que aparece en el capítulo segundo de *Palabra y Objeto*, sólo que prescindiendo de la noción de significado estimulativo – al menos tal y como la concebía Quine – a favor de la referencia a los rasgos objetivos del mundo, que se alteran en conjunción con cambios de actitud hacia la verdad de las oraciones. Por lo demás, los tres pasos arriba descritos equivaldrían, en la terminología quineana, a las oraciones eternas, observacionales y ocasionales, respectivamente.

Una vez que el sujeto correspondiente ha culminado la tarea de diseñar una teoría de la verdad, siguiendo los pasos reseñados, será capaz de producir un esquema V para toda oración significativa de la lengua en cuestión. Así, por ejemplo:

«La nieve es blanca» es V si y sólo si la nieve es blanca.

El esquema, así formulado, produce sensación de vacuidad, pues la oración entrecomillada es repetida tal cual tras el bicondicional. Podemos esquivar esta molesta

sensación ofreciendo una descripción estructural en lugar del entrecomillado, como, por ejemplo, recurriendo al deletreo:

«ele, a, espacio, ene, i, e, uve, e, espacio...» es V si y sólo si la nieve es blanca.

En todo caso, el esquema ofrecido no aparentará la vacuidad anterior en las oraciones que contienen elementos indécicos, pues a la hora de describir sus condiciones de verdad nos vemos obligados a introducir contextualizaciones espacio-temporales. Así pues, la teoría entrañará oraciones como las siguientes:

«Estoy cansado» es verdadera en tanto que (potencialmente) proferida por h en t si y sólo si h está cansado en t .

«Este libro fue robado» es verdadera en tanto que (potencialmente) proferida por h en t si y sólo si el libro señalado por h en t es robado anteriormente a t .

Una vez que hemos caracterizado someramente lo que ha de constituir una teoría de la verdad y, por ende, del significado, nos interesa preguntar ahora: según la teoría davidsoniana, ¿se halla el significado indeterminado?

1.5. La indeterminación en Davidson

Davidson acepta las teorías quineanas de la indeterminación del significado y la inescrutabilidad de la referencia. De hecho, podemos decir que las acepta tal cual, es decir, tal y como Quine las expone y por los argumentos que ofrece:

...Acepto la tesis de Quine de la inescrutabilidad de la referencia y, por tanto, de la indeterminación de la traducción. Y creo que acepto ambas principalmente por los argumentos que he aprendido de Quine. (Davidson 1979, p. 227).

Y a la inversa: Quine defiende que el programa de Davidson de equiparar la clarificación del significado con la definición de verdad le confiere una gran claridad a su vieja acusación de que las intenciones no son claras:

Si seguimos a Davidson al igualar la clarificación del significado con la definición de la verdad, entonces nuestra vieja crítica de que las intenciones no son claras gana una cierta objetividad. Y también lo hacen nuestros escrúpulos contra las entidades mentales. (Quine 1969c, p. 334).

En todo caso, antes de exponer las consideraciones de Davidson sobre la indeterminación, veamos cómo ésta constituye también una conclusión de su teoría del significado. Imaginemos alguien que en su proceso de adquisición del lenguaje ha procedido de la manera supuesta por Davidson, produciendo una oración V para cada enunciado a interpretar y siempre basándose en lo que los componentes de los

enunciados añaden a las condiciones de verdad de todas las oraciones donde aparecen. Llamemos L al lenguaje que este sujeto intenta interpretar y θ a la teoría de la verdad por él pergeñada. Supongamos ahora que L posee un predicado, digamos P , que significa ‘ser parte de’. Pues bien, la cláusula crucial para P es la que afirma que para cualesquiera objetos n y m , una asignación satisface Px_nx_m si y sólo n es una parte de m . En resumidas cuentas, lo que la cláusula nos dice es que el predicado P se aplica a todo par de objetos de los que el primero sea parte del segundo. Pero imaginemos un predicado P' con la misma extensión que P pero con distinto significado. Tomemos para ello cualquier oración contingente o necesariamente verdadera en castellano, por ejemplo, ‘la Tierra se mueve’ y traduzcamos P' como ‘ n es parte de m y la Tierra se mueve’. Tenemos ahora dos predicados coextensivos, pues no puede haber un par de objetos al que se aplique P y no P' , cuyo significado, digamos, *intuitivo*, difiere.

Podemos proceder exactamente de la misma manera con los cuantificadores y las conectivas. Introduzcamos una conectiva ‘ \wedge^+ ’ que se traduce realmente como ‘y la tierra se mueve y...’; y un cuantificador ‘ \forall^+ ’ que se traduce como ‘para todo objeto..., la tierra se mueve y...’. Cualquier evidencia que haga verdadero un enunciado como ‘ $P \wedge Q$ ’, hará igualmente verdadero uno como ‘ $P \wedge^+ Q$ ’. En general, cualquier evidencia hará verdadera una oración O' que resulte de sustituir en O a P , \wedge o \forall por P' , \wedge^+ o \forall^+ respectivamente. Vemos así cómo la teoría davidsoniana conduce también a la indeterminación: toda la evidencia posible no determina si los hablantes utilizan el predicado P o P' , la conectiva \wedge ó \wedge^+ .

En la exposición davidsoniana se aprecia perfectamente, pues, cómo la indeterminación surge debido a que el elemento relevante para la interpretación o construcción de una teoría del lenguaje es el valor de verdad de las oraciones consideradas holofrásticamente, no el supuesto significado o referencia de los elementos suboracionales. Como dice el propio Quine: «Como Davidson observa, la definición de la verdad de Tarski hace verdaderas a las mismas oraciones bajo cualquier transformación uno a uno de los valores de las variables» (1995c, p. 351).

Davidson advierte de que no hay forma de hallar una determinación para la referencia, ni tan siquiera relativizándola a un marco ontológico, ya que – como vimos con Quine – esa relativización estaría hecha en esa misma lengua y podría ser interpretada, de nuevo, de diferentes maneras. La única forma que hay de relativizarla es aceptando que una misma expresión puede pertenecer a varias lenguas, cada una con diferentes compromisos ontológicos. Al fin y al cabo, lo que hacemos diariamente es traducir de idiolecto a idiolecto.

La evidencia empírica seguirá sin determinar la referencia, pero una vez elegido un esquema de interpretación ya queda determinada la respuesta a las preguntas acerca de la traducción de las emisiones de otro hablante. Eso sí, a la hora de cambiar de marco de referencia imputado a un hablante hemos de tener en cuenta toda la evidencia, lo que incluye entidades intencionales:

Lo que nos permite escoger entre varios lenguajes para un hablante es el hecho de que la evidencia – las actitudes o acciones dirigidas a oraciones o emisiones – se refiere no sólo a la interpretación de la lengua sino también a la atribución de creencia, deseos e intenciones (y sin ninguna duda a otras actitudes). (Davidson 1979, p. 239).

Ahora bien, Davidson defiende que la indeterminación no llega nunca a extremos tan dramáticos como propugna Quine. La causa de esto es la limitada importancia que Quine le otorga al principio de caridad (éste es utilizado por su traductor radical para traducir las conectivas veritativo-funcionales y no imputar, pues, una lógica demasiado «exótica» al nativo) y la fundamental que le otorga Davidson. Como él mismo dice y como ya hemos subrayado arriba, dicho principio se aplica en la interpretación radical «de arriba abajo» (*across the board*. Davidson 1973, p. 136, nota al pie).

Como dice M. Hernández, esta visión del asunto conduce a un «quineanismo mitigado», con actitudes proposicionales irreductibles, pero sin más base empírica aceptada que lo intersubjetivamente observable (M. Hernández 1990, p. 108).

Así es como la indeterminación resulta insoslayable según la manera davidsoniana de entender el aprendizaje y funcionamiento del lenguaje. Davidson, además, ha hecho hincapié en un argumento para la indeterminación: resulta inútil postular significados. Las razones, ahora, no tienen que ver con el antimentalismo, como en Quine, sino que resultan mucho más pragmáticas:

Mi objeción a los significados en la teoría del significado no es que son abstractos o que sus condiciones de identidad son oscuras, sino que no tienen una utilidad demostrada.

(Davidson 1967, p. 21).

Conviene, pues, subrayar aquí que la acusación de inutilidad que Davidson dirige contra los viejos significados no va acompañada, como en Quine, de conductismo. Quine rechaza toda apelación a términos mentalistas como intención y creencia, Davidson no:

Mi forma de intentar dar una explicación del lenguaje y el significado hace uso esencial de conceptos tales como los de creencia e intención, y no creo que sea posible reducir estas nociones a algo más científico o conductual. [...] No arruinaré ninguno de mis planes si para decir lo que sabe un intérprete es necesario utilizar una así llamada noción intensional – una que va ligada [*consorts*] a creencia e intención y similares.

(Davidson 1976, pp. 175-6)

Así pues, Davidson rechazará la apelación a entidades como los significados porque no han probado ser de utilidad alguna en la elaboración de una teoría del significado. Ésta pues, no tiene más remedio que diluirse: «No necesitamos el concepto de referencia; tampoco necesitamos la propia referencia, sea ésta lo que sea.» (Davidson 1977, p. 224). Así que... la referencia desaparece [*drops out*]. No tiene ningún papel esencial en la explicación de la relación entre lenguaje y realidad» (Davidson 1977, p. 225).

La idea de hacer de los significados entidades autónomas – es decir, la respuesta platónica – , puede ser vista como una manera de solucionar los problemas que presenta una respuesta introspectiva. Ésta última buscaría los significados en la conciencia. Así, el significado vendría a consistir en la imagen o representación mental que acompañe la emisión. A diferencia de las palabras, la imagen mental se parecería a aquello de lo que es imagen. El problema con esta respuesta es el siguiente:

Basar el significado en la relación de semejanza explica demasiado, pues cualquier cosa se parece a otra en algún sentido. Por otra parte, si la relación entre los sonidos «la hierba es verde» y el color de la hierba constituía un problema, no lo es menos la relación entre dichos sonidos y la imagen mental en cuestión. En ambos casos se trata de dos cosas distintas y separadas. En realidad, el problema se ha complicado: si lo que da significado a los sonidos «la hierba es verde» es una imagen mental, ¿cómo puedo saber que esos sonidos significan lo mismo en tu boca que en la mía? La comunicación se convierte en un azar incognoscible. (C. Moya 1992, pp. 28-29).

Estas dificultades, pues, pueden llevar a una concepción platónica del significado, según la cual éste consiste en una entidad abstracta con existencia independiente de las mentes de los hablantes, pero captada por estas mentes. El rechazo davidsoniano al platonismo semántico se explica así:

Junto al habla, los hablantes y el mundo, la ontología se ve enriquecida con entidades ideales: significados, proposiciones, intenciones, sentidos. La escasa simpatía de Davidson hacia esta actitud no se debe tan sólo a una tendencia hacia la sobriedad ontológica, sino también, y sobre todo, a su pobre valoración de las virtudes explicativas

de las entidades postuladas. Al postular estas entidades no se ha hecho sino deificar el problema con el que tropezábamos, no resolverlo. (C. Moya 1992, p. 28).

En Davidson, nociones como las de satisfacción, referencia o significado han de ser entendidas como postulados (*posits*) que la teoría utiliza sin precisar una corroboración empírica independiente.

Sugiero que palabras, significados de palabras, referencia y satisfacción son postulados [*posits*] que necesitamos para implementar una teoría de la verdad. Sirven para este propósito sin necesitar confirmación independiente o base empírica. (Davidson 1977, p. 222).

La idea que Davidson está exponiendo aquí es, pues, que la referencia es una noción teórica, completamente interna a su análisis filosófico de la interpretación (exponiéndolo con las palabras de Hookway 1988, p. 180). Dice Davidson hablando de conceptos como nombre, predicado, oración, referencia o significado:

Todos estos son conceptos teóricos, y los objetos [*items*] a los que se les aplica son objetos abstractos. No los necesitamos para utilizar o aprender una lengua; obviamente, no están a nuestra disposición cuando estamos aprendiendo una primera lengua. Cuando queremos estos conceptos es al hablar de la conducta verbal. Los filósofos, psicólogos y lingüistas necesitan estos términos teóricos si desean describir, teorizar sobre y explicar actividades verbales. (Davidson 1992, p. 108).

Podemos leer una intuición de esta idea ya en Quine. Éste se muestra un instrumentalista radical en epistemología: nuestras teorías científicas no son más que instrumentos que nos permiten predecir nuestra estimulación sensorial y este instrumento está constituido en parte por nuestros mecanismos lingüísticos de referencia; la misma referencia puede ser vista como cumpliendo esta única función. (V. F. Mühlhölzer 1995).

Habremos de volver sobre esto, ya que la idea de que significado o referencia son conceptos altamente teóricos, sin correlato observacional claro, ha servido a muchos autores para diferenciar tajantemente entre dos niveles lingüísticos: un nivel inmanente, en el que los hablantes hacen afirmaciones acerca de la referencia de sus términos, y uno trascendente, en el que los filósofos del lenguaje exponen, por ejemplo, la tesis de la indeterminación. Con esta dualidad de niveles se intenta conjugar la indeterminación con la posibilidad de enunciar la tesis.

1.6. Indeterminación y lenguas naturales

Davidson ha subrayado los problemas que surgen a la hora de definir con precisión un lenguaje natural (el español o el alemán, por ejemplo), ya que dicha definición habrá de basarse en la de nociones como la de referencia o significado, que, como hemos observado, ofrecen insalvables dificultades. Vemos que esta dificultad conduce a otra: la de determinar qué lenguaje está hablando un sujeto.

Para hablar un lenguaje, las emisiones de uno deben ser consistentes con la definición de algún lenguaje. El problema es que las emisiones son finitas en número, mientras que la definición de un lenguaje asigna significados a un número infinito de oraciones. Habrá, por tanto, infinitos lenguajes diferentes que están de acuerdo con todas las emisiones del sujeto, pero que difieren con respecto a oraciones no emitidas. ¿Qué hace que el hablante sea un hablante de uno de esos lenguajes y no de otro? (Davidson 1992a, p. 110).

Tenemos aquí, en realidad, el viejo problema humeano de la inducción aplicado ahora no a determinar qué reglas rigen la naturaleza, sino qué lengua habla un sujeto particular. Ya que no podemos observar todas las posibles emisiones, nunca podemos estar seguros de que el hablante se comportará en el futuro como nuestra interpretación predice que lo hará. Aquí no servirá de ayuda apelar a hechos como que el hablante se ha comportado siempre como el resto de la comunidad lingüística, que aprendió la lengua en el mismo lugar que todos los demás o que pertenece a la misma cultura o comunidad, ya que el problema es cómo se comportará el hablante en el futuro, no cómo lo hacía en el pasado. Así, si nosotros como intérpretes no sabemos cómo se va a comportar en el futuro el hablante que ha de ser interpretado, ya que sólo disponemos de los datos acerca de su conducta verbal pasada y presente, no podemos adscribirle sin error posible una lengua.

Pero el problema resulta aún más grave: aun suponiendo que pudiéramos observar todas sus emisiones en todas las circunstancias posibles, esto no determinaría qué lengua está hablando. Citando de nuevo a Davidson:

Y el problema puede ser peor aún. Ya que incluso si un hablante emitiera (cosa imposible) todas las oraciones de algún lenguaje, muchos otros lenguajes serían consistentes con toda su conducta y sus estados internos, como Quine ha mantenido; y yo estoy de acuerdo. (Davidson 1992a, p. 110).

Recordemos que una fuente de indeterminación es la inextricabilidad entre creencia y significado:¹⁶ podemos variar uno u otro elemento y obtener así teorías interpretativas válidas pero incompatibles:

Cuando se tome en consideración toda la evidencia, quedarán, como Quine ha subrayado, los arreglos [*trade-offs*] entre las creencias que atribuimos a un hablante y las interpretaciones que damos a sus palabras. Pero la indeterminación resultante no puede ser tan grande que cualquier teoría que pase los tests nos ofrezca interpretaciones. (Davidson 1973, p. 139).

Así, dos teorías interpretativas empíricamente válidas podrían contener teoremas contradictorios, a condición de que los reajustes pertinentes fueran realizados en otras

¹⁶ Una fuente, pero no la única. (Remitimos a la diferenciación entre el argumento desde arriba y el argumento desde abajo, v. 1.2). Ramberg (1989, p. 94) propone un ejemplo para mostrar cómo aún eliminando esta fuente la indeterminación sería inevitable. Supongamos dos individuos perfectamente consistentes en sus teorías y omniscientes, con la única salvedad de que no conocen la lengua del otro. En su tarea de interpretación no surgirían dudas acerca de las creencias a adscribir, éste sería un factor “dado”, quedando únicamente por indagar el significado. Ahora bien, la indeterminación surge por la posibilidad de partir de axiomas diferentes para generar las oraciones V: tanto el aparato lógico del lenguaje – cuantificadores y conectivas – como el referencial – términos – son susceptibles de recibir interpretaciones incompatibles entre sí pero que ofrecen una racionalización satisfactoria de la conducta del interpretado. Por tanto: «Nada que un ser consistente y omnisciente pueda decir o hacer dará al otro razones para escoger una permutación y no la otra» (Ramberg 1989, p. 95). Con todo, se podría replicar que el ejemplo parte del supuesto de que es posible hacer una diferenciación estricta entre creencias y significados, justo lo que la tesis de la inextricabilidad niega.

partes de la teoría. Aquí se podría pensar que Davidson está permitiendo una indeterminación tan grande que acaba cayendo en contradicciones:

[Indeterminación de este tipo implica que] no habrá nada que escoger entre una oración V ‘*s* es verdad si y sólo si *p*’, ofrecida en un sistema, y ‘*s* es verdad si y sólo si *q*’ ofrecida en otro sistema. Sin embargo, teníamos los requerimientos iniciales de que las oraciones V fueran verdaderas. Ya que *p* y *q* pueden ser contrarias, *ambas* oraciones V no pueden ser verdaderas. Desde el punto de vista de Davidson, esto es una *reductio ad absurdum* de la indeterminación. (Hacking 1975, p. 154. Citado de Ramberg 1989, pp. 89-90).

La respuesta de Davidson ahora es que: «Nuestro error era suponer que hay un lenguaje único al que pertenece una emisión determinada» (Davidson 1984a, p. 239). Efectivamente, ya hemos visto que la tesis de la inescrutabilidad de la referencia muestra cómo toda la evidencia posible no puede determinar la referencia de los términos de un hablante; y si consideramos que dos hablantes cuyos términos no tienen la misma referencia no hablan la misma lengua, la tesis vendría ahora a ser algo así como la inescrutabilidad de la lengua. Davidson no afirma que las distinciones dentro de una lengua estén indeterminadas, sino que está indeterminado qué lenguaje – que se identifica por medio de un esquema de referencia – estamos hablando.¹⁷

¹⁷ Se vuelve sobre la crítica de Davidson a la existencia de convenciones lingüísticas rígidas en 3.3.3. Añadamos que algunos autores han considerado que esta idea lleva la indeterminación demasiado lejos:

Los gobiernos opresivos, como los de Franco y Mussolini, intentan suprimir las lenguas minoritarias; bajo esos regímenes los maestros castigan a los niños por hablar esas lenguas en el recreo. En la india, las multitudes se manifiestan contra las propuestas de hacer el Hindi la única lengua oficial. Los bretones, catalanes, vascos y kurdos afirman que su lengua es el alma de su cultura. No parece estarnos abierta la opción de

En la práctica, la adscripción de una lengua resulta sumamente sencilla. Habitualmente utilizamos criterios como la parte del planeta donde vive el hablante. A eso podemos sumar las personas con las que habla o incluso factores como su forma de vestir o sus costumbres. Pero la idea de Davidson ahora es negar que la comunicación precise la determinación unívoca de la lengua – es más, resulta imposible realizar semejante determinación – . Como hemos visto, aun siendo posible adscribir diferentes significados a las palabras de un hablante, la comunicación resulta fluida; y, asimismo, no necesitamos la certeza de que un hablante seguirá hablando de la manera que nuestra interpretación predice para poder entendernos. El lenguaje se basa en la capacidad de al menos una segunda persona para interpretarlo. De hecho, no es inhabitual el caso de personas que entienden una lengua sin ser capaces de hablarla. ¿Por qué, entonces, hemos de exigir que otro hablante hable la misma lengua (siga las mismas regularidades) que nosotros para entenderlo?

Si usted y yo fuéramos los únicos hablantes en el mundo, y usted hablara Sherpa mientras que yo hablo inglés, podríamos entendernos, aunque cada uno siguiera diferentes ‘reglas’ (regularidades). Lo importante, por supuesto, es que cada uno debería proveer al otro de algo que fuera inteligible como lengua. Ésta es una intención que los hablantes han de tener; pero cumplir esta intención, mientras que puede requerir un cierto grado de lo que el otro percibe como consistencia, no conlleva seguir reglas compartidas o convenciones. Puede incluso darse el caso de que debido a diferencias en

afirmar que esos gobiernos y sus gentes se hallan bajo la ilusión de que hay algo que están suprimiendo o defendiendo. (M. Dummett 1986, p. 465).

nuestras cuerdas vocales no podríamos hacer los mismos sonidos y, por tanto, hablar la misma lengua. (Davidson 1992a, 114).

Quienes hablan alegremente de lenguas se comprometen, según la visión davidsoniana, con una reificación de los significados que las reflexiones anteriores acerca de la indeterminación muestran inaceptable. Cuando nos comunicamos con alguien, independientemente de qué lengua consideramos que habla, pergeñamos una teoría interpretativa para dicho hablante y para cada momento. Con toda seguridad, durante la conversación surgirán emisiones que nos obligarán a retocar nuestra teoría precedente. Una concepción de la comunicación basada en teorías interpretativas no halla inconvenientes para acomodar los cambios que, como vemos, se han de realizar a cada momento, mientras que hablar de lenguas nos ofrece una imagen estática de la comunicación, pues una lengua es un conjunto de convenciones ya cerrado.

Por tanto, Davidson no considera que la apelación a reglas, como convencionalmente se ha pensado, suponga una explicación válida de la conducta lingüística. Recordemos que éste fue uno de los problemas favoritos de Wittgenstein, tratado en profundidad por Kripke. Wittgenstein observó que la apelación a reglas resulta problemática, ya que ¿cómo saber qué regla está siguiendo un sujeto si sólo hemos observado unos pocos casos de su aplicación de la regla? Ya que él aprendió la regla a partir, igualmente, de unos pocos casos, no podemos estar seguros de que no haya malentendido la regla y en un momento dado la utilice de forma diferente a como nosotros lo haríamos. Pero si ambas reglas, la que yo intentaba enseñarle y la que él acabó interiorizando, ofrecen los

mismos resultados en los casos en los que la aplicamos juntos, cuando aparezca una discrepancia, no estará claro quién de los dos se halla equivocado. Es más, es la misma noción de hallarse equivocado aquí la que pierde sentido, ya que no hay materia objetiva sobre la que hallarse equivocado o no. Así pues, no parece que apelar a reglas nos libere de la indeterminación.

Resultará interesante analizar brevemente la opinión que Kripke le atribuye a Wittgenstein acerca de las reglas lingüísticas (como dice Davidson, no sin cierta ironía, no está claro que Wittgenstein defendiera esta visión, y Kripke no la defiende claramente, así que puede que no sea la opinión de nadie). Según la idea tradicional, hablar un lenguaje es seguir unas reglas. Las reglas especifican cómo se debe actuar en los casos futuros, es decir, una vez visto cómo se aplica la regla, no hay más que «seguir así». Con todo, no hay ningún estado mental interno o «captación» del contenido de la regla o de proposiciones, así que ningún estudio de la cabeza – o mente – del hablante revelará si éste está siguiendo la regla prevista o alguna alternativa (si, por ejemplo, está hablando de conejos o de instancias de la conejeidad).

Los intérpretes simplemente juzgan que un hablante está siguiendo la misma regla que ellos (sus intérpretes) si el hablante continúa como ellos lo harían. Dicho en términos de significado: juzgamos que un hablante quiere decir lo que nosotros si nosotros emitiéramos las mismas palabras que él, si él continúa como nosotros lo haríamos. (Davidson 1993a, p. 113).

Wittgenstein trata aquí las supuestas reglas lingüísticas de la misma manera que las reglas de la aritmética, como sumar. Davidson cree que esto es un error. La diferencia fundamental es que en aritmética disponemos de una descripción exacta del procedimiento, aprendemos y describimos el procedimiento. Éste es formulado de manera explícita. En el aprendizaje de la lengua no hay nada similar. Y volvemos a donde empezamos: la noción de «regla lingüística» es como la de significado, oración o sinonimia, una noción teórica utilizada para describir el lenguaje o para aprender un segundo lenguaje, pero no un primero. No se nos enseña a utilizar palabras o a formar oraciones dándonos un conjunto de reglas, como sí se hace en matemáticas.

¿Cuál es, pues, la alternativa davidsoniana a la apelación a lenguas naturales, sistemas de reglas o convenciones? Teorías interpretativas continuamente cambiantes. Davidson introduce una distinción entre teorías «de entrada» (*prior*) y «de paso» (*passing*). La teoría de entrada del intérprete «expresa cómo se halla preparado de antemano para interpretar una emisión del hablante» (Davidson 1986, p. 442), mientras que la teoría de entrada del hablante «es lo que él *cre*e que es la teoría de entrada del intérprete» (Davidson *ibid.*, p. 442). Pero no es necesario que ambas coincidan para lograr la comunicación. De hecho, Davidson subraya que en nuestras conversaciones ordinarias nos vemos obligados a utilizar una interpretación no homofónica más a menudo de lo que pueda parecer. Así, la indeterminación nos hace ver que podemos adscribir a un hablante diversos idiolectos y, sin embargo, la comunicación sería posible. Podemos pensar que este hecho evidencia dos factores relevantes para cuestiones semánticas: uno, que no necesitamos las lenguas naturales ni la reificación de significados o

convenciones que comportan y dos, la indeterminación resulta tan inocua que no impide la comunicación fluida.

Lo único que se requiere para que un intérprete se comunique con un hablante es que sus teorías de paso coincidan. Con todo, Davidson señala:

La teoría de paso no puede corresponder, en general, a la competencia lingüística de un intérprete. No sólo posee su lista cambiante de nombres propios y vocabulario peculiar, sino que incluye cada uso satisfactorio – esto es, correctamente interpretado – de cualquier palabra o frase, no importa cuán lejos de lo ordinario. Cada desviación del uso ordinario, mientras sea entendida en el momento (a sabiendas de su desviación o no, por una de las partes o por ambas), es en la teoría de paso como un rasgo de lo que las palabras significan en esa ocasión. Tales significados, aun cuando sean transitorios, son literales... (Davidson 1986, p. 442).

1.7. Holismo y caridad

Los nombres de Quine y holismo van indefectiblemente de la mano. Gibson llega a decir que aunque Quine sea el empirista *par excellence*, su diferencia con otros empiristas reside en que Quine es un empirista ilustrado, ¿ilustrado con qué?... con el holismo.

El rechazo quineano de nociones como las de analiticidad y significado se apoya, principalmente, en su defensa del holismo. Éste nace en Quine como una reacción a la tesis verificacionista del significado, defendida por algunos miembros del Círculo de

Viena – y por el «Wittgenstein intermedio» – , que postula que el significado de una oración es el método de confirmarla empíricamente (Quine 1951, p. 37). En contra de esta idea, la tesis holista defiende que las oraciones no pueden ser confirmadas individualmente (y, por tanto, que ninguna oración posee por sí sola un significado), y que «nuestras oraciones sobre el mundo externo se enfrentan al tribunal de la experiencia sensorial no individual sino sólo corporativamente [*corporate body*]» (Quine 1951, p. 41). Aquí – en nota al pie – Quine reconoce que esta doctrina no es realmente suya, ya fue formulada por Pierre Duhem en su obra *La Théorie physique: son objet et sa structure* (Duhem 1960, pp. 303-328) – de ahí que la tesis haya sido bautizada como tesis de Duhem-Quine – , y en *The methodology of Pierre Duhem*, de Armand Lowinger (Lowinger 1941, pp. 132-40).

El lenguaje se compone, según la visión holista de Quine, de tres tipos de oraciones: oraciones observacionales, para cuyo a/disentimiento no se requiere más que la estimulación presente; oraciones ocasionales, aquellas para las que además de la estimulación necesitamos cierta información colateral (como ‘Soltero’); y las oraciones eternas o permanentes, que una vez aprendidas suelen mantener su valor de verdad invariable independientemente de la estimulación (por ejemplo ‘la Tierra gira’). Por otro lado, podemos comparar la ciencia a un campo de fuerza cuya periferia estaría constituida por las oraciones más observacionales, que estamos dispuestos a abandonar en caso de que nuestra teoría ofrezca algún error, y cuyo núcleo estaría formado por las oraciones de más alto contenido teórico, especialmente las de la lógica y matemática, a las que si calificamos de analíticas es por la razón de que, en caso de fallar nuestra

teoría sobre el mundo, no estamos dispuestos a abandonarlas, sino que siempre preferimos variar la parte más periférica. (Con todo, es inútil buscar una frontera exacta entre los tres tipos de oraciones; conviene concebirlas insertas en un todo graduado de cero a cien. Así, oraciones como ‘Rojo’ estarían cercanas al cero – a la observacionalidad absoluta – y oraciones como ‘la Tierra gira’ estarían cercanas al cien – a la teoriedad absoluta –).

Podemos diferenciar entre un holismo epistemológico y un holismo semántico. Si el primero, que sería, propiamente, la llamada tesis de Duhem-Quine, se puede resumir diciendo que las teorías científicas se enfrentan a la corroboración empírica en bloque y no oración por oración, la del holismo semántico explica esta circunstancia por el hecho de que las oraciones no tienen un significado perfectamente aislable y determinado, pues si lo tuvieran no debería haber problema para verificar una oración de manera independiente. Según las oraciones se alejan de la observacionalidad, pierden anclaje directo con la realidad y aumentan su dependencia con otras oraciones de la lengua. Como vemos, la tesis de Duhem viene a decir que la verdad no es una característica de las oraciones que constituyen una teoría, sino de la teoría en su conjunto o de una parte suficientemente amplia de la misma; el holismo semántico dice lo mismo del significado.¹⁸

¹⁸ Veamos con un sencillo ejemplo que ilustra cómo funciona el holismo epistemológico en la práctica (tomado de O. Müller 1997). Tomemos las oraciones

- (1) En esta cámara oscura hay un electrón en movimiento.
- (2) Los electrones en movimiento dentro de cámaras oscuras originan estelas perfectamente visibles.

Dijimos ya que de la inextricabilidad más el holismo se obtenía la indeterminación; e inextricabilidad no significa otra cosa que el rechazo a la dicotomía analítico/sintético. Veamos, pues, la tesis holista en relación con el problema de la analiticidad. Una oración analítica es, supuestamente, aquella que resulta verdadera por convención lingüística y no por haber sido confirmada por la experiencia. Pues bien, es esta dualidad la que Quine rechaza: «Tomada colectivamente, la ciencia tiene una doble dependencia, del lenguaje y de la experiencia; pero esta dualidad no se puede rastrear significativamente en las oraciones tomadas una a una» (Quine 1951, p. 42). Así, la tesis holista viene a defender: a) que la experiencia no verifica una oración aislada, sino una teoría científica entera o una parte suficientemente amplia y b) es imposible trazar una frontera nítida entre oraciones verdaderas por mera convención lingüística (sea del tipo que sea) y oraciones verdaderas por la información empírica que transmiten. Quine lo expresa diciendo que

Tanto (1) como (2) son de un alto nivel teórico. Separadas del resto de la teoría no pueden ser testadas empíricamente, aunque sólo sea porque los electrones son inobservables. (La oración 'los aviones a reacción ocasionan estelas perfectamente visibles' no exige, sin embargo, posesión de más teoría que la que se precisa para aprender castellano, de ahí que sí pueda ser testada empíricamente por sí sola).

Sin embargo, la conjunción entre (1) y (2) sí puede someterse a contrastación empírica:

- (3) Por esta cámara oscura se mueve un electrón y los electrones en movimiento dentro de cámaras oscuras ocasionan estelas perfectamente visibles.

Esta cámara oscura sin una estela perfectamente visible contradice (3). Así pues, ni (1) ni (2) se pueden testar empíricamente, sólo la conjunción entre ellas se puede someter al test. Además, una vez que hemos refutado (3) – viendo que en esta cámara oscura no se origina estela alguna – no queda unívocamente determinado si hemos de rechazar (1), (2) o ambas. Así, si el lenguaje científico se desarrollara por inducción, aplicando un término a nuevas instancias por similitudes observables, el lenguaje no se vería afectado por el holismo; la cuestión es que muy pronto entra en juego el método hipotético-deductivo.

Si esta noción [el holismo] es correcta, es equívoco hablar del contenido empírico de una oración individual – especialmente si es una oración muy alejada de la periferia experimental del campo. Además, sería absurdo buscar una frontera entre oraciones sintéticas, que dependen contingentemente de la experiencia, y oraciones analíticas, que rigen pase lo que pase. Cualquier oración puede ser considerada verdadera pase lo que pase, si realizamos ajustes lo suficientemente drásticos en el sistema. (Quine 1951, p. 43).

Como es sabido, Carnap defendía que las oraciones analíticas eran las oraciones que estipulaban las reglas de nuestro lenguaje, aunque ello no las convirtiera en intocables: cuáles de nuestras reglas estamos dispuestos a cambiar y cuáles no, es algo que no se puede decir *a priori*. A veces podemos decidir cambiar no una regla del lenguaje físico – el que habla sobre la realidad –, sino una de las que describe nuestro lenguaje, o sea, de las consideradas analíticas (Carnap 1937, p. 318). Quine, por el contrario, no acepta ninguna explicación de la analiticidad. Si construimos un lenguaje artificial, podemos confeccionar una lista de oraciones y colocarlas bajo el rótulo de «reglas» o «definiciones», pero si estamos describiendo nuestro lenguaje, es imposible decidir cuáles de sus oraciones constituyen reglas o definiciones y cuáles oraciones sintéticas. Uno de los puntos principales de *Dos dogmas del empirismo* (Quine 1951) es que no se puede decir, a partir de la conducta de los hablantes, qué oraciones articulan principios analíticos constitutivos de sus marcos lingüísticos. De ahí que la distinción entre oraciones analíticas y sintéticas sea uno de los dogmas del empirismo que Quine pretende superar. El otro es el *reduccionismo*, la idea ya comentada de que no existen

oraciones puramente observacionales o, al menos, de que nuestras oraciones teoréticas no son reductibles a oraciones observacionales. El holismo acaba con ambos dogmas.

Quine, según Gibson, ofrece tres razones a favor de su holismo:

- 1) La ya familiar idea de que las hipótesis se obtienen siempre a partir de otras hipótesis que se asumen como verdaderas. En caso de que una hipótesis no supere la prueba empírica, podemos sentirnos libres para desechar la hipótesis en cuestión o cualquiera de las que habíamos asumido. Cuál va a ser la decisión no es algo que se pueda decir *a priori*. No hay receta para esto, pero la maximización de la simplicidad y la minimización de la mutilación son máximas con las que la ciencia se esfuerza por la vindicación en predicciones futuras.
- 2) La idea, comentada arriba, de que si el holismo fuera falso, cada oración individual de una teoría tendría su propio significado empírico y podríamos desarrollar una teoría verificacionista adecuada. El ejemplo del electrón en la cámara oscura muestra que esto no es así, luego el holismo es correcto.
- 3) La razón más poderosa a favor del holismo semántico es que aprendemos el lenguaje por medio de síntesis analógica, esto es, por analogías sobre cuyo funcionamiento no conocemos prácticamente nada. El lenguaje es

una red hecha de «términos teóricos unidos por hipótesis fabricadas, y asociadas a eventos observables aquí y allí» (Quine 1975a, p. 71).

Ésta última parece ser para Gibson la razón más importante. Así, hace notar que la auténtica razón para defender el holismo es que el lenguaje científico es el producto de «saltos irreductibles de analogía» y, por tanto, no es reducible a la pura observación. El proceso mediante el cual el aprendiz pasa de hablar de colores y objetos físicos a utilizar términos abstractos y cuantificar sobre propiedades, números, funciones o mencionar hipotéticas partículas físicas y fuerzas, no es una derivación continua, que nos permita, siguiéndola hacia atrás, reducir la teoría científica a pura observación; es un progreso que sucede, más bien, por breves saltos de analogía.

La manera como el holismo apoya la indeterminación – hasta el punto de que podemos considerarlo la razón principal de ésta – resulta bastante clara:

- a) Significados y creencias no son cosas diferenciables – la llamada *inextricabilidad* – . Como dice el propio Quine: «No aprendemos primero de qué hablar y luego qué decir sobre ello» (Quine 1960, p. 16).

- b) No podemos diferenciar entre creencias que forman parte invariablemente del significado, del concepto, y otras que sean meramente empíricas y contingentes – esto es, el rechazo a la dicotomía analítico/sintético – .

- c) Hay holismo en nuestro sistema de creencias y, por tanto, de significados, es decir, guardan entre ellos complejas relaciones (unos se apoyan, dependen, provienen, etc. de otros).

Por tanto,

- d) las oraciones o los términos, considerados individualmente, no tienen un significado o una referencia propios determinados. En palabras de Feyerabend: «el significado de un término no es una propiedad intrínseca sino que depende de la manera en que el término ha sido incorporado a la teoría» (Feyerabend 1962, p. 68), o en la versión quineana, subrayando la dependencia de los términos respecto a las oraciones: «Se tiene que ver la doctrina molecular en funcionamiento en la teoría física para obtener una noción adecuada de las moléculas (...). Se trata de aprender la palabra contextualmente como un fragmento de oraciones que se aprende a emitir como un todo en las circunstancias apropiadas» (Quine 1960, p. 15).

El holismo puede ser visto como un problema de difícil solución para el intérprete radical: ¿cómo es posible comenzar la labor interpretativa si para descubrir cualquier creencia del interpretado se hace necesario conocer otras varias antes? La respuesta es, como ya hemos comentado, la caridad. Comenzamos a interpretar *presuponiendo* que el

interpretado comparte con nosotros criterios básicos de racionalidad (lo que nosotros consideramos racionalidad, claro).

Davidson ha hecho hincapié a lo largo de su obra en la relevancia de la inextricabilidad esencial entre significados, creencias e intenciones – de ahí que afirme que una teoría de la interpretación del tipo defendido por él es una tarea común del lingüista, el psicólogo y el filósofo (Davidson 1974b, p. 142) – . Así, afirma:

Hay un obstáculo de principio, y no meramente práctico, para verificar la existencia de creencias e intenciones detalladas, generales y abstractas, mientras seamos incapaces de decir lo que significan las palabras de un hablante [...]

Mi tesis es simplemente que darle un sentido detallado a las intenciones y creencias de una persona no puede ser independiente de darle sentido a sus emisiones (Davidson 1974b, p. 144).

Y algo más adelante:

Ya que no podemos esperar interpretar la actividad lingüística sin saber lo que un hablante cree, y no podemos hallar una teoría de lo que quiere decir sobre un descubrimiento anterior de sus creencias e intenciones, concluyo que al interpretar las emisiones partiendo de cero [*from scratch*] – en la interpretación *radical* – debemos, de alguna manera, ofrecer simultáneamente una teoría de la creencia y una teoría del significado. (Davidson 1974b, p. 144)

Cuando aprendemos una lengua, aprendemos a la vez toda una red de creencias y una forma de expresar y adscribir intenciones; lenguaje, creencias e intenciones se hallan indisolublemente unidos: «De manera general, mi idea es que deberíamos pensar en los significados y las creencias como constructos interrelacionados de una única teoría[.]» (Davidson 1974, p. 146). En efecto, Davidson tiene en mente una teoría no sólo de la conducta lingüística, sino de la conducta en general. Y el fenómeno de la indeterminación no haría más que subrayar la interrelación que guardan estos elementos – lenguaje, creencias, intenciones – : «la indeterminación [del significado] es importante sólo para llamar la atención hacia cómo la interpretación del lenguaje debe ir de la mano con la interpretación de la acción en general, y así, con la atribución de deseos y creencias» (Davidson 1984a, p. 154).

Para interpretar a un hablante hemos de imputarle creencias. Davidson ha hecho hincapié repetidamente en esta necesidad de principio:

Si todo lo que sabemos es qué oraciones considera verdaderas un hablante, y no podemos asumir que su lenguaje es el nuestro, entonces ni tan siquiera podemos dar un primer paso hacia la interpretación sin saber o asumir muchas cosas sobre las creencias del hablante. Ya que el conocimiento de las creencias se obtiene sólo con la habilidad de interpretar las palabras, la única posibilidad al comienzo es asumir un acuerdo general sobre las creencias. (Davidson 1974b, p. 196)

Y si no asumimos ese acuerdo general:

Pero cuando otros piensan de manera diferente a nosotros, ningún principio general ni apelación a la evidencia nos puede forzar a decidir que la diferencia se halla en nuestras creencias y no en nuestros conceptos. (Davidson 1974b, p. 197)

Hemos de recordar, eso sí, que la apelación al *principio de caridad* hace que la indeterminación de Davidson sea mucho más limitada que la defendida por Quine. Además, no sólo se trata aquí de maximizar el acuerdo del interpretado con nosotros, sino que, a la hora de imputar creencias y deseos, se hace imprescindible la presunción metodológica de la racionalidad (esto sería el llamado *principio de humanidad*). El *principio de caridad* y el *principio de humanidad* son, pues, constantemente utilizados en la construcción de la teoría. Le imputamos al hablante que se interpreta la mayor cantidad de creencias verdaderas que sea posible y, de manera semejante, maximizamos su racionalidad. Davidson dice, pues, que la creencia y el deseo que explican una acción deben «exhibir la racionalidad de la acción a la luz del contenido de la creencia y el objeto de la intención» (Davidson 1974b, p. 159). Así, si alguien levanta la mano, su acción quedará teleológicamente explicada si suponemos que quería llamar la atención de un amigo suyo que había cerca y creía que levantando la mano llamaría la atención de su amigo. La corrección de una explicación teleológica se basa, como vemos, en su habilidad para descubrir patrones coherentes en la conducta de un agente. Y, explica Davidson, la coherencia aquí incluye la idea de la racionalidad, tanto en el sentido de que la acción explicada ha de ser racional a la luz de las creencias e intenciones adscritas, como en el de que unas y otras han de casar (*fit with one another*, se lee en Davidson 1975, p. 159). Estas consideraciones hacen que la teoría de la interpretación,

dentro del sistema davidsoniano, haya de incardinarse dentro de una teoría comprensiva de la acción y el pensamiento.

La asunción de la racionalidad del hablante interpretado no ha de ser vista como un don que le presuponemos a éste, sino como una condición *sine qua non* de la interpretación. Davidson volvía sobre este punto en una de sus últimas obras:

El pensador conceptualmente desarrollado posee dos mecanismos interpretativos básicos a su alcance para comprender las emisiones y las acciones de otros: la asunción de una racionalidad suficiente para hacer inteligible estas acciones, y el conocimiento de cómo la percepción determina [*yields*] los contenidos de la creencia. En el caso del lenguaje, esto es fácil de ilustrar, y la lección se puede aplicar a los contenidos proposicionales. Las oraciones, o mejor, las actitudes que expresan, le deben su contenido, esto es, su significado, a dos factores: sus relaciones con otras oraciones o actitudes, y sus relaciones, directas o indirectas, con el mundo a través de la percepción. Es, por tanto, imposible para un intérprete desechar completamente las relaciones lógicas entre las oraciones o las actitudes de un hablante. Esto no es una cuestión de que los agentes hacen un esfuerzo para ser consistentes; es una cuestión de que su lenguaje y conducta tienen el significado que tienen por cómo dependen mutuamente. Al menos que haya suficiente coherencia, es imposible asignar contenidos proposicionales a su lenguaje, creencias o intenciones. La asunción del intérprete de un grado de racionalidad por parte de aquellos a quienes desea entender no es, pues, más que una condición para poder entenderlos. (Davidson 1998, p. 89).

Davidson niega, pues, que se pueda confeccionar un manual de interpretación sin atribuir a la vez deseos y creencias; la interpretación es vista como un proceso

global de adscripción de creencias, intenciones y significados simultáneamente, que no podría ni tan siquiera comenzar a menos que se partiera de ciertos supuestos acerca del sujeto al que se pretende interpretar. Estos supuestos son los dos mencionados como «mecanismos interpretativos» en la cita anterior – el externismo y el principio de humanidad – más el principio de caridad.¹⁹

Este último, como dice Hookway, pasa en Davidson de constituir una máxima regulativa que nos permite vivir con la indeterminación de la traducción a ser un principio que determina qué teorías del significado son empíricamente correctas. La caridad se convierte, pues, en un elemento constitutivo de la corrección y no en una máxima regulativa (Hookway 1988, p. 170). Y, como comentábamos antes, al adscribir creencias y deseos a la gente – y al decidir cuál es el significado de sus oraciones – , *racionalizamos* su conducta. Eso sí, la única vara de medir a la hora de esa racionalización, es la nuestra propia. Nuestro principio de caridad no es demasiado sofisticado en ese sentido. Como dice Hookway:

En lo que podemos confiar es en nuestra visión de lo que uno debería creer, si estuviera en el lugar en que está el sujeto, y cómo uno debería razonar... Todo lo que podemos establecer es qué parece verdadero y racional *a nuestros ojos* [*by our lights*], y, si las creencias de alguien cumplen esas condiciones, esto nos provee con una base para una predicción y corroboración fáciles. Nos usamos a nosotros mismos como instrumentos de medida al establecer en qué medida las interpretaciones cumplen las condiciones normativas: las interpretaciones recogen y explican cuánto difieren las creencias y las

¹⁹ Para una exposición acerca de por qué quedaría bloqueada la interpretación en caso de no asumir dichos supuestos, v. C. Moya, *op. cit.*

capacidades del sujeto de las nuestras. [...] Estamos satisfechos cuando sus acciones nos parecen inteligibles a nosotros, incluyendo sus acciones lingüísticas. (Hookway 1988, p. 173).

Ahora bien, hemos visto cómo Davidson defiende la indeterminación del significado y de la referencia a pesar de su constante apelación al principio de caridad en el proceso interpretativo. Esto se debe, en parte, a que «no surge ningún principio único de caridad óptima; los requisitos, pues, no determinan una única teoría.» (Davidson 1967, p. 27). Así, aun aplicando el principio de caridad como una condición de corrección obtendremos manuales correctos incompatibles entre sí. Uno de ellos puede hacernos atribuirle un mínimo de creencias erróneas al nativo, pero quizá a cambio de imputarle cierta irracionalidad o inconsistencia; otro de los manuales nos ofrecerá al sujeto como completamente razonable y consistente, pero adscribiéndole deseos que nos resultan difíciles de comprender; otro manual puede hacernos al sujeto interpretado una persona completamente racional, inteligible para nosotros y con creencias normalmente verdaderas, pero utilizando sistemas de clasificación excéntricos; quizá haya un manual que cumple todas nuestras condiciones, pero con toda seguridad necesitará postular tantas hipótesis *ad hoc* que será difícilmente aplicable. Por tanto

Careciendo de una regla estricta e inmediata [*hard and fast*] sobre las relaciones de prioridad entre las diferentes condiciones normativas, no tenemos base firme para la elección entre las teorías que cumplen las diferentes condiciones relativamente bien, pero imperfectamente. La indeterminación surge de la necesidad de equilibrio [*trade-off*] entre las diferentes condiciones que constituyen nuestra práctica de la interpretación. A

menos que una traducción satisfaga perfectamente todas las condiciones (y eso es muy improbable), la indeterminación parece inevitable [...]. (Hookway 1988, p. 179).

En palabras de un comentador:

La caridad, con todo, no puede ser concebida como proveyendo un procedimiento algorítmico para la interpretación de una oración particular. No nos equipa con un procedimiento preciso para tratar con ninguna de las diferencias observadas en la conducta lingüística del hablante y el intérprete. El único criterio para la interpretación correcta de una oración específica es la coherencia general [*overall fit*] entre la interpretación de una oración específica y la interpretación de todas las demás oraciones. La caridad, así, deja indeterminado dónde se debe trazar el límite entre una diferencia en la creencia y una diferencia en el concepto. En el contexto de la interpretación radical, la tesis del holismo, la indeterminación y el rechazo de la distinción analítico/sintético van de la mano. (K. Stüber 1997, p. 296).

Veremos en 2.1 cómo Searle ataca esta visión holista, defendiendo que el hecho de que una oración tenga significado es, obviamente, relativo a una lengua ('es regnet' tiene significado en alemán y no en castellano), pero su significado no puede ser relativo a nada dentro de esa lengua ('es regnet', en alemán, significa lo que significa).

1.8. ¿In- o infradeterminación?

La tesis de la infradeterminación es uno de los pilares de la filosofía de la ciencia de Quine y halla asimismo aplicación en su filosofía del lenguaje, en cuanto que un manual de traducción no es otra cosa que una construcción teórica y, por tanto, puede ser tratado como una teoría científica más. La tesis defiende, *grosso modo*, que todas las observaciones posibles – la suma de todas las observaciones realizadas por observadores idealizados que pueden observar en cualquier tiempo y en cualquier lugar – no determina unívocamente la elección de una teoría científica. En otras palabras, no importa de cuántos y cuán valiosos datos dispongamos, siempre habrá varias teorías incompatibles entre sí que expliquen igualmente bien los datos observados. En palabras de Quine:

Ahora mi idea sobre la teoría física es que la teoría física está infradeterminada por todas... las observaciones posibles... Las teorías físicas pueden ser incompatibles [*be at odds*] entre ellas y, sin embargo, ser compatibles con todos los datos posibles incluso en el sentido más amplio. En una palabra, pueden ser lógicamente incompatibles y empíricamente equivalentes. Sobre esta idea espero un amplio acuerdo [*wide agreement*], aunque sólo sea porque los criterios observacionales de los términos teóricos son normalmente muy flexibles y fragmentarios. Quienes están de acuerdo con esta idea general no tienen por qué estarlo respecto a hasta qué punto la teoría física está determinada empíricamente en este sentido fuerte; algunos reconocerán una brecha tal sólo en los ámbitos más elevados y especulativos de la teoría física, mientras que otros la verán extendiéndose incluso a los rasgos de sentido común de los cuerpos macroscópicos. (Quine 1970a, p. 179.).

Y, a continuación Quine establece el nexo que antes apuntábamos entre esta tesis y el problema de la indeterminación:

Miremos ahora la traducción radical de la teoría de un físico radicalmente extraño. Como siempre en la traducción radical, el punto de comienzo es el emparejamiento de oraciones observacionales de las dos lenguas mediante un emparejamiento de significados estimulativos. Para formar después las oraciones teóricas del extraño hemos de proyectar hipótesis analíticas, cuya última justificación es, en lo sustancial, simplemente que las oraciones observacionales implicadas casen. Pero ahora reaparece en segunda intención la vieja brecha empírica, la vieja indeterminación entre teorías físicas. En la medida en que la verdad de la teoría física está infradeterminada por los datos observables, la traducción de la teoría física del extraño está infradeterminada por la traducción de sus oraciones observacionales. Si nuestra teoría física puede variar aunque todas las observaciones posibles estén determinadas, entonces nuestra traducción de su teoría física puede variar aunque nuestra traducción de todas sus oraciones observacionales esté determinada. Nuestra traducción de sus oraciones observacionales no determina más nuestra traducción de su teoría física que lo que nuestras observaciones posibles determinan nuestra teoría física. (Quine 1970a, p. 181).

Vemos cómo Quine trata la tesis de la infradeterminación de las teorías científicas casi como una verdad perogrullesca, afirmando que es un punto sobre el que espera «amplio acuerdo», y ofreciendo como argumento a su favor el que los criterios observacionales de los términos teóricos son muy flexibles y fragmentarios. A continuación Quine expone su idea de que aun cuando la traducción de las oraciones observacionales se hallara determinada, la traducción de las oraciones más teóricas se hallaría

infradeterminada con respecto a ésta. Esta idea nos podría hacer pensar que Quine ha rechazado la célebre tesis de la indeterminación de la traducción a favor de la más laxa infradeterminación. Pero no, aunque lo que hemos expuesto arriba habla a favor de la infradeterminación de la traducción de la teoría física respecto a la de las oraciones observacionales y que, como dijimos al comienzo, podríamos considerar un manual de traducción como una construcción teórica más y, por tanto, defender su infradeterminación respecto a toda la evidencia posible, Quine continúa impertérrito a la hora de defender la indeterminación – no mera infradeterminación – de la traducción:

La indeterminación de la traducción no es sólo una instancia del carácter infradeterminado empíricamente de la física. La idea no es sólo que la lingüística, al ser una parte de la ciencia de la conducta y, por tanto, en última instancia, de la física, comparte el carácter infradeterminado de la física. Por el contrario, la indeterminación de la traducción es adicional. Donde las teorías físicas A y B sean ambas compatibles con todos los datos posibles, podríamos adoptar A para nosotros y aún ser libres para traducir al nativo tanto como creyendo A como creyendo B. (Quine 1970a, p. 182).

La idea de que la indeterminación de la traducción no es una instancia más de la infradeterminación de cualquier teoría empírica respecto a la evidencia posible, sino que se trata de un rasgo adicional de las teorías del significado ha sido repetido por Quine en varios ensayos. Ahora bien, ¿qué significa ese «adicional»? A partir de lo expuesto arriba debemos plantear tres cuestiones que nos aclararán la duda:

1. ¿Por qué la traducción se halla indeterminada y no infradeterminada, como todas las teorías científicas?
2. ¿Por qué las teorías científicas se hallan infradeterminadas y no indeterminadas, como la traducción?
3. ¿Por qué la traducción de la parte teórica del lenguaje se halla infradeterminada respecto a la traducción de la parte observacional?

Centrémonos ahora en los puntos 1 y 2, cuyas respuestas se solapan. Realmente, se trata de la misma pregunta desde dos perspectivas diferentes, mientras que en la tercera pregunta la cuestión es otra: ¿por qué aunque concediéramos la existencia de oraciones puramente observacionales y con un significado, por tanto, completamente determinado, el lenguaje teórico continuaría indeterminado? Como hemos dicho, Quine acepta la infradeterminación de las teorías físicas casi como una perogrullada.²⁰ Sin embargo, lo que no resulta tan perogrullesco es la idea de que mientras las teorías físicas se hallan infradeterminadas, la traducción se halla indeterminada, y ni tan siquiera adquiere el carácter de perogrullada partiendo de la misma base naturalista-conductista de Quine, esto es, admitiendo que para la lingüística no hay más evidencia que la de los patrones de a/disentimiento ante ciertas estimulaciones. Podemos continuar preguntándonos por

²⁰ La tesis ha hallado diferentes problemas que han obligado al propio Quine afirmar que ésta, tal como él la había concebido, resultaba insostenible. Con todo, una formulación más moderada siempre ha contado con su beneplácito, a saber: Nuestro sistema del mundo ha de tener alternativas empíricamente equivalentes que, si las descubriéramos, no podríamos hacer lógicamente equivalentes a nuestro sistema. (V. Gibson 1982, §3.3.4). Sin embargo, la tesis de la infradeterminación suena extraña en boca de alguien como Quine, que no admite más hechos acerca del significado que el significado estimulativo, ya que si no hay más significado que éste, y las teorías se distinguen por las predicciones que hacen de la realidad, entonces no tiene sentido hablar de dos teorías distintas – lógicamente incompatibles – que coincidan, sin embargo, en todas las oraciones observacionales (o, más exactamente, categóricas observacionales) que de ellas se derivan. De hecho, R. Kirk comenta (en la nota 8, pág. 149 de Kirk 1987) que Quine reconoció en una carta privada que esta crítica invalida la línea argumental seguida en «On the Reasons For the Indeterminacy of Translation».

qué la infradeterminación de toda construcción teórica respecto a toda la evidencia posible se torna en indeterminación cuando la teoría resulta ser un manual de traducción. En efecto, algunos lectores han tenido dificultades en ver la diferencia exacta entre la infradeterminación de las teorías científicas y la indeterminación de la traducción.²¹ La cuestión clave aquí sería: ¿qué distingue a una teoría del significado de una teoría física para hacer que Quine no hable de la infradeterminación o de la indeterminación de ambas, sino precisamente de la indeterminación de la primera pero de la infradeterminación de la segunda? Quine responde así:

La indeterminación de la traducción difiere de la infradeterminación de la ciencia en que para los manuales de traducción sólo hay una cosa con la que estar de acuerdo o en desacuerdo, la conducta verbal de los nativos; nada se dice sobre ocultos mecanismos neuronales. Si los traductores no se ponen de acuerdo sobre la traducción de una oración de la Jungla pero nada en la conducta de la gente de la Jungla puede dilucidar el desacuerdo, entonces, simplemente, no hay materia objetiva [*fact of the matter*]. En el caso de la ciencia natural, por otra parte, hay materia objetiva, aun cuando todas las observaciones posibles sean insuficientes para revelarla unívocamente. Los hechos de la naturaleza desbordan [*outrun*] nuestras teorías, así como todas las observaciones posibles, mientras que la semántica tradicional va más allá de los hechos del lenguaje.

[...] La ciencia natural, asumámoslo de nuevo, está infradeterminada por toda la observación posible. Con todo, supón que nos hemos decidido por una de las muchas teorías globales que concuerdan con toda la observación posible. La traducción continúa indeterminada, incluso relativamente a la teoría de la naturaleza escogida. Así pues, la

²¹ «Algunos de mis lectores han tenido dificultad en ver que la tesis de la indeterminación de la traducción no es otra cosa que un caso especial de la tesis de que la ciencia natural se halla infradeterminada por toda la observación posible.» (Quine 1987, p. 9).

indeterminación de la traducción es una indeterminación adicional a la infradeterminación de la naturaleza. (Quine 1987, pp. 9-10).

Así pues, para la física hay materia objetiva (*fact of the matter*); para la traducción, no. Hagamos aquí una precisión para aclarar esto. En la labor del traductor radical podemos distinguir dos tipos de hipótesis²²: 1. Las que nos indican el significado estimulativo de una oración para un hablante; éstas son «genuinas» y proceden por el método habitual de la inducción; 2. Las que nos indican la referencia de los términos y, tras hacer esto, el significado de las oraciones; éstas son pseudohipótesis, pues pretenden ir más allá de los datos disponibles. Las hipótesis del primer tipo, las «genuinas», se limitan a emparejar oraciones estimulativamente sinónimas, cosa que se puede examinar empíricamente; son hipótesis acerca de la conducta verbal y la estimulación de los hablantes. Pero las del segundo tipo son hipótesis acerca de lo que los hablantes *quieren decir* con sus emisiones, o sea, no sólo indagan acerca de en qué ocasiones emiten los hablantes ciertas oraciones, sino que analizan éstas y atribuyen referencia a los términos componentes. Pero, para Quine, la referencia es inescrutable, lo que no quiere decir que los datos empíricos (los patrones de a/disentimiento de los hablantes) sean insuficientes para escrutarla, sino que no hay nada que escrutar. La atribución de una referencia a los términos es siempre arbitraria, ya que muchas otras eran posibles y ningún dato acerca de las disposiciones lingüísticas o la condición neuronal del hablante determinan la elección de una referencia. Luego, una hipótesis acerca de la referencia de un término es una hipótesis acerca de nada.

²² Se sigue aquí el ensayo de Manley Thompson (en Hahn y Schlipp (eds.) 1986).

De ahí que el primer conjunto de hipótesis sean genuinas hipótesis acerca del significado, mientras que las segundas sean pseudohipótesis, ya que pretenden trascender toda la evidencia posible. El primer conjunto no está formado por hipótesis *analíticas*, pues nada analizan, sino que se limitan a emparejar oraciones ateniéndose fielmente a la única evidencia disponible: la que ofrece el significado estimulativo. Es el segundo conjunto el que introduce la infradeterminación, ya que diferentes conjuntos de hipótesis analíticas casan igualmente bien con la evidencia posible:

Respecto a estar infradeterminadas por todos los datos posibles, la sinonimia en la traducción y la teoría física son similares. La totalidad de las observaciones posibles de la naturaleza, hechas y no hechas, es compatible con teorías físicas que son incompatibles entre sí. Por su parte, la totalidad de las observaciones posibles de la conducta verbal, hechas y no hechas, es compatible con sistemas de hipótesis analíticas de traducción que son incompatibles entre sí. Hasta aquí se mantiene el paralelismo. (Quine 1969b, pp. 302-3).

El paralelismo se quiebra en el momento en el que consideramos que no hay verdad superior a la que nos ofrece la teoría física. Siempre nos hallamos dentro de una teoría, no hay manera de observar las teorías a vista de pájaro. Cuando dos científicos discuten sobre la masa de los neutrinos, lo hacen, inevitablemente, desde una teoría previamente aceptada, y desde ésta consideran todos los datos de los que disponen – sus observaciones – , mientras que cuando dos lingüistas discuten sobre la traducción de ‘gavagai’ lo hacen sobre una cuestión para la que no hay materia objetiva, ni tan

siquiera dentro de su teoría infradeterminada de hipótesis analíticas. Esto es, las hipótesis del tipo 2 se refieren a una materia inexistente, aunque se basen en las del tipo 1, para las que sí hay materia objetiva y, por tanto, se hallan infradeterminadas. La quiebra del paralelismo entre teorías científicas y manuales de traducción consiste pues, para Quine:

Esencialmente en esto. La teoría en física es un parámetro último. No hay una primera filosofía legítima, más elevada o más firme que la física, a la que apelar sobre la cabeza de los físicos. Incluso nuestra noción de la arbitrariedad parcial o de la infradeterminación de nuestra teoría global de la naturaleza no es una intuición de nivel más elevado; es integral a nuestra misma teoría de la naturaleza infradeterminada, y de nosotros mismos como objetos naturales. Así que vamos razonando y afirmando lo mejor que podemos dentro de nuestra siempre infradeterminada y evolutiva [*evolving*] teoría de la naturaleza, la mejor que en cada momento podemos formar; y normalmente es redundante citar la teoría como parámetro de nuestras aserciones, ya que no hay ningún estándar mayor. Deja de ser redundante sólo cuando estamos contrastando teorías alternativas en un nivel más profundo, por ejemplo, con la previsión de un cambio.

Aunque la lingüística es, claro está, una parte de la teoría de la naturaleza, la indeterminación de la traducción no está simplemente heredada como un caso especial de la infradeterminación de nuestra teoría de la naturaleza. Es paralela pero adicional. Así, adopta por ahora mi actitud completamente realista hacia los electrones y los muones y el espacio-tiempo curvo, aceptando así la teoría actual del mundo a pesar de saber que se halla, por principio, metodológicamente infradeterminada. Considera, desde este punto de vista realista, la totalidad de las verdades de la naturaleza, conocidas y no conocidas, observables e inobservables, pasadas y futuras. La cuestión sobre la indeterminación de la traducción es que soporta incluso toda esta verdad, toda la verdad sobre la naturaleza. Esto es lo que quiero decir cuando digo que, donde se aplica la

indeterminación de la traducción, no hay una cuestión real de elección correcta [*real question of right choice*], no hay materia objetiva [*fact of the matter*] incluso dentro de la reconocida infradeterminación de una teoría de la naturaleza. (Quine 1969b, p. 303).

Repitiendo, cuando dos científicos discuten, lo hacen desde una determinada teoría, por mucho que haya sido, en parte, arbitrariamente escogida, y una vez en ella describen los eventos que perciben de acuerdo con el marco que la teoría les ofrece. Sin embargo, dos traductores que discuten sobre la traducción correcta de ‘gavagai’, aun cuando ya hayan escogido un cierto marco de traducción – el reista, pongamos por caso – , no disponen de datos algunos sobre la materia que están discutiendo. La diferencia es, pues, tan contundente como la que media entre un lugar donde hay algo, por mucho que este algo pueda ser descrito de diferentes formas, y un lugar donde no hay nada.

Hemos visto, asimismo, cómo el argumento de Quine se basa, en última instancia, en su rechazo de una filosofía primera. Aquí Quine se desmarca de los primeros empiristas, especialmente Hume, quien persiguió con ahínco la fundamentación de nuestro conocimiento, por tanto de la ciencia, pero desde fuera de la ciencia, en orden a evitar caer en una petición de principio. Así pues, la noción de que la ciencia investiga nuestras sensaciones privadas no constituye en el sistema humeano una teoría científica, sino metafísica. Utilizar los datos de la ciencia para fundamentar el conocimiento científico implicaría caer en un círculo vicioso. No obstante, el naturalismo radical de Quine sí que le permite hacerlo y afirmar, por tanto, que no hay más fundamento para la ciencia que el que ella misma pueda darse. Ha de ser la propia ciencia, y no la

metafísica disfrazada de epistemología, la que explique cómo se produce el conocimiento de la realidad física. (V. lo dicho en 1.1 sobre el naturalismo de Quine).

Este cambio de talante entre el antiguo y el nuevo empirismo es expuesto así por Gibson:

En la antigua epistemología el objetivo era comenzar desde verdades auto-evidentes, no científicas, sobre la experiencia sensorial y deducir (por pasos auto-evidentes) o construir (por métodos de la lógica y la teoría de conjuntos) todas las verdades de la ciencia natural. En un programa semejante, como antes se dijo, no sería permisible usar los hallazgos de las teorías científicas como parte de los datos iniciales. La idea era que nuestra teoría del conocimiento, nuestra filosofía primera, debería, de alguna manera, contener pero no estar contenida en la ciencia. Ahora, con todo, tras el fracaso del empirismo radical, Quine sugiere que el objetivo para el empirismo emergente es prácticamente el mismo, si no idéntico al del viejo empirismo. *El objetivo del nuevo empirismo es ofrecer una explicación de cómo, dada sólo la evidencia de nuestros sentidos, construimos (pero no deducimos) nuestra teoría del mundo.* (Gibson 1982, p. 4).

No existe otro conocimiento que el científico; sobre esta tesis se yergue el naturalismo radical quineano. De aquí que Quine diera por errada la tentativa de Carnap de derivar todos los términos teóricos de los términos observacionales y los auxiliares lógico-matemáticos, ya que esto supone colocar estos dos últimos en el lugar que antes ocupaba la metafísica, esto es, en un reino de conocimiento precientífico que establece el fundamento del conocimiento científico. Contra el intento de la *Aufbau* de Carnap

Quine exclama: «¿Pero por qué toda esta reconstrucción creativa, toda esta ensoñación [*make-believe*]? La estimulación de sus receptores nerviosos es toda la evidencia que cualquier persona tiene para, en última instancia, llegar a su imagen del mundo. ¿Por qué no, simplemente, ver cómo se forma realmente esta construcción? ¿Por qué no decantarse por la psicología?» (Quine 1969a, p. 75). E incluso la ontología constituye a los ojos de Quine una labor científica. El hablar de objetos físicos supone ya asumir una teoría, aunque muy rudimentaria y arraigada en el ser humano. La ciencia trabaja con el único dato realmente objetivo que hay: nuestra estimulación nerviosa. Y aun esto supone ya una teoría científica, pues apela a seres humanos, nervios y estimulaciones, luego, como se ve, no hay más fundamento para el conocimiento científico que el que se pueda hallar en el conocimiento científico, pero esto no supone escepticismo alguno, ya que, como dijimos, siempre vemos el mundo desde una teoría, no hay forma de buscar lo auténticamente objetivo por medio de no utilizar ninguna teoría.

Nuestro discurso de cosas externas, nuestra misma noción de cosas, es simplemente un aparato conceptual que nos ayuda a prever y controlar la estimulación de nuestros receptores sensoriales a la luz de estimulación previa de nuestros receptores sensoriales. La estimulación, en primer y en último lugar, es todo lo que tenemos para proceder. Al decir esto yo también estoy hablando de cosas externas, a saber, personas y sus terminaciones nerviosas. Así, lo que estoy diciendo se aplica particularmente a lo que estoy diciendo, y no pretende resultar escéptico. No hay nada de lo que podamos estar más seguros que de las cosas externas – con todo, de algunas de ellas – , de otras personas, palos, piedras. Pero se mantiene el hecho – un hecho de la misma ciencia – de que la ciencia es un puente conceptual que hemos construido nosotros, uniendo

estimulación sensorial con estimulación sensorial; no hay percepción extrasensorial.

(Quine 1981, pp. 1, 2).

No veamos estas consideraciones sobre el antifundamentismo quineano como una disquisición un tanto colateral al asunto. El antifundamentismo de Quine es su razón principal para defender la infradeterminación de la física – pero la indeterminación de la traducción – . Una crítica como la de Chomsky – que veremos a continuación y basada en el paralelismo entre in- e infraindeterminación – parecerá a quienes se sientan quineanos en este sentido «una objeción desviada», ya que

[I]a indeterminación de la traducción se mueve en un plano más amplio que las teorías científicas (puesto que afecta también a la concepción ontológica de sentido común de la comunidad hablante) y que es claramente anterior a ellas, al menos en el sentido de que las hipótesis analíticas preceden incluso al uso de aparato referencial.

Por otra parte la indeterminación semántica no es, por así decirlo, una cuestión meramente fáctica, como lo es la infradeterminación de las teorías científicas, sino de principio, y de ahí que no haya lugar en ella al problema de elegir la alternativa correcta.

(M. Garrido 1976, pp. 49-50).

Con todo, Chomsky insiste en lo desacertado del paralelo. Cuando un físico, esté utilizando el marco teórico que esté utilizando, predica la existencia de muones, campos electromagnéticos y de la fuerza gravitatoria, está yendo tan allá de los datos de los que dispone como el lingüista que, una vez observada la conducta verbal de los nativos, les atribuye uno de los marcos ontológicos posibles. ¿Por qué, pues, la física tiene una materia objetiva de la que carece la traducción? Las hipótesis del físico son

tan arbitrarias como las del lingüista. En la defensa de Quine de la infradeterminación de la física y la indeterminación de la traducción sólo vemos el tenaz rechazo a la admisión de referencias o significados constituidos por algo más que conducta verbal observable. Chomsky expresa esta crítica así:

Así, lo que distingue el caso de la física del caso del lenguaje es que en el caso del lenguaje, por alguna razón, no se nos permite tener una «teoría tentativa» [...]. No puede haber un conjunto determinado de hipótesis analíticas sobre el lenguaje en general. Necesitamos un nuevo conjunto para cada lenguaje (para ser más precisos, para cada hablante de cada lenguaje), ya que no hay nada universal sobre la forma del lenguaje. (Chomsky 1969, p. 62).

O, por expresar la misma idea con otras palabras:

Su [de Quine] fideísmo, si se me permite la palabra, con respecto a las construcciones explicativas de la realidad del mundo físico, tales como electrones, positrones o muones, contrasta con su incredulidad con respecto a las construcciones explicativas de la vida mental. (M. Garrido 1976, p. 51).

Centrémonos ahora en la tercera de las preguntas que habíamos planteado. Según el párrafo citado, Quine considera que la forma de traducir una teoría física nativa, pues, es la misma que en cualquier otro ámbito del lenguaje nativo: «emparejando oraciones observacionales de las dos lenguas por medio de un emparejamiento inductivo de significados estimulativos». Cuando pasamos a traducir oraciones no observacionales, oraciones con un mayor nivel de teoreticidad, hemos de proyectar hipótesis analíticas,

«cuya única justificación, sustancialmente, es sólo que las oraciones observacionales implicadas coinciden». Pero la forma en la que hemos traducido las oraciones observacionales, traducción que ahora podemos considerar como la evidencia disponible, no determina unívocamente la traducción de las oraciones más teóricas. O sea que aunque la traducción de todas las oraciones observacionales nativas estuviera determinada – como lo están las observaciones – , la traducción de las oraciones teóricas continuaría infradeterminada – como lo están nuestras teorías científicas – . De ahí que Quine hable de que hallamos aquí algo así como una indeterminación en segunda intensión.

Fijémonos en que Quine no defiende aquí la indeterminación de la traducción de la teoría física nativa, sino su infradeterminación respecto a la traducción de las oraciones observacionales. Esto es, aun aceptando que la traducción de las oraciones observacionales ha sido determinada (y Quine admite que de existir oraciones netamente observacionales, su significado estaría determinado), pues aun así podríamos traducir las oraciones más teóricas de diferentes maneras que, con todo, implicarán las mismas oraciones observacionales. Luego la traducción de la teoría se halla infradeterminada respecto a la traducción de las oraciones observacionales. En palabras arriba citadas de Quine: «Nuestra traducción de sus [las del nativo] oraciones observacionales no fija nuestra traducción de su teoría física más de lo que nuestras observaciones posibles fijan nuestra propia teoría». El que dos teorías A y B compatibles con toda la evidencia disponible se hallen infradeterminadas por ésta ofrece también la curiosa consecuencia de que podemos escoger A para nosotros y traducir al

nativo como defendiendo B, aun suponiendo que hemos determinado la traducción de sus oraciones observacionales.

Es importante, pues, percatarse de que el problema no reside en que no haya oraciones puramente observacionales (oraciones como ‘Rojo’ o ‘Frío’ son las que más se acercan a serlo), sino que aun habiéndolas, el significado de las oraciones ocasionales y eternas continuará indeterminado. Sin embargo, la única razón para defender que aun determinando el significado de las oraciones observacionales queda indeterminado el de la parte más teórica del lenguaje es que las oraciones de alto nivel teórico no tienen un anclaje a la realidad tan claro como el de las observacionales, sus criterios de uso son muchos más imprecisos (como decía Quine, «aunque sólo sea porque los criterios observacionales de los términos teóricos son por lo común muy flexibles y fragmentarios»), pero no hemos de extraer de ahí la conclusión de que su significado o la referencia de sus términos queda en un limbo insustancial. Los términos teóricos de las ciencias, por decirlo en vena popperiana, se usan tentativamente, de una manera muy inconcreta al principio para ir precisando su significado con el tiempo.

Podemos concluir esta sección con las palabras de Chomsky que, de nuevo, subrayan el paralelismo entre las teorías físicas y las teorías lingüísticas:

Interpretada en un contexto epistemológico, como una pretensión acerca de la posibilidad de desarrollar una teoría lingüística, la tesis de Quine es meramente una versión de los argumentos escépticos familiares, que pueden aplicarse con idéntico derecho a la física, al problema de la percepción verídica, o, a este respecto, a sus “hipótesis genuinas”. Es bien

cierto que para que las hipótesis concernientes al lenguaje de un hablante nativo sean hipótesis serias (...) han de “ir más allá de la experiencia”. Si no lo fueran, carecerían de interés. Y puesto que van más allá del mero sumario de los datos, se dará el caso de que haya suposiciones que sean consistentes con los datos y que compitan entre sí. Ahora bien, ¿qué hay de sorprendente o de preocupante si tal sucede en esta ocasión? (Chomsky, *Synthese* XIX, 1, 1968, pp. 66-67. Citado en Garrido 1976, pp. 48-49).

2. INDETERMINACIÓN Y PRIMERA PERSONA

Hemos visto a lo largo de la sección anterior una perspectiva fundamentalmente conductista del lenguaje – en el caso de Quine – o, cuando menos, extensionalista – en el de Davidson –. Estos acercamientos al lenguaje han hecho aparecer el problema de la indeterminación del significado y la referencia. Veremos ahora una de las más importantes críticas que se han dirigido contra esta tesis y, por extensión, a un acercamiento semejante al lenguaje. El análisis se centrará en la crítica que Searle dirigió a la tesis de la indeterminación en el celebrado ensayo «Indeterminacy, Empiricism and the First Person», de 1987. La crítica defiende que la tesis supone una reducción al absurdo de sí misma y que, siendo el conductismo semántico la premisa sobre la que se construye el argumento, estamos obligados a admitir que los significados se hallan determinados más allá de lo que puedan determinar las condiciones de uso de las oraciones.

Además, ya que la autorefutación se debe, en el argumento searleano, a la no diferenciación entre la perspectiva de la primera y de la tercera persona o, más bien, a la consideración de que la perspectiva de la primera persona no ha de ser considerada científica o empírica, es también esta idea la que se ve reducida al absurdo: el conductismo quineano o el extensionalismo davidsoniano quedarían descartados como acercamientos admisibles al lenguaje.

En efecto, una de las implicaciones más sorprendentes – por antiintuitivas – de las tesis quineanas es el rechazo de la autoridad de la primera persona. Resulta manifiesto, *prima facie*, que los hablantes saben lo que quieren decir con sus palabras o, al menos, que lo saben mejor que cualquier otra persona. Como hemos visto, la indeterminación del significado y la referencia ponen esto en duda. Será, pues, la cuestión capital de la crítica.

Este capítulo gira en torno, pues, a dos cuestiones: por un lado, la supuesta reducción al absurdo en que cae el argumento de la indeterminación – se analiza esta crítica tal y como la han tratado, especialmente, Searle y Soames – , por otro, la apelación a la primera persona para eliminar la paradoja planteada.

2.1. Indeterminación y reducción al absurdo

John Searle ha realizado una crítica a la tesis quineano-davidsoniana de la indeterminación de la traducción y el significado que, de ser correcta, resultaría definitiva. La tesis, como ya hemos apuntado, se autorefuta, según Searle. Exponemos la crítica a continuación.

Para exponer el argumento de Quine hemos tenido que suponer en el lector la capacidad de distinguir entre los significados de oraciones como ‘Conejo’ y ‘Parte no separada de conejo’, pues, de lo contrario, o bien el argumento le habría resultado completamente baladí, o bien, sencillamente, no lo habría comprendido. Dicho de otra forma, la tesis defiende, siempre según Searle, que no hay distinción alguna entre el significado de ‘Conejo’ y ‘Parte no...’ y, sin embargo, si realmente no hubiera distinción alguna entre los significados de ambas oraciones, el argumento resultaría tan baladí que tendríamos la impresión de que lo que se nos ha dicho es que ‘Conejo’ y ‘Conejo’ significan lo mismo. Pero no es así, sino que tenemos la impresión de que el argumento sí ha afirmado algo no perogrullesco, con lo que el argumento acaba demostrando que ‘Conejo’ y ‘Parte no...’ no significan lo mismo, luego el argumento resulta ser una reducción al absurdo de sí mismo.

Por tanto, una teoría, como la de Quine, que equipare significado a patrones de uso, nos debe explicar por qué entonces la idea de que ‘Conejo’ y ‘Parte no...’ significan lo mismo, no suena igual que si se afirmara que ‘Conejo’ y ‘Conejo’ significan lo mismo. Si no se ofrece una explicación semejante, la conclusión es que el argumento constituye

una *reductio ad absurdum* de sí mismo, y que lo que realmente acaba demostrando es que oraciones con las mismas condiciones de uso no tienen el mismo significado, con lo que el significado no es equiparable a las condiciones de uso o los patrones de a/disentimiento – por lo que el conductismo lingüístico se hallaría errado en su base – .

Resumiendo con las palabras del propio Searle:

El estímulo – que, recordemos, se define enteramente en términos de estimulaciones de las terminaciones nerviosas – resulta igualmente apropiado para traducir ‘¡Gavagai!’ como ‘¡Hay un estadio en la historia vital de un conejo!’ o ‘¡Hay una parte no separada de un conejo!’ . El mismo patrón de estimulación de las células fotorreceptoras sirve [*does duty for*] para las tres traducciones. Así, si todo lo que hubiera respecto al significado fueran patrones de estímulo y respuesta, entonces sería imposible discriminar significados que de hecho se pueden discriminar. Ésta es la reducción al absurdo. (Searle 1987, pp. 124-125).

De hecho, la idea de que ‘Conejo’ y ‘Parte no...’ significan lo mismo se deduce directamente de que no hay más datos para determinar el significado que los observables – el a/disentimiento de los hablantes ante ciertas estimulaciones – , y una vez que hemos visto que ‘Conejo’ y ‘Parte no...’ no significan lo mismo (recordemos que es esencial para el propio argumento que no lo hagan), esa idea aparece como una asunción injustificada, una hipótesis que se ha adoptado y que ha resultado ser errónea. Al fin y al cabo, el conductismo lingüístico fue algo de lo que Quine partió, no a lo que Quine llegó:

La tesis de que no hay significados objetivamente reales además de disposiciones a la conducta verbal estaba ya asumida al comienzo de la discusión. Quine rechazaba cualquier apelación a los significados, en cualquier sentido psicológico, desde el principio. Eso nunca estuvo en cuestión. Lo que estaba en cuestión era la posibilidad de traducciones correctas motivadas empíricamente dado el conductismo; la cuestión era si queda o no una noción de igualdad de significado empíricamente motivada una vez que hemos adoptado un conductismo lingüístico extremo. (Searle 1987, p. 128).

Y, siempre según Searle, una vez comprobado que este «conductismo lingüístico» conduce a contradicciones, debe ser abandonado.

La reducción al absurdo se desarrollaría, pues, de la siguiente manera:

Premisa 1. Premisa del conductismo lingüístico: El significado de una oración consiste en las situaciones (o estimulaciones nerviosas) en las que un hablante está dispuesto a emitir dicha oración. (Esto es, pues, lo único que ha de ser considerado evidencia empírica en cuestiones de semántica.)

Premisa 2. Por tanto, la evidencia empírica no determina si debemos traducir ‘Gavagai’ como ‘Conejo’ o como ‘Parte de conejo...’ Ni siquiera determina si debo traducir la oración ‘Conejo’ de mi idiolecto como ‘Conejo’ o como ‘Parte de conejo...’

Ergo,

Conclusión: Las expresiones ‘Conejo’ y ‘Parte de conejo no separada’ significan lo mismo.

Pero, si la premisa 2 y la conclusión no resultan una perogrullada, como decir que ‘Conejo’ y ‘Conejo’ significan lo mismo, se debe a que los hablantes somos capaces de distinguir los diferentes significados de ‘Conejo’ y ‘Parte de conejo no separada’; luego debíamos contar con otra premisa adicional:

Premisa 3. Las oraciones ‘Conejo’ y ‘Parte de conejo no separada’ no significan lo mismo.

Dado que nuestras premisas han generado una contradicción, debemos rechazar alguna de ellas. Quine, en la lectura searleana de la tesis, rechaza 3; Searle cree que 2 y 3 son claramente aceptables, luego el problema está en 1.: «Creo que está claro cuál de las dos, (1) o (2) debemos abandonar. Quine, simplemente, ha refutado el conductismo lingüístico extremo» (Searle 1987, p. 126).

Desde el punto de vista davidsoniano, la cuestión es prácticamente idéntica. Veamos cómo la expone Searle:

Lo llamativo sobre Davidson es que, si se presenta el argumento como una serie de pasos, no se sigue que haya inescrutabilidad *a menos que* se añada una premisa extra sobre la naturaleza de una teoría empírica del lenguaje. Aquí están los pasos:

- (1) La unidad de análisis empírico en la interpretación radical es la oración (como opuesta a los elementos suboracionales).
- (2) La única evidencia empírica para la interpretación radical es el hecho de que los hablantes “consideren ciertas” ciertas oraciones en ciertas situaciones.
- (3) Hay formas alternativas de emparejar palabras con objetos que resultan inconsistentes, pero cualquiera de ellas puede explicar de manera igualmente correcta por qué un hablante consideraba cierta una oración.

Pero estas tres no implican inescrutabilidad o indeterminación sobre lo que el hablante realmente quería decir y sobre a lo que se refiere. Para eso se necesita una premisa extra. ¿Cuál? Creo que la siguiente:

- (4) Todos los hechos semánticos deben ser públicamente accesibles tanto al hablante como al oyente. Si el intérprete no puede hacer una distinción sobre la base de la evidencia pública, empírica, entonces no hay distinción alguna que hacer. (*Ibid*, p. 138).

Como vemos, esta premisa (4) es la contrapartida davidsoniana de la (1) de Quine; ambas defienden lo que podríamos llamar, con Alston, el *explicitismo* semántico (v. 2.3).

Así pues, dos son hasta ahora las ideas principales de Searle:

1. *La tesis de la indeterminación es paradójica*. Según Searle, la tesis viene a defender que no hay diferencia alguna de significado, ni tan siquiera para el emisor, entre expresiones estimulativamente sinónimas. Por tanto, puedo interpretarme a mí mismo tanto queriendo decir ‘estadio de conejo’ como queriendo decir ‘epifanía de la

conejeidad' cada vez que digo 'Conejo'. Pero, la paradoja aquí, recordemos, se halla en que todo el argumento precisa de que sepamos diferenciar entre los significados de 'Conejo', 'Estadio de conejo', etc. Si no somos capaces de apreciar ninguna diferencia entre esos significados, entonces éste resulta para nosotros un argumento completamente vacuo. En palabras de Searle:

Si la tesis de la indeterminación fuera cierta, ni tan siquiera podríamos entender su formulación; ya que cuando se nos dijera que no hay "materia objetiva" acerca de la corrección de la traducción entre conejo y estado de conejo, para empezar, no habríamos sido capaces de percibir ninguna diferencia (objetivamente real) entre las dos expresiones castellanas. (Ibíd., p.181).

He aquí lo que podemos denominar, con Searle, una *paradoja pragmática*: para poder exponer la tesis hemos de especificar distinciones que, según la tesis, no pueden ser especificadas. Es importante observar que no se trata sólo de que Searle dé por hecho que hay una diferencia entre 'conejo' y 'estadio de conejo', es que es el propio argumento de Quine el que lo da por hecho, para punto y seguido, negar esta diferencia; en eso consiste la reducción al absurdo.

La cuestión, pues, se plantea de la siguiente manera: ¿Cómo entender, siquiera, la tesis de la indeterminación si la referencia de nuestras expresiones es pura «metafísica trascendental»? Vimos en 1. 6. someramente las ideas quineanas sobre la interiorización de la ontología, y ahora podemos apreciar una tensión en este sentido. Por un lado, Quine hace cábalas acerca de la ontología de nuestro discurso, y llega a afirmar que

nuestra mente se halla predispuesta de manera natural para hablar de cuerpos (somos «body-minded», Quine 1974, p. 74). Pero, por otro lado, la tesis de la inescrutabilidad niega que los hablantes hagan referencia a algún tipo de entidad concreta, incluidos los cuerpos.

2. *La premisa culpable de la autorefutación es la del conductismo lingüístico (o mejor, explicitismo semántico)*. Sin asumir éste, obtendríamos, a lo sumo, la vieja infradeterminación de toda construcción teórica:

Es sólo *asumiendo* la no-existencia de significados intencionales [*intentionalistic meanings*] que el argumento de la indeterminación tiene éxito. (...) En cuanto a los significados construidos psicológicamente, se da la familiar infradeterminación de las hipótesis por la evidencia, y esa infradeterminación se añade a la infradeterminación al nivel de las partículas físicas o la conducta física bruta. ¿Y qué? Éstos son puntos familiares sobre cualquier teoría psicológica. (*Ibid.*, pp. 129-130).

El que la tesis de la indeterminación se sostiene sobre el conductismo lingüístico es una idea refrendada por el propio Quine: «Los críticos [con la tesis] han dicho que la tesis es una consecuencia de mi conductismo. Algunos han dicho que es una *reductio ad absurdum* de mi conductismo. Discrepo del segundo punto, pero estoy de acuerdo con el primero.» (Quine 1987, p. 5). Efectivamente, Quine basa su defensa de la indeterminación en su conductismo lingüístico, pero no así Davidson (v. 1.5). Y, efectivamente, se trata de dos cuestiones independientes (v. 2.5).

2.1.1. Relatividad ontológica y valor literal

Si el argumento de la indeterminación fuera válido, Searle cree que resultaría que no hay diferencia alguna entre un conejo o un estadio de la conejeidad. Recordemos que no es necesario el recurso a una lengua extraña, pues, como dice Searle exponiendo la tesis quineana, «la lengua materna es la madre de la indeterminación». En palabras de Quine: «Si ha de tener algún sentido decir incluso de uno mismo que se está refiriendo a conejos y fórmulas y no a estados de conejos y a números de Gödel, entonces debería tener sentido igualmente decirlo de otro» (Quine 1969a, p. 47). Volviendo a citar a Searle:

Si el argumento es válido, entonces debe tener el resultado de que no hay ninguna diferencia *para mí* entre *querer decir* conejo o estadio de conejo, y esto tiene, a su vez, el resultado de que no hay ninguna diferencia para mí entre *hacer referencia a* un conejo y hacerla a un estadio de conejo, y que no hay ninguna diferencia para mí entre que algo *sea* un conejo y *sea* un estadio de conejo. (Searle 1987, p. 130).

Y es, antes que Searle, el propio Quine quien se percata de que el argumento nos está llevando a consecuencias absurdas:

Parece que nos estamos dirigiendo a la absurda posición de que no hay diferencia alguna, interlingüística o intralingüística, objetiva o subjetiva, entre referirse a conejos y referirse a partes o estados de conejos; o entre referirse a fórmulas y referirse a sus números de Gödel. Con toda seguridad, esto es absurdo, ya que implicaría que no hay diferencia entre el conejo y cada una de sus partes o estados, y que no hay diferencia entre una fórmula y su número de Gödel. La referencia parecería ahora haberse vuelto

una absurdidad no solo en la traducción radical sino en casa. (Quine 1969a, pp. 47-48).

Quine afronta esta crítica aduciendo la relatividad de la ontología. Ésta nos permitiría afirmar que ‘conejo’ significa conejo pero sólo relativamente a un esquema de interpretación ya establecido (pues la estimulación que ha de contar como la referencia de ‘conejo’ no es más un conejo que un estado de conejo). Pero esta respuesta no resulta satisfactoria. En primer lugar porque acaba admitiendo que la referencia es escrutable, aun cuando sólo lo sea relativamente a un esquema preestablecido de traducción o interpretación. Como dice Searle, no podemos defender ambas ideas simultáneamente: o bien hay materia objetiva para fijar la referencia – y, entonces, podríamos hablar, como mucho, de infradeterminación – , o no la hay – con lo que cabe hablar de indeterminación – .

Realmente, no lo podemos tener de las dos formas. No podemos, por una parte, insistir en un conductismo riguroso que implica que no hay materia objetiva [*fact of the matter*] y después, cuando tenemos problemas, apelar a una ingenua noción de lengua materna o lenguaje materno [*home language*] con palabras que tienen un valor literal [*face value*] además de su contenido empírico conductual. Si somos serios acerca de nuestro conductismo, la lengua materna es la madre de la indeterminación, y el valor literal es un engaño si sugiere que hay diferencias empíricas cuando de hecho no hay ninguna. (Searle 1987, p. 133)

Debemos descartar, pues, la apelación a la relatividad ontológica como una respuesta aceptable al problema; quizá, ni siquiera supone una respuesta:

Ahora, no responde a esta dificultad decir que podemos fijar el significado y la referencia haciendo una selección arbitraria de un manual de traducción. La arbitrariedad de la selección del manual de traducción es precisamente el problema, ya que es un reflejo de la arbitrariedad de la selección entre el conjunto original de hipótesis analíticas alternativas. La tesis de la relatividad de Quine no elimina el absurdo; simplemente lo reformula. (*Ibid.*, p. 132).

Searle subraya aquí que lo relativo es el hecho de que una palabra tenga significado pero no cuál sea su significado. Ésta es una distinción crucial. Si una palabra tiene significado, lo tiene sólo relativamente a una lengua (por ejemplo, ‘conejo’ lo tiene en castellano y no en alguna otra lengua), pero su significado no es relativo a la lengua, sino que es el que es.

Por supuesto, una palabra significa lo que significa sólo relativamente a una lengua de la que es parte, pero la misma relatividad de la *posesión* de significado presupone la no relatividad del *significado* poseído [...]

Alguien podría objetar que parece que estoy asumiendo el “mito del museo” que Quine combate, la idea de que existe una clase de entidades mentales llamadas “significados”. Pero mi idea es neutral entre las diversas teorías del significado. El significado puede ser una cuestión de ideas en la cabeza *à la* Hume, disposiciones de conducta *à la* Quine, usos de palabras *à la* Wittgenstein o capacidades intencionales *à la* yo. No importa aquí. Sea lo que sea el significado, necesitamos distinguir la tesis verdadera de que una

palabra tiene el significado particular que tiene solo relativamente a un lenguaje de la tesis falsa de que el mismo significado es relativo a una lengua. (*Ibid.*, p. 135).

2.1.2. Lenguaje de fondo y sistema coordinado

Así, Quine se ha percatado de que su argumento lo hacía deslizarse hacia «la absurda posición» de que no es que la referencia sea relativa, sino que es inexistente. De ahí que haya intentado rescatar la noción de referencia o, al menos, ofrecer una explicación de la misma sin entrar en conflicto con la idea de la indeterminación. Podemos distinguir en sus consideraciones dos vías diferentes de hacerlo. Por un lado, como ya hemos apuntado, apela a un ‘lenguaje de fondo’ [*background language*] y, por otro, a un ‘marco de referencia’ o ‘sistema coordinado’ [*coordinate system*]. Según la primera opción, no tendría sentido hablar de referencia en términos absolutos sino sólo según el metalenguaje que utilicemos para interpretar el lenguaje objeto. Supongamos que vamos a traducir el término ‘rabbit’ de un anglohablante. Escogemos un manual de traducción que correlaciona el término ‘rabbit’ con el de ‘conejo’, pero necesitamos saber a qué se refiere el término ‘conejo’ (v. 1.3). El metalenguaje habrá de ser también interpretado y, para ello, precisaremos de un nuevo metalenguaje, digamos un metalenguaje₁, que nos diga que el término ‘conejo’ del metalenguaje se refiere a conejos, pero, de nuevo, cada término del metalenguaje₁ debe ser interpretado, para lo cual precisaríamos un metalenguaje₂, con lo que se abre un regreso al infinito de metalenguajes.

Ya Davidson se había percatado de este problema en fecha tan temprana como 1979, señalando que la estrategia quineana de hacer la referencia y la ontología doblemente relativas, al manual de traducción y al lenguaje de fondo, no funciona:

Debido a que relativizar la referencia en el metalenguaje no puede concretar la referencia y la ontología para el lenguaje objeto, a menos que el predicado de referencia relativizado del metalenguaje tenga una semántica determinada, Quine ve una jerarquía infinita de teorías o lenguajes, cada uno de los cuales intenta (en vano) estabilizar el esquema de referencia del lenguaje para el que provee la teoría. (Davidson 1979, p. 233).

La conclusión de Davidson apunta a una idea central de Quine: «La relatividad ontológica no se sigue, ya que sugiere que, cuando se han realizado las suficientes decisiones, arbitrarias o como fuere, una referencia única es posible, contrariamente a nuestro argumento para la inescrutabilidad de la referencia.» (*ibid.*, p. 235). O, como afirma de forma algo más contundente: «A lo que me opongo es a la idea de que la referencia puede ser relativizada de manera tal que fije la ontología. Es la relatividad ontológica lo que no entiendo.» (*ibid.*, p. 230-231).

La idea de Quine ahora será que en la práctica detenemos este posible regreso al infinito de metalenguajes tomando las palabras del lenguaje objeto con su valor literal [*at face value*]. Pero, como ya hemos apuntado arriba, ¿qué hemos de entender, al menos según Quine, por el ‘valor literal’ de las palabras? Quine hace la siguiente reflexión:

... imaginémonos en casa con nuestro lenguaje, con todos sus predicados y sus mecanismos auxiliares. Este vocabulario incluye «conejo», «parte de conejo», «estadio de conejo», «fórmula», «número», «buey», «ganado»; y asimismo los predicados diádicos de identidad y diferencia, y otras partículas lógicas. En estos términos podemos decir con muchas palabras diferentes que esto es una fórmula y eso un número, esto un conejo y eso una parte de conejo, esto y eso el mismo conejo y esto y eso partes diferentes. *Sólo con esas palabras*. Esta red de términos y predicados y mecanismos auxiliares es... nuestro marco de referencia, o sistema coordinado [*coordinate system*]. Relativamente a él podemos y, de hecho, hablamos de forma significativa y distintiva de conejos y partes, números y fórmulas. (Quine 1969b, p. 48).

Así pues, para detener el regreso al infinito de los metalenguajes y para poder explicar qué significa considerar el significado de las palabras de forma literal, hemos pasado de la apelación a un lenguaje de fondo a la de un sistema coordinado o marco de referencia. Pero tampoco esta respuesta parece convincente. Es evidente que uno puede expresar los diferentes marcos en su lengua pero, de nuevo, esto cuenta más como una refutación del argumento quineano que como una defensa del mismo, ya que si podemos expresar la diferencia entre conejos y partes de conejos, es porque los términos tienen habitualmente una referencia concreta. Como expone esta crítica N. Georgalis:

Concedido que uno puede *decir* estas cosas, usar «sólo esas palabras», pero hacer eso no es de ninguna ayuda si esas palabras no tienen algún significado y referencia particular. Apelar a la locución ‘sólo esas palabras’ sólo servirá si las expresiones ‘conejo’ y ‘parte de conejo’, en castellano, tienen sus denotaciones usuales, esto es, conejos y partes de

conejo, entonces, y sólo entonces, pueden *esas palabras* ser usadas para expresar una diferencia *particular*. (Georgalis 2004, p. 246).

Y como añade algo más adelante, «dada la indeterminación de la traducción: las palabras no tienen ni significados ni referencias particulares, *las palabras no tienen valor literal*. Los diferentes tipos fonéticos o inscripcionales no son ni necesarios ni suficientes para asegurar diferencias de significado o referencia[.]» (p. 247). Efectivamente, es contradictorio afirmar por una parte que los términos no tienen una referencia determinada y, por otro, que para discusiones sobre la referencia de nuestros términos los tomamos con su referencia literal, ya que esta supuesta referencia literal es lo que acaba de ser negada.

Consideraciones acerca del aprendizaje del lenguaje hechas por el propio Quine también apoyan la idea sostenida aquí. Hemos visto que la inescrutabilidad de la referencia hace que la noción de «tomar las palabras literalmente» no tenga sentido. Georgalis refuerza esta crítica basándose en lo dicho por el propio Quine sobre lo que ha de contar como conocimiento del significado de una palabra. En principio, conocer una palabra ha de ser una condición tanto necesaria como suficiente para ser capaz de interpretarla literalmente. Quine defiende que podemos considerar que una palabra es conocida cuando se domina: (1) la parte fonética – estar familiarizado con el sonido y ser capaz de repetirlo – y (2) la parte semántica – saber utilizar la palabra – . (Quine 1969b, pp. 27-28). Estas condiciones se cumplen en la lengua materna: así, yo puedo pronunciar la palabra ‘conejo’ y sé cómo usarla. Pero la cuestión es que cuando

aprendemos una lengua hemos de percatarnos de cuáles son las palabras que están funcionando de cuantificadores, y para hacer esto hemos de interpretarlos – ya que la indeterminación también afecta a los cuantificadores – . Localizar los cuantificadores, pues, implica darles una interpretación, y esto implica fijar una referencia para los términos. Luego la característica (2), la que hace referencia a la semántica de los términos no puede explicarse apelando, exclusivamente, a rasgos sintácticos, ya que se requiere una interpretación, es decir, una asignación de referencias. Así que, de nuevo, el supuesto ‘valor literal’ de las palabras según Quine reduce al absurdo su tesis. Citando de nuevo la obra de Georgalis:

Ya que las condiciones estipuladas para conocer una palabra no son sensibles a estas diferentes construcciones [las diferentes interpretaciones posibles], no son suficientes para conocer una palabra. El alegado «valor literal» no tiene valor semántico en la visión de Quine. La discusión de Quine del «valor literal» de una palabra no resuelve el problema de la referencia – lo evade. (*Ibid.*, p. 249).

Quine ha ofrecido una respuesta a este problema apelando al desentrecomillado. Volvemos sobre ello en 2. 7.

2.2. Primera persona y conductismo

¿Por qué el argumento de la indeterminación acaba constituyendo una reducción al absurdo de sí mismo según Searle? ¿Por qué éste se «siente tan seguro» [*feel so confident*] de que la premisa que se ha de rechazar en el argumento es la del conductismo semántico (v. 2.1)? Searle ha puesto en estrecha relación su crítica acerca

de la reducción al absurdo del argumento de la indeterminación con su defensa de la asimetría entre la primera y la tercera persona. Es decir, si el argumento supone una reducción al absurdo de sí mismo es debido al rechazo de dicha asimetría. Esta es la forma exacta en la que Searle vincula la supuesta *reductio* del argumento con la apelación a la primera persona:

[S]i el conductismo fuera correcto, tendría que ser correcto para nosotros como hablantes de español tanto como para hablantes del gavagaiense. Y nosotros sabemos *por nuestro propio caso* que con ‘conejo’ queremos decir algo diferente de ‘estadio de conejo’ o ‘parte de conejo no separada’ [*And we know from our own case that we do mean by ‘rabbit’ something different from ‘rabbit stage or ‘undetached rabbit part’.*] Si mi vecino castellano hablante, habiendo leído a Quine, decide que no puede decir si con ‘conejo’ yo quiero decir conejo, parte no separada de conejo, o estadio de conejo, entonces, tanto peor para él. Cuando vi un conejo hace poco, como sucedió en verdad, y lo llamé un conejo, quise decir conejo [*I meant rabbit*]. En todas las discusiones en la filosofía del lenguaje y la filosofía de la mente, es absolutamente esencial en algún punto recordarse a sí mismo el caso de la primera persona. (Searle *op. cit.*, p. 126.)

El argumento de Searle aquí, resumidamente, es el siguiente: a) El conductismo semántico genera la indeterminación de la traducción; b) La indeterminación de la traducción implica la del significado; c) La del significado rige también en el ‘caso doméstico’; d) Sabemos que el caso doméstico es falso por la autoridad de la primera persona; luego e) la autoridad de la primera persona refuta el conductismo semántico y, con él, la indeterminación.

De hecho, Searle afirma explícitamente que si la indeterminación constituye un problema, se debe sólo al conductismo y verificacionismo: «Creo que sólo parece haber un problema filosófico sobre la traducción porque aún estamos sufriendo los malos efectos del viejo conductismo y verificacionismo.» (Searle 1997, p. 3). Y continúa insistiendo en su idea de que es esta premisa conductista la que aboca el argumento a la autorefutación:

Si, con todo, estás tentado a pensar que no puede haber hechos sobre el significado que no sean hechos sobre la conducta manifiesta [*overt behavior*], entonces es posible producir una *reductio ad absurdum* de tu conductismo, porque ves que no puedes distinguir, sólo con la conducta, entre querer decir conejo y querer decir, por ejemplo, partes no separadas de un conejo. Quine produjo una *reductio ad absurdum* de su conductismo hace muchos años con su argumento de la «indeterminación de la traducción». (*Ibid.*, p. 3).

Y, de nuevo, se apela a la autoridad de la primera persona como alternativa al conductismo quineano:

En este caso, puedes, con Quine, mantenerte firme en el conductismo pase lo que pase, e insistir en que no hay materia objetiva sobre si cuando digo ‘conejo’ quiero decir conejo o partes no separadas de un conejo. O puedes, conmigo, reconocer lo que yo tomo como los hechos obvios de que hay una distinción clara para mí entre querer decir conejo, y querer decir partes no separadas de un conejo, y que cualquier explicación conductista del significado que no pueda reconocer esos hechos es, por tanto, falsa. (*Ibid.*, p. 4).

O, por utilizar las palabras de otro autor de línea searleana:

Incluso sin exagerar, el lugar que el argumento de la indeterminación nos deja es lo suficientemente incómodo, ya que el argumento, asumiendo que nos lo podemos aplicar a nosotros de igual manera que a los hablantes de la Jungla, nos dice que no hay materia objetiva acerca de si cuando uso la palabra ‘conejo’ me refiro a conejos o a partes de conejo. Sin embargo, es evidente para mí [*it's plainly evident for me*] que cuando uso la palabra ‘conejo’ estoy hablando de conejos y de nada más. (V. McGee 2005, p. 402).

William P. Alston ha realizado también algunas reflexiones en esta misma línea.

Comienza ofreciendo la siguiente esquematización del argumento de Quine:

- (1) Habrá varias traducciones alternativas de la expresión E_1 de un lenguaje L_1 a expresiones de un lenguaje L_2 , todas las cuales explicarán perfectamente bien las disposiciones verbales de todos los hablantes, así como cualesquiera otros datos observacionales relevantes.
- (2) Por tanto, no puede haber razón suficiente para considerar una de estas traducciones como la correcta.
- (3) Por tanto, no hay traducción objetivamente correcta.
- (4) Por tanto, es imposible determinar lo que significa E_1 .

(5) Por tanto, no hay hechos acerca de lo que significa E_1 .

(6) Por tanto, E_1 no significa nada determinado.

Su crítica se dirige ahora al paso de (3) a (4):

Reflexionemos un momento sobre por qué parece tan absurdo decir que porque, irremediablemente, seamos incapaces de conseguir una traducción única de 'gavagai' observando la conducta del nativo y sus circunstancias, demos concluir que el nativo no quiere decir nada definido con 'gavagai', y, por extensión, que nosotros no queremos decir nada definido con 'conejo'. (Alston 1982, p. 59).

Y, como Searle, Alston apela a la primera persona:

Claramente, es porque parece obvio que *yo* sé lo que quiero decir con 'conejo' y otras palabras en mi lenguaje. *Yo* sé que, por ejemplo, uso 'conejo' para denotar organismos completos duraderos como ese, y no las partes o los estados de tales organismos o las clases a las que pertenecen. Mi seguridad respecto a lo que quiero decir con 'conejo' no se basa en lo que yo u otra persona podemos hacer al traducir de una lengua a otra, y mucho menos en lo que es posible hacer en la traducción radical. (*Ibid.*, p. 60. Dos notas al pie omitidas).

La apelación a la primera persona es, pues, idéntica a la de Searle: si utilizando las funciones vicarias hábilmente otro hablante me interpreta como refiriéndome a

epifanías de conceptos platónicos y no a objetos físicos persistentes, peor para él, porque yo sé que no me estoy refiriendo a epifanías. La crítica del paso (3) al (4) se basa en la concepción de la expresión «traducción objetivamente correcta». En el argumento quineano-davidsoniano, esto no implica más que concordancia con la conducta *visible* de los hablantes (patrones de *a*/disentimiento); es desde este punto de vista desde el que Alston acepta el paso que lleva de (2) a (3). Sin embargo, impugna el de (3) a (4): Aun cuando se haya admitido que no hay razones para tratar una traducción como *la* traducción correcta (en el sentido antes explicado), sino que varias resultan igualmente correctas, esto no implicaría que cualquiera de esas traducciones puede contar como el significado de la expresión, ya que para determinar el «significado correcto» se ha de tener en cuenta la perspectiva de la primera persona; el *non sequitur*, por tanto, se da de (3) a (4).

Con todo, no se ve claro por qué Alston no acepta el sentido conductista de «significado correcto» pero sí el de «traducción correcta». Es decir, si se puede admitir, aunque sea por mor del argumento, que una traducción es correcta siempre que se acople a los datos de la conducta observable de los hablantes, ¿por qué no aplicar ese mismo criterio para uno mismo? ¿Por qué no impugnar ya el paso entre (1) y (2)? Alston parece detenerse en la premisa (4) porque el significado depende de algo más que de la conducta observable, sin embargo, parece olvidar que traducir consiste en emparejar oraciones con el mismo significado. Para ser coherente con el argumento, Alston debería afirmar aquí que del hecho de que varias traducciones se acoplen perfectamente bien a las disposiciones lingüísticas de los hablantes, no se sigue que

cualquiera de esas traducciones sea correcta. *La* traducción correcta sería, según Alston, aquella que recoge lo que el hablante quiere decir; no parece tener mucho sentido aceptar la indeterminación de la traducción pero no la del significado: ambas son la misma.

Sin embargo, hay un sentido en el que se puede explicar el rechazo de Alston al paso de (3) a (4), y no de (1) a (2). Recordemos que la primera premisa añade a la conducta verbal de los hablantes, como información a tener en cuenta, «otros datos observacionales relevantes», y, por supuesto, todo depende de cuáles consideramos que son estos datos... aunque, eso sí, han de ser «observacionales». De nuevo, como vimos con Searle, para salvar el *non sequitur* entre (1) y (2) necesitamos una premisa adicional: (1') Para que una traducción sea correcta es condición suficiente el que se acople a la conducta verbal. Y lo que Alston estaría defendiendo aquí es que en el caso de la primera persona disponemos de datos relevantes no observacionales.

La cuestión ahora es, pues, si se ha de rechazar (1'). Esto es, si el acoplamiento a la conducta verbal es una condición suficiente de la corrección de una traducción.

B. C. Smith ha desarrollado una crítica a Davidson en la misma línea que Searle y Alston. Smith sintetiza el mecanismo de interpretación de otra persona en dos premisas:

I. Considero que tú quieres decir S con tu uso de P.

II. Tú quieres decir S con tu uso de P.

El paso de I a II sólo es legítimo apoyándonos en una premisa implícita:

III. S es lo que cualquiera que usas P quiere decir con ello.

Y aquí es donde se apela a la primera persona. Smith defiende que si interpretamos a otros como queriendo decir S con P se debe a que uno quiere decir S con P y presupone que los demás hacen lo mismo. Así, se dice:

Pero lo que hay que ver es que nuestra aceptación de II es automática más que respaldada por la confirmación empírica. (...) Sin esta presuposición (al menos inicialmente), nuestras prácticas comunicativas no tendrían éxito. Estamos constituidos de tal manera que comenzamos considerando que todo el mundo es como nosotros a este respecto: se trata de una perspectiva tan básica que ni tan siquiera se halla apoyada por alguna asunción o inferencia. Esto es lo que hacemos al comienzo como norma. (B. C. Smith 2006, p. 402).

La idea de Smith, como la de Searle, es basar el proceso interpretativo en (i) el conocimiento privilegiado, directo, de los significados de mis palabras y (ii) la analogía, esto es, considerar que los demás quieren decir con sus palabras lo mismo que yo cuando utilizo esas mismas palabras:

En una situación real en la que entiendo a otro, quiero proponer que el caso de la primera persona – saber lo que quiero decir con mis palabras – es básico, y que me baso

en este conocimiento para saber lo que *tú* quieres decir. Esto resulta crucial en el orden de adquisición:

(A) Tomo mis palabras como significando tal y tal.

(B) Al oír las palabras que pronuncias, las tomo como queriendo decir lo que yo quiero decir con ellas. (*Ibid.*, 401).

Son varios los autores que se han sentido tentados por la apelación a la analogía; sería ésta, pues, la que nos permite determinar la traducción más allá de las condiciones de verdad:

[L]as similitudes psicológicas que podemos esperar encontrar entre nosotros y la gente en nuestra comunidad inmediata será suficiente para asegurar que traducir el «conejo» de nuestro vecino por «conejo» captará su significado mejor que traduciéndolo por «parte de conejo no separada.» (V. McGee 2005, p. 402).

Se debe recordar aquí que la cuestión que la tesis de la indeterminación plantea no es que no se pueda saber con certeza la referencia de los términos de otra persona, sino que no hay nada que saber más allá de las condiciones de verdad; no se trata de inaccesibilidad, sino de inexistencia. En todo caso, dada esta crítica a la tesis, la cuestión sería ahora cómo consigue la primera persona determinar el significado.

2.3. Primera persona y explicitismo

Quine parte de la premisa mencionada en el argumento de Alston porque, para él, aprender una lengua consiste en adquirir disposiciones verbales. Alston esquematiza ahora esta argumentación:

- (1) Al adquirir una lengua, uno no tiene nada más para proceder que la conducta verbal de los otros y los rasgos observables de la situación de la emisión.

- (2) Por tanto, todo lo que uno puede llegar a saber, al aprender una lengua, es qué emisiones están dispuestos a hacer los hablantes de una lengua ante qué circunstancias.

- (3) Por tanto, en virtud de haber aprendido una lengua uno no tiene conocimiento de significados o nada que vaya más allá del conocimiento de disposiciones verbales.

Alston apunta ahora que la transición espuria se halla entre (1) y (2):

La transición vulnerable es de (1) a (2). Ahí Quine hace una asunción fuerte y cuestionable sobre lo que los seres humanos pueden llegar a saber a partir de un cuerpo dado de datos. Se asume que lo más que se puede aprender de observar la conducta bajo ciertas circunstancias es un conjunto de verdades sobre disposiciones, la actualización de las cuales comporta conducta de ese tipo en circunstancias de ese tipo. Esto es, asume la

imposibilidad de ir, de manera justificada, de los datos observados a un sistema de hipótesis explicatorias que postulan entidades bastante diferentes de las observadas y que atribuyen a esas entidades rasgos bastante diferentes de esas entidades observadas. Esto es, asume que nada como lo que sucede en las ciencias más altamente desarrolladas tiene lugar al aprender una lengua. En ausencia de más razones para hacer esta asunción creo que puede ser tranquilamente abandonada. (Alston 1982, p. 65).

Aprender una lengua, subraya Alston, supone más que aprender disposiciones conductuales o condiciones de verdad de oraciones, aprendemos a imputar creencias e intenciones y, probablemente, utilizamos evidencia no quineana: «Mucho de nuestro conocimiento es adquirido mediante procesos implícitos inconscientes que no somos capaces de hacer enteramente explícitos» (*ibíd.*, p. 67).

Según esta visión, el error de Quine y Davidson ha consistido en extraer conclusiones ontológicas acerca de los significados a partir de una metodología epistemológica – la perspectiva de la tercera persona – , que, además, se ha asumido como premisa, sin justificación alguna o, al menos, suficiente. Así, se ha llegado a conclusiones acerca de cómo debe ser el significado para que pueda ser conocido exclusivamente desde una perspectiva de la tercera persona. En esta misma línea dice Barry C. Smith:

De esta manera, se extraen conclusiones metafísicas a partir de premisas epistemológicas. Tanto para Davidson como para Quine, la epistemología de la tercera persona del significado conduce a la afirmación de que el significado es un fenómeno esencialmente público y social. (B. C. Smith 2006, p. 391).

La postura defendida por Quine en la primera premisa de este argumento es bautizada por Alston como *explicitismo*, una postura de «consíguelo-todo-de-fuera»:

Un aspecto de la posición de Quine es lo que podría ser llamado Explicitismo, una actitud de consíguelo-todo-de-fuera [*a get-it-all-out-in-the-open attitude*]. Recordemos que el principio básico de Quine es que no puede haber hechos que no podrían ser descubiertos por el uso de un método examinador de hipótesis empíricas en el que todo está consciente y explícitamente formulado: datos observados, hipótesis alternativas, teorías de fondo, razones para considerar una alternativa superior a otra, etc. (Alston 1982, p. 67).

La conclusión de Alston es, pues, similar a la que vimos en Searle y a la que defendía una semántica pre-quineana: saber una lengua tiene que implicar saber el significado de sus oraciones y la referencia de sus términos (esto no implica, puntualiza Alston, que los hablantes sean capaces de hacer explícito lo que quieren decir: probablemente no tendrán otro término que ‘conejo’ para formular esa idea.). El explicitismo no sería más que el punto fundamental del conductismo semántico: la premisa de que todo rasgo semántico ha de aparecer explícito en la conducta observable de los hablantes.

De hecho, aunque Davidson nunca haya concedido que durante el proceso de interpretación pudiéramos hacer uso de información de alguna manera implícita, sí que es cierto que ha puesto énfasis en que las condiciones sobre las que ha de operar la interpretación impiden que la indeterminación tenga demasiado alcance. En palabras de Ramberg:

Tanto Davidson como Lewis toman la presencia de resistencia intuitiva a la indeterminación empírica entre teorías concretas como una indicación de que en la práctica tenemos más condiciones [*constraints*] sobre nuestra interpretación de personas de las que son captadas por las racionalizaciones teóricas. Cuando haya un reconocimiento total de todas las condiciones empíricas, las maneras alternativas de acomodar las anomalías no será preocupante. Si sentimos que algo se ha perdido entre teorías empíricamente equivalentes, esto significa o bien que hay más condiciones por descubrir o que la teoría del significado empírico-holista es falsa. (Ramberg 1989, p. 94).

Ahora bien, si la tesis comienza a predicar condiciones de las teorías interpretativas no descubiertas, y sobre cuya naturaleza nada se dice, que mitigan o eliminan la indeterminación, la tesis comienza también a perder gran parte de su atractivo. Por no decir que comienza a recular. Lewis, por ejemplo, afirma:

Credo: si llegas a probarme que todas las condiciones que hemos hallado hasta ahora podrían permitir dos soluciones perfectas, difiriendo sólo en el aparato auxiliar de M [la teoría de la verdad], entonces me habrás probado que aún no hemos hallado todas las condiciones. (Lewis 1974, p. 343. Citado de Ramberg 1989, p. 94).

Aunque aquí parece haber una pugna más entre convicciones que entre demostraciones. Por un lado, la línea de crítica searleano-alstoniana parece insistir en que la autoridad de la primera persona es el hecho primario:

Nadie, por ejemplo, nos puede convencer con argumentos, por ingeniosos que sean, de que los dolores no existen si, de hecho, los tenemos, y consideraciones similares se aplican al ejemplo de Quine. Si alguien tiene una teoría según la cual no hay diferencia alguna entre que yo quiera decir conejo o quiera decir parte de conejo, entonces sé que su teoría está simplemente equivocada; y el único interés que su teoría puede tener para mí consiste en intentar descubrir dónde se equivocó. (Searle 1987, pp. 126-127).

O, en respuesta a una pregunta de Føllesdal:

Si alguien me pregunta cómo sé que 'lapin' significa 'conejo', sólo puedo decir que he hablado francés durante unos cuarenta años, y sé francés lo suficientemente bien para saber que esto es lo que quiero decir cuando uso esta palabra, y que comparto este significado con otros hablantes de francés. (Searle 1997, p. 3).

Añadamos aquí que a pesar del rechazo explícito de Searle del mito del museo, esta respuesta parece comprometerse con él. Searle no sólo defiende aquí la autoridad de la primera persona – que, como hemos visto, hace en base a un 'contenido Intencional' –, sino que afirma conocer también lo que quieren decir los demás hablantes, y esto en base al significado convencional de las palabras (v. sobre esto 3.3.2).

Así, dado que la autoridad de la primera persona implica en esta concepción la existencia de datos relevantes para la semántica no observacionales, deberíamos rechazar el *explicitismo*, que constituía la premisa fundamental del argumento de la indeterminación. Así, si queda abierta la posibilidad de confeccionar diferentes manuales de traducción para un hablante, se debería a que no hemos tomado en

consideración todos los datos relevantes; o, como decía Lewis, no hemos hallado todas las condiciones que se han de imponer sobre una interpretación.

Por otro lado, Lewis no ofrece ninguna razón en apoyo de su idea: «Pero esto se ofrece como un “credo”, no como un argumento. Quine puede insistir en que ningún conjunto justificado de condiciones asegurará que incluso la totalidad de las verdades físicas sólo deja espacio para un esquema de interpretación correcto» (Kirk 2004, p. 173). Lo mismo se puede afirmar de las ideas expuestas por Searle y Alston: ¿por qué tratar cómo un hecho primario la autoridad de la primera persona y no el explicitismo semántico? De hecho, el argumento de la indeterminación, especialmente una vez que se aplica al caso doméstico, está pensado como una reducción al absurdo de la perspectiva de la primera persona en cuestiones semánticas.

Volvámonos ahora, con Searle, hacia Davidson. Como hemos visto, éste ha desarrollado una teoría del significado basada en la de Quine. En la reconstrucción davidsoniana del aprendizaje del lenguaje se apela a condiciones en el mundo (estímulos distales), emisiones y la actitud psicológica de considerar verdadera una oración. Davidson acepta la inescrutabilidad quineana, ya que, siendo la unidad del test empírico la oración y no las partes de ésta, siempre se podrán hallar esquemas diferentes de interpretar las oraciones a partir de las interpretaciones diferentes que demos de sus partes que sean consistentes con los mismos hechos acerca de qué oraciones considera verdaderas un hablante dado y en qué condiciones. La idea básica, como en Quine, vuelve a ser que hay diferentes formas de emparejar entidades con palabras que darán

perfecta cuenta de la conducta verbal de los hablantes. El argumento de Davidson se desarrolla, pues, así:

- 1) La unidad del análisis empírico en la interpretación radical es la oración (no los elementos suboracionales).
- 2) La única evidencia empírica para la interpretación radical es el hecho de que los hablantes «consideran verdaderas» ciertas oraciones en ciertas ocasiones.
- 3) Hay formas alternativas e inconsistentes de emparejar palabras con objetos que explican igualmente bien por qué un hablante considera verdadera una oración.

Como vimos, incluso con la aceptación del papel esencial de creencias e intenciones para el lenguaje y del uso constante del principio de caridad, Davidson acepta la indeterminación de la traducción y la inescrutabilidad del significado. Con todo, lo que asombra a Searle es que de aquí no obtenemos inescrutabilidad alguna. Para ello, Davidson debe añadir una premisa adicional acerca del carácter de su teoría:

- 4) Todos los hechos semánticos deben ser públicamente accesibles, tanto para el hablante como para el intérprete. Si el intérprete no puede hacer una distinción sobre la base de la evidencia pública y empírica, entonces es que no hay distinción alguna que hacer.

Uno de los ejemplos de Davidson nos invita a que supongamos que todo tiene una sombra, con lo que en una circunstancia en la que un hablante considerara verdadera la oración ‘Wilt es alto’, podríamos interpretar ‘Wilt’ como refiriéndose a Wilt y ‘es alto’ como haciéndolo a las cosas altas o bien, podríamos interpretar ‘Wilt’ como refiriéndose a la sombra de Wilt y ‘es alto’ como haciéndolo a la sombra de las cosas altas. El primer esquema de interpretación afirmaría que ‘Wilt es alto’ es verdad si y sólo si Wilt es alto. El segundo afirmaría que ‘Wilt es alto’ es verdad si y sólo si la sombra de Wilt es la sombra de algo alto. El sentido común nos dice que cuando digo ‘Wilt es alto’, con ‘Wilt’ me estoy refiriendo a Wilt y con ‘es alto’ a las cosas altas. Cuando digo ‘Wilt’, por tanto, no me refiero ni explícita ni implícitamente a las sombras, e igualmente para ‘es alto’. Ahora bien, eso es lo que dice el sentido común, ya que éste considera que esas afirmaciones son evidentes *para mí*, aunque puedan no serlo para otro que no consigue interpretarme correctamente. No se trata de hipótesis analíticas diseñadas para explicar mi conducta o mis «actitudes veritativas» (*hold-true attitudes*), se trata de que cuando hablo yo sé a qué me refiero, esto es, conozco más datos que meramente ante qué circunstancias considero verdadera una oración. Como dice Searle: «Por decirlo crudamente: además, yo sé lo que quiero decir».

Pues bien, para Davidson – como para Quine – tiene el mismo sentido considerar que con ‘Wilt es alto’ yo quería decir que Wilt es alto como que quería decir que la sombra de Wilt es la sombra de una cosa alta. Esta es la consecuencia de defender a ultranza el carácter público de la semántica. Ahora bien, el argumento davidsoniano, idéntico en esencia al quineano, comete la misma reducción al absurdo: para entenderlo debemos

asociar una referencia diferente en cada paso – en un caso de objetos, en otro de sombras de objetos – pero, en todo caso, hemos de asociar una referencia determinada y percatarnos, además, de que se trata de referencias diferentes. Lo que se habría mostrado, de nuevo, es la posibilidad de utilizar diferentes marcos interpretativos compatibles con la conducta verbal del hablante, pero no que los hablantes no asocien a sus términos una referencia concreta.

También la teoría de Davidson, por tanto, resulta ser su propia reducción al absurdo. O, al menos, la reducción de la premisa adicional que se vio obligado a añadir para obtener la inescrutabilidad: la premisa “explicitista”. Esto es, la premisa que defiende que ya que el lenguaje es un arte público, todos los hechos acerca de los significados deben ser hechos públicos. Expuesto según el propio Davidson:

Quizás alguien (Quine no) estará tentado a decir, ‘Pero al menos el hablante sabe a lo que se está refiriendo’. Hemos de mantenernos firmes contra esta idea. Los rasgos semánticos del lenguaje son rasgos públicos. Lo que, en la naturaleza del caso, nadie puede descubrir a partir de la totalidad de la evidencia relevante no puede ser una parte del significado. Y ya que cada hablante debe saber esto, al menos en algún vago sentido [*in some dim sense*], no puede ni tan siquiera intentar usar sus palabras con una referencia única, ya que sabe que no hay manera de que sus palabras le comuniquen a otro la referencia. (Davidson 1979, p. 235).

Searle se pregunta si esta respuesta no empeora la cosa. Para Davidson, ahora, debemos saber «en algún vago sentido», al menos, que la inescrutabilidad es inevitable

y, por tanto, ni tan siquiera nos esforzamos en transmitir una referencia única. O sea que ni tan siquiera puedo pretender referirme a Wilt y no a su sombra, a conejos y no a sus partes. Aquí halla de nuevo perfecta aplicación la paradoja pragmática: el argumento hace distinciones que según él mismo, sencillamente, no existen.

Davidson ha argumentado repetidamente a favor de la inescrutabilidad de la referencia.

En uno de sus últimos escritos dice:

La tesis de la inescrutabilidad de la referencia dice que no hay manera de decir qué forma de conectar palabras y cosas es la manera correcta; si una manera funciona, habrá incontables otras que también lo hacen. Desde un punto de vista técnico, esto significa que podemos sustituir la relación estándar de satisfacción (la satisfacción es una forma sofisticada de referencia) por incontables relaciones sin alterar las condiciones de verdad de ninguna oración o las relaciones lógicas entre las oraciones. Ya que toda la evidencia para interpretar el lenguaje ha de darse en el nivel oracional (ya que sólo las oraciones tienen un uso en la comunicación), el resultado es que no puede haber evidencia para afirmar que una de las relaciones de satisfacción (o referencia) es la correcta. (Davidson 1997, p. 78).

Davidson pasa a ofrecer otro curioso ejemplo, similar al de Wilt y su sombra. Supongamos una relación de satisfacción s que mapea 'Roma' en Roma, y el predicado 'ser una ciudad de Italia' en ciudades que están en Italia. Esta interpretación nos ofrecería como condición de satisfacción de la oración 'Roma es una ciudad de Italia' el que Roma sea una ciudad de Italia. Ahora consideremos una relación de satisfacción s' que mapea el término 'Roma' en un punto situado a 100 kilómetros al sur de Roma, y el

predicado ‘ser una ciudad de Italia’ en un punto que está a 100 kilómetros al sur de una ciudad de Italia. Ahora, el nuevo mapeo nos ofrecería como condiciones de satisfacción de la oración ‘Roma es una ciudad de Italia’ el que un punto situado a 100 kilómetros al sur de Roma es un punto situado a 100 kilómetros al sur de una ciudad de Italia. Las condiciones de satisfacción son claramente equivalentes. La conclusión, según Davidson, suena así:

La tesis de la inescrutabilidad de la referencia defiende que no puede haber evidencia para determinar que *s* es en algún aspecto mejor que *s'* para interpretar la oración ‘Roma es una ciudad de Italia’. No hay manera de decir *sobre qué* es una oración o qué está pensando alguien. (Davidson 1997, p. 78).

Pero volvemos a la problemática anterior: si *s* y *s'* son igualmente admisibles se debe a que la oración en cuestión, considerada holofrásticamente, mantiene sus condiciones de verdad, a pesar de que Roma y un punto a 100 km. de Roma, no son, obviamente, la misma entidad. El argumento, pues, es que no hay ninguna diferencia empírica entre ambas interpretaciones. En palabras de Davidson:

Con toda seguridad, Roma y un área 100 millas al sur no son la misma entidad. Esto es cierto. Con todo, cualesquiera diferencias que convencionalmente concebimos como diferencias en la referencia de los nombres o de las extensiones de los predicados serán preservadas por cualquier relación correcta de satisfacción. El hecho de que no haya ninguna diferencia empírica entre las interpretaciones ofrecidas por *s* y *s'* no conlleva que la persona que es interpretada no pueda distinguir entre Roma y un área 100 kilómetros al sur; pero sí que conlleva que no se puede decir a cuál se refiere su palabra ‘Roma’. Una

interpretación correcta halla [*keeps track*] un patrón complejo, y localiza oraciones y actitudes particulares dentro de él. Pero los medios que utilizamos para representar este patrón, esto es, nuestras oraciones, pueden representarlo de más de una manera.

(Davidson 1997, p. 79.)

Ya que el propio Davidson basa explícitamente su argumento en que no haya «diferencia empírica» entre una interpretación u otra, Searle dirigirá ahora su crítica hacia la concepción quineano-davidsoniana de «dato empírico».

Cabría distinguir dos sentidos de «empírico»: por un lado puede significar apto para ser testado por una tercera persona y, por otro lado, real, de hecho. Las teorías de Davidson y Quine han confundido ambos sentidos, induciendo a la falsa opinión de que lo que no es testable por una tercera persona no es real. (Así por ejemplo, cuando Davidson dice: «Lo que un intérprete no puede decidir a partir de datos empíricos acerca de la referencia de las palabras de otro, no puede ser un rasgo empírico de esas palabras», en realidad no está diciendo más que una tautología: lo que no es empírico no es empírico, aunque él quería decir otra cosa: lo que no puede ser testado por una tercera persona no puede ser un rasgo real del significado). Abandonemos ese recelo injustificado. Apelar a lo que es accesible sólo a la primera persona es apelar a algo tan real – y, por tanto, empírico en el segundo sentido mencionado – como hacerlo a lo que también lo es a la tercera.

Es parte de la tendencia persistentemente objetivizadora [*objectivizing*] de la filosofía y la ciencia desde el siglo diecisiete el que consideremos el punto de vista objetivo de la

tercera persona como preferible a, como, de alguna manera, más “empírico” que el punto de vista “subjetivo” de la primera persona. Lo que parece, pues, una simple declaración de hecho científico – que el lenguaje es una cuestión de estimulaciones de las terminaciones nerviosas – resulta ser, al examinarla, la expresión de una preferencia metafísica y, creo, una preferencia que no está garantizada por los hechos. El hecho crucial en cuestión es que realizar actos de habla – y querer decir cosas con emisiones – sucede a un nivel de intencionalidad intrínseca de la primera persona. El conductismo de Quine está motivado por una metafísica profundamente antimentalista que hace parecer el análisis conductista el único análisis que es científicamente respetable. (Searle1987, p. 145).

2.4. Acceso privilegiado

La crítica de Searle y Alston que venimos siguiendo defiende, pues, la necesidad de aceptar una asimetría entre la primera y la tercera persona. La cuestión, ahora, se centraría en hallar una caracterización satisfactoria de ésta.

La asimetría entre la primera y la tercera persona consiste, básicamente, en una asimetría entre la evidencia públicamente disponible acerca de lo que un hablante dice y lo que él realmente quiere decir. Este nuevo planteamiento no eliminaría la infradeterminación, pero ésta no implicaría más que la conocida infradeterminación de todas las teorías científicas, incluidas las psicológicas, o sea, las que hacen afirmaciones acerca de otras mentes, aplicada ahora a la semántica. Como se vio en 1.8, es posible defender la infradeterminación de las teorías científicas y la indeterminación del significado mientras no nos veamos obligados a reconocer una materia objetiva para la

traducción aparte de los meros patrones de asentimiento/disentimiento de los hablantes. Como dice Georgalis en una vena muy searleana:

Desde una perspectiva de la primera persona sabemos que se puede determinar una interpretación particular en el caso de uno mismo, la intencionada por el agente, aunque ésta sea invisible desde una perspectiva estrictamente de la tercera persona del agente. Desde esta última perspectiva, podría ser cualquiera de las alternativas permisibles establecidas por Quine. (Georgalis 2004, p. 247).

Davidson comienza distinguiendo dos tipos de asimetría: en primer lugar hallaríamos la diferencia entre adscripciones de una actitud proposicional hechas por un individuo a sí mismo o las que haga una segunda o tercera persona; y, por otro lado, la diferencia entre el criterio de primera persona para decir que su adscripción a sí misma es correcta y el de una segunda o tercera persona. En el primer caso estaríamos discutiendo si la oración de Juan ‘creo que Peter Pan es infantil’ y la de otra persona ‘Juan cree que Peter Pan es infantil’ significan lo mismo. En el segundo caso estaríamos discutiendo si la justificación de Juan para pensar que ha dicho algo verdadero al decir ‘creo que Peter Pan...’ es la misma que la de otra persona para pensar que Juan ha dicho algo verdadero. Wittgenstein exponía así este problema: «¿Cuál es el criterio de igualdad entre dos representaciones? ¿Cuál es el criterio para que una representación sea roja [*das Kriterium der Röte einer Vorstellung*]? Para mí, si es otro el que la tiene: lo que dice y hace. Para mí, si soy yo el que la tiene: absolutamente nada.» (Wittgenstein 1953, parte II, § 377.)

Respecto a la primera cuestión, no parece aceptable defender que ciertas oraciones tienen un significado en boca de una persona – la primera – y otro significado diferente en boca de otra persona – una segunda o tercera – . Los hablantes han de conocer lo que las palabras quieren decir en su lengua y, por tanto, éstas han de referirse a lo mismo tanto si se predicán de uno mismo como si se predicán de otro sujeto. Pensemos en actitudes – o quizá sea mejor decir «estados» – proposicionales como estar triste por algo, desear o creer alguna cosa: cuando un hablante afirma que él está triste, que tiene un deseo o que cree algo, la oración ha de tener el mismo significado que si lo afirma otro hablante. Concibamos como concibamos estos estados no podemos admitir que la oración que la adscribe tenga un significado en boca de la primera persona y otro en boca de la segunda o tercera.

Sin embargo, la segunda asimetría sigue en pie: ¿cómo negarle a la primera persona una posición privilegiada al adscribirse, por ejemplo, una intención o un significado? Davidson lo hace argumentando que «ni el hablante ni el oyente conocen de una forma especial o misteriosa lo que significan las palabras del hablante; y ambos pueden estar equivocados» (Davidson 1984b, p. 13). Ambas ideas, la de que no existe una asimetría entre la primera y la tercera persona a la hora de determinar el significado de las palabras del hablante, y la de que el hablante podría no saber el significado o la referencia de sus expresiones, eran las atacadas por Searle en su crítica al argumento de la indeterminación (v. 2.1).

¿Cómo es posible que un hablante se halle equivocado respecto al significado de sus palabras? Un hablante puede estar equivocado, obviamente, respecto a lo que una palabra significa; es decir, puede estar utilizándola de una manera no aceptada por la mayoría de la comunidad lingüística, pero no parece coherente pensar que no sabe qué significado asocia con la palabra. Esto no implica, por supuesto, que todos los hablantes sean capaces de expresar con otras palabras lo que quieren decir, o que no haya casos dudosos a la hora de aplicación del término, sino, simplemente, que todo hablante sabe lo que quiere decir con sus palabras. (Podemos pensar, por el contrario, en casos como el de los locos o los niños, pero, precisamente en estos casos, no estamos dispuestos a admitir que los sujetos estén hablando una lengua).

Resulta, pues, que la base de la autoridad de la primera persona proviene de los diferentes criterios de aplicación que utilizamos. Los predicados de tipo mentalista – o la atribución de significados – dispondrían de dos criterios de aplicación: a partir de la evidencia observable para otros, y sin utilizar ninguna evidencia para nosotros. Y resulta, desde luego, llamativo que concedamos mayor autoridad a este segundo tipo de adscripción, realizado sin utilizar evidencia alguna. Esta dificultad lleva a Davidson a pensar, como vimos, que si un mismo predicado (como ‘estar triste’) se aplica a veces a partir de la evidencia observable y a veces a partir de otro tipo de evidencia (o ninguna), el predicado es ambiguo, ya que lo que cuenta como evidencia para aplicar un predicado es lo que determina (o ayuda a determinar) su significado. En todo caso, nos interesa aquí la segunda asimetría: la de un tipo diferente de conocimiento o, por decirlo de otra manera, un método diferente de adscripción de estados mentales.

Algunos autores han negado la existencia de una asimetría entre la primera y la tercera persona y, por tanto, la autoridad de la primera persona. Ryle o Quine se contarían entre ellos. Joseph Agassi, recientemente, ha emprendido una batalla contra ambas ideas:

La doctrina del acceso privilegiado es que yo soy la autoridad de todas mis propias experiencias... La tesis fue refutada por Freud (yo conozco tus sueños mejor que tú), Duhem (yo conozco tus métodos de descubrimiento científico mejor que tú), Malinowski (yo conozco tus costumbres y hábitos mejor que tú), y los teóricos de la percepción (yo puedo hacerte ver cosas que no están ahí y describir tus percepciones mejor de lo que tú puedes). (Citado en Davidson 1984b, págs. 6- 7. Nota omitida).

Ryle ha defendido (1949) que la presunta asimetría se debe, simplemente, al hecho de que estamos generalmente mejor situados para observarnos a nosotros mismos de lo que lo están los demás. Ryle afirma, con su gracejo habitual, que, por principio – no sólo de hecho – , los métodos que John Doe utiliza para averiguar algo sobre John Doe, son los mismos que utiliza John Doe para averiguar algo sobre Richard Roe. Y continúa diciendo:

Las diferencias son de grado, no de tipo. La superioridad del conocimiento del hablante de lo que está haciendo sobre el del oyente no indica que tiene acceso privilegiado a hechos de un tipo inevitablemente inaccesible al oyente, sino sólo que está en una posición muy buena para saber lo que el oyente normalmente está en una posición muy mala para saber. (Citado en Davidson 1984b, p. 5).

Sin embargo, parece, más bien, que ningún hablante necesita autointerpretarse para escrutar la referencia de sus términos; esto es, nadie necesita observar su conducta y, a partir de ahí, establecer hipótesis analíticas, como hace con los demás hablantes. Así, para interpretar las oraciones de un tercero, observamos su conducta verbal, pero para interpretar las nuestras, no lo necesitamos. Tenemos, pues, un *acceso privilegiado* a nuestros estados proposicionales o a los significados de nuestras palabras. Ahora bien, ¿en qué consiste exactamente ese acceso privilegiado en que vendría a consistir la autoridad de la primera persona? Davidson, por ejemplo, comenta:

Los filósofos contemporáneos que han discutido la autoridad de la primera persona no han hecho muchos intentos de contestar a la pregunta de por qué las autoadscripciones son privilegiadas. Ya está pasado de moda explicar el autoconocimiento sobre la base de la introspección. Y es fácil ver por qué, ya que esta explicación sólo lleva a la cuestión de por qué deberíamos ver mejor cuando inspeccionamos nuestras propias mentes que cuando inspeccionamos las mentes de otros. (Davidson 1984b, p. 5).

¿Deberíamos apelar a la introspección? Pocos son los filósofos dispuestos a postular la existencia de una especie de «sentido interno» que nos diga de manera inmediata e indubitable la referencia o el significado que asociamos a nuestras palabras.

Con todo, aun cuando no podamos ofrecer una explicación convincente de la autoridad de la primera persona, hemos de admitir – en contra de la opinión de Ryle – que dicha

autoridad existe. Davidson intentará conjugar esta asimetría con el rechazo de algún tipo de acceso privilegiado a la propia mente o de conocimiento especial.

La tesis de la indeterminación es verdadera, en la concepción davidsoniana y quineana, porque los significados no se hallan determinados en la cabeza de los hablantes (el repetido argumento de que el lenguaje es una actividad pública, social); así, quienes, como Searle y Fodor, ven en la tesis una disolución inaceptable del significado y la referencia, defenderían una concepción obsoleta de la mente, en la que ésta se ve como un lugar donde se hallan entidades como los significados y las creencias. Esta concepción cartesiana de la mente ayudaría también a caracterizar la autoridad de la primera persona:

El cartesianismo ofrece una explicación que, en parte, es aún dominante en la filosofía de la mente contemporánea. Es la afirmación de que un sujeto tiene un acceso privilegiado a sus estados mentales, un acceso encuadrado en el modelo de percepción de objetos ordinarios con la diferencia importante de ser infalible y transparente. Dado que estos objetos internos y estos estados de cosas se presentan al sujeto directamente, pueden ser conocidos directa e infaliblemente. En analogía con la percepción física, se supone que el conocimiento introspectivo es un logro cognitivo, la existencia del cual presupone objetos mentales que se manifiestan al sujeto, pero son diferentes de los juicios con los que el sujeto asevera su existencia. Así, de acuerdo con la imagen cartesiana, las auto-adscripciones mentales se hacen verdaderas por estados de cosas independientes, que, con todo, son completamente accesibles al sujeto. (K. Puhl 1994, p. 340).

Davidson aboga por el abandono de esta imagen mediante tres argumentos: en primer lugar, hemos de rechazar la epistemología empirista basada en intermediarios entre el mundo y nuestra mente, se conciban estos intermediarios como *sense data* o como irritaciones nerviosas; en segundo lugar, hemos de abandonar la idea de objetos intencionales «en» o «captados por» la mente; y, en último lugar, si no hay tales objetos que deban representar el mundo para hacer verdadera una creencia, hemos de abandonar también la idea de la correspondencia como explicación de la verdad. A partir de estas premisas, Davidson habrá de ofrecer una respuesta a la crítica realizada por Searle a la tesis de la indeterminación – *yo sé lo que quiero decir* – que no se base en la existencia de entidades tales como significados, referencias o representaciones en la mente. Efectivamente, si aceptamos la indeterminación y las ideas davidsonianas citadas, la autoridad de la primera persona parece cuestionable. En palabras de B. Ramberg:

Una vez que postulamos objetos tales de la mente, entonces, a menos que la esencia del objeto, el contenido, consista en ser percibido por la mente autoobservante, nuestra relación epistémica con ellos deviene cuestionable. Y ya que la visión puramente externista del contenido niega, precisamente, que el *esse* de un objeto intencional de un pensamiento es su *percipi* por la primera persona, se sigue que debemos ser escépticos sobre los contenidos de nuestras propias mentes. (B. Ramberg 2001, pp. 220-221).

La cuestión es, pues, si la asimetría entre el método utilizado para interpretar al resto de los hablantes y para interpretarse a sí mismo conllevaría, necesariamente, la

existencia de un tipo de conocimiento especial en el caso de la primera persona. El argumento sería el siguiente:

- 1) Nada en mi conducta observable hace que mi emisión de la oración O signifique S_1 y no S_2 .
- 2) Yo sé que con mi emisión de O quiero decir (mis palabras significan) S_1 y no S_2 ,
luego
- 3) Yo tengo acceso a un conocimiento acerca del significado de mis oraciones al que no tienen acceso el resto de los hablantes.

Como ya hemos adelantado, para Davidson, el problema no está en admitir la asimetría, que, con Searle, considera obvia:

Mientras que está bastante claro, al menos esquemáticamente [*in outline*], lo que tenemos para intentar adivinar los pensamientos de otros, es oscuro por qué, en nuestro propio caso, podemos saber tan a menudo lo que pensamos sin apelar a la evidencia o recurrir a la observación. (Davidson 1987, p. 16).

El problema radica en buscar una buena explicación para ella y, sobre todo, en no basarla en un modo especial de conocimiento, ni renunciar por ello a lo que hemos llamado explicitismo semántico:

Estoy de acuerdo con Ryle en que cualquier intento de explicar la asimetría entre las afirmaciones en presente de la primera persona sobre las actitudes, y las afirmaciones de otra persona o en otro tiempo, por referencia a una forma especial de conocer o a un tipo especial de conocimiento ha de conducir a un resultado escéptico. Cualquier explicación semejante debe aceptar la asimetría, pero no puede explicarla. Pero Ryle ni acepta ni explica la asimetría; simplemente niega que exista. Ya que creo que es obvio que la asimetría existe, creo que es un error argumentar desde la ausencia de un tipo especial de conocimiento o un modo o manera especial de conocimiento a la ausencia de una autoridad especial; por el contrario, deberíamos buscar otra fuente para la asimetría. (Davidson 1984b, p. 6).

2.5. Contenido intencional y mentalismo

La crítica de Searle a la indeterminación se basa en la apelación a la autoridad de la primera persona, por lo que debería ser éste quien ofreciera una caracterización satisfactoria de la misma. En uno de los párrafos finales de su celebrado artículo, afirma:

El lenguaje es público; y no es una cuestión de significados-como-entidades-introspeccionables, objetos privados, acceso privilegiado o cualquiera de la parafernalia cartesiana. La cuestión, no obstante, es que, cuando entendemos a otro o a nosotros mismos, lo que requerimos – entre otras cosas – es un conocimiento de contenidos intencionales. El conocimiento de estos contenidos no es equivalente al conocimiento del emparejamiento de emisiones con condiciones en el mundo. Vemos esto más obviamente en el caso de la primera persona, y nuestro repudio del caso de la primera persona nos lleva a tener un modelo falso de la comprensión de un lenguaje. Creemos, erróneamente, que entender a un hablante es una cuestión de construir una ‘teoría’, que

la teoría está basada en ‘evidencia’ y que la evidencia debe ser ‘empírica’. (Searle 1987, p. 146).

El concepto fundamental que aparece en la explicación searleana es, pues, el de *contenido intencional*. Ha sido especialmente Føllesdal quien ha puesto de manifiesto la insuficiencia de la explicación searleana, subrayando dos puntos:

- a) «Searle no nos dice cómo llegamos a conocer contenidos intencionales. Todo lo que nos ofrece es un fragemento de teología negativa: *no* es una cuestión de introspección, acceso privilegiado o auto-visión cartesiana [*Cartesian self-insight*]. Pero nunca nos dice qué es.» (Føllesdal 1990, p. 100).
- b) Searle ha malentendido el argumento de la indeterminación, de ahí que piense, por una parte, que de la tesis quineana se seguiría que no hay diferencia entre conejos y partes no separadas de conejos, y esto le lleva, por otra parte, a que haga hincapié en “lo que quiere decir” un hablante o en su contenido intencional – con ‘conejo’ *yo sé* que quiero decir conejo – . Føllesdal recuerda que el argumento no defiende que ‘conejo’ y ‘parte no separada...’ sean intercambiables sin más:

Cuando correlacionamos el lenguaje/teoría de otro con el nuestro, un manual de traducción podría correlacionar su palabra ‘conejo’ con mi palabra ‘conejo’, otro manual podría correlacionarla con ‘estadio de conejo’, un tercero con ‘parte no separada de conejo’, etc. Cada uno de estos manuales diferirá también en numerosos puntos, por

ejemplo, en lo que correlacionan con el signo de identidad. Es de crucial importancia que no mezclemos manuales; una vez que hemos comenzado a traducir con la ayuda de un manual, debemos mantenernos en este manual. No hay manual, compatible con toda la evidencia conductual, que correlacione la palabra de otra persona ‘conejo’ con las tres: ‘conejo’, ‘estadio de conejo’ y ‘parte no separada de conejo’. (*Ibid.*, p. 105).²³

2.5.1. Contenido intencional e internismo

Centrémonos ahora en la primera crítica mencionada. La caracterización searleana del contenido intencional constituiría una típica teoría *internista* del significado. Esta se basa en dos premisas:

1. Saber la referencia de un término es hallarse en un determinado estado psicológico.
2. El significado (en el sentido de la intensión) determina la extensión.

Ahora bien, Searle continúa sin aclarar *cómo* un estado psicológico – un contenido intencional – puede determinar unívocamente una referencia. ¿Qué hace que mi estado mental sea un estado mental de referirse a conejos y no a partes no separadas? Føllesdal opina que esto constituye una versión de Humpty Dumpty:

²³ Señalar que Føllesdal apoya el argumento de la indeterminación basado en la inextricabilidad más el holismo: «Nuestro principal problema al estudiar semántica es separar la teoría a la que se adhieren nuestros profesores de los significados que unen a sus oraciones. Tanto Quine como Davidson han utilizado a veces esta visión de la indeterminación, y creo que es una buena visión.» (Føllesdal 1990, p. 102). (V. 1.2).

Esto me parece una versión modificada de la visión del lenguaje de Humpty Dumpty. El Humpty Dumpty de Lewis Carroll defendía que él podía hacer que una expresión lingüística significara cualquier cosa que él quisiera, Searle afirma que dentro del abanico de posibilidades que deja abiertas la evidencia conductual, el “contenido intencional” del hablante establece lo que significa una expresión. Pero yo le preguntaría a Searle, si el lenguaje es una institución pública y si toda la evidencia que se halla públicamente accesible a quienes aprenden y usan el lenguaje es compatible con diferentes manuales de traducción, que correlacionan ‘conejo’ con varias palabras en el otro lenguaje, ¿cómo puedo entonces *hacer* que ‘conejo’ signifique algo de tal manera que sólo una de estas correlaciones sea correcta? (Føllesdal 1990, p. 106).

Siguiendo ahora la premisa 2, la respuesta de Searle apelaría a la capacidad intrínseca para determinar unívocamente la referencia de los contenidos intencionales. Pero ahora estaríamos acercándonos a una teoría mágica de la referencia. Searle describe así la idea que rechaza: «El contenido Intencional interno del hablante es insuficiente para determinar a lo que se refiere, tanto en sus pensamientos como en sus emisiones.» (Searle 1983, p. 199). La visión internista del significado ha sido desarrollada por Searle, especialmente, en la discusión con H. Putnam acerca de su conocido ejemplo de ‘Tierra Gemela’. Putnam utilizó dicho ejemplo con el objetivo de refutar las teorías de corte internista e ilustrar la suya, una teoría causal y externista de la referencia. Imaginemos dos individuos, uno de ellos habitante de la Tierra y otro de Tierra Gemela, un planeta clonado partícula por partícula de la Tierra. La única diferencia entre ambos planetas es que en la Tierra el agua es H₂O, mientras que en Tierra Gemela, aunque resulta visualmente idéntica, una vez analizada descubrimos que su composición química no es H₂O sino otra muy compleja, abreviémosla con XYZ, y que, por tanto, no

se trata de agua – sino de, digamos, ‘gagua’ – . Así, el hablante de la Tierra se refiere a agua cuando dice ‘agua’, pero su doble de Tierra Gemela se refiere a gagua cuando dice ‘agua’. Se plantea ahora una doble cuestión: a) hallamos que puede haber una diferencia en las referencias o en los significados de dos sujetos sin que haya diferencia alguna en sus estados físicos ni psicológicos (al ser clones suponemos que estos estados son idénticos en ambos sujetos), y b) los hablantes utilizan expresiones cuyo significado son incapaces de determinar. Así, a) se sigue en el ejemplo de Putnam del hecho de que ambos hablantes son clones, esto es, fisiológicamente calcados y, además, al haber sido similar el proceso de aprendizaje del término ‘agua’ (sustancia con la que se lavan, en la que nadan, que beben, etc.), no estamos justificados a suponer diferencias psicológicas relevantes aquí. Por otro lado, b) se sigue de que el agua posee propiedades que ellos desconocen.

Respecto a a), debemos plantear: ¿hemos de pensar que nuestros estados mentales no determinan el significado? Putnam sí lo hace: afirma que abandona «por completo la idea de que si hay una diferencia en el significado... entonces debe de haber alguna diferencia en nuestros conceptos (o en nuestro estado psicológico)». Y respecto a b): ¿hemos de extraer la conclusión de que los hablantes no conocen realmente los significados de las expresiones que utilizan? Putnam también lo hace: lo que determina el significado y la extensión «no es, en general, completamente conocido por el hablante» (Davidson defendiendo la visión putnamiana en Davidson 1987, p. 21). Putnam niega así tanto que los significados estén en la cabeza, como la autoridad de la primera persona al estilo de Searle.

Searle apela aquí a su teoría de la percepción para mostrar cómo ésta puede dar cuenta del ejemplo de Putnam. Es decir, hacer ver que éste no refuta el internismo. Según Searle, cada contenido intencional es autoreferencial en el sentido de que él mismo aparece mencionado en sus condiciones de satisfacción. La situación ahora sería que cuando a Óscar Terráqueo y a Óscar Gemelo (llamemos así a dos hablantes, de la Tierra y de Tierra Gemela, respectivamente) se les está definiendo indéxicamente el agua, éstos pueden tener una experiencia visual idéntica en tipo, pero no tendrán contenidos intencionales idénticos en tipo. La definición indéxica dada a Óscar Terráqueo tendría esta estructura: ‘agua’ es todo aquello idéntico en estructura a la sustancia que está ahora causando *esta* experiencia visual, sea cual sea esa estructura (en este caso, sería H₂O). Mientras que el análisis en el caso de Óscar Gemelo sería: ‘agua’ es todo aquello idéntico en estructura a la sustancia que está causando ahora *esta* experiencia visual, sea cual sea esa estructura (en este caso, sería XYZ). Así es como, aun considerando que el término ‘agua’ de Óscar Terráqueo y el de Óscar Gemelo tienen diferente extensión, ésta seguiría estando determinada por un estado mental. En palabras de Searle:

Así, en cada caso tenemos experiencias idénticas en tipo, emisiones idénticas en tipo, pero, de hecho, en cada caso se quiere decir algo diferente. Esto es, en cada caso las condiciones de satisfacción establecidas por el contenido mental (en la cabeza) es diferente a causa de la auto-referencialidad causal de las experiencias perceptuales. (Searle 1983, p. 208).

No nos extenderemos más aquí sobre el debate entre Putnam y Searle acerca del internismo o externismo, sino que nos interesa centrarnos en si la idea searleana de que es el contenido intencional – mental – el que determina el significado resulta aceptable. M. Devitt critica la supuesta refutación de Searle a las tesis de Putnam por dar por supuesto precisamente aquello que está en cuestión. ¿En virtud de qué representa un «contenido mental»? Es decir, ¿en virtud de qué el término ‘agua’ tiene como condiciones de satisfacción el agua y no la gagua? Como ya habíamos apuntado, la teoría de la referencia de Searle constituye una teoría mágica de la referencia.

... el fallo en la respuesta de Searle a Putnam es obvio. Asume precisamente lo que se está preguntando: que el estado mental de pensar ‘agua’ tenga el poder intrínseco de alcanzar y captar el agua (en vez del agua de Tierra Gemela). [...]

Es nuestra interacción con el mundo lo que determina que tenemos estados mentales que se refieren a partes de ese mundo y, por tanto, tienen contenidos Intencionales. No son los contenidos Intencionales internos de los estados mentales lo que determina que esos estados se refieren a partes del mundo. (Devitt 1990, p. 94).

La crítica de Devitt es, esencialmente, que (a) «nada interno e intrínseco a un cuerpo puede determinar una relación con un objeto que se halla fuera del cuerpo». Si, además, tenemos en cuenta que (b) «el contenido intencional determina una relación con un objeto que se halla fuera del cuerpo», hallamos la conclusión de que (c) «el contenido intencional no es interno e intrínseco al cuerpo». La referencia es una relación entre el término y el referente y las propiedades relacionales no pueden ser intrínsecas a un objeto. Para ilustrar este punto Devitt propone dos ejemplos: A) Yvonne está jugando al

baloncesto y recibe un pase de Raelene. ¿En virtud de qué fue la recepción de la pelota por parte de Ivonne un pase de Raelene precisamente? Nada intrínseco a Ivonne puede determinar esto. Para determinar el pase hemos de hacer mención de Raelene. B) Norm es el hijo de Bruce. ¿En virtud de qué? Nada interno a Norm determina que él es el hijo de Bruce en particular. Bruce es el padre de Norm porque hace años realizó una acción que ha causado que Norm tenga esa propiedad. Devitt concluye que

Suponer que estas propiedades relacionales de Ivonne y Norm son internas es tener teorías mágicas sobre el pase y la paternidad. Similarmente, suponer que los pensamientos de uno pueden dirigirse a [*reach out to*] objetos particulares fuera de la mente es tener teorías mágicas de la referencia y la intencionalidad. (Devitt 1990, p. 83).

Devitt ofrece aún otro ejemplo. Consideremos la solubilidad de la sal en agua. Ésta es una propiedad interna, intrínseca de la sal y, sin embargo, esa propiedad determina su relación con algo externo a ella. De hecho, la sal podría tener esta propiedad aun cuando el agua no existiera. La propiedad de la solubilidad ni tan siquiera presupone la existencia del agua; ni de nada que no sea la propia sal. Así, si conocer el significado fuera una propiedad de este tipo cumpliría la condición del solipsismo metodológico (condición que Putnam impone a los estados mentales genuinamente privados) y estarían en la cabeza, a pesar de lo cual podrían determinar una relación con algo externo.

La respuesta que ofrece el propio Devitt es: la solubilidad no es algo interno e intrínseco a la sal, ya que la sal es soluble en el agua no sólo en virtud de su propia naturaleza sino también en virtud de la naturaleza del agua. Lo que es intrínseco a la sal es la disposición a disolverse dados ciertos estímulos, estén causados por el agua o no. Aun así, la crítica puede continuar argumentado que lo que una postura searleana defiende es que hay algo en la cabeza que, *dado como es el mundo*, determina que alguna parte de ese mundo es el objeto intencional. Así como, dado que el agua es como es, lo que hace que la sal sea soluble en ella es exclusivamente algo interno a ella. En terminología de Searle: la sal, de hecho, cumple las condiciones de satisfacción internamente determinadas.

No podemos detenernos más en este debate entre teorías externista e internistas, uno de los debates filosóficos que más bibliografía ha generado en las últimas décadas; sólo añadir que Devitt responde acudiendo a la teoría causal de la referencia de Putnam y subrayando de nuevo el hecho de que ninguna representación puede referir por poderes ingénitos: «La visión de Searle del significado, con su énfasis en la fenomenología, tiene un toque cartesiano» (Devitt 1990, p. 97) y continúa subrayando que Searle no ha explicado qué hace que el contenido intencional tenga exactamente el objeto intencional que tiene: «La teoría de la Intencionalidad de Searle le adscribe a la mente un poder que es único en la naturaleza, un poder mágico» (*ibid.*, p. 97). Devitt no ve otra razón para defender esa teoría que el que

cuando adopta la perspectiva de la primera persona y mira a su propia mente, *simplemente puede ver* que el contenido es intrínseco. El núcleo de la argumentación es un dogma cartesiano.

Y eso no es el final. Searle *simplemente puede ver* que el contenido Intencional no es causal en absoluto. No se puede explicar, al menos no completamente, incluso por relaciones causales *internas*. Searle cree que dar una explicación causal de los estados mentales, como hacen los funcionalistas, es eliminarlos (1983, pp. viii-ix, 262-3); es negar su mentalidad esencial. (...). Para Searle, la mentalidad consiste en propiedades fenomenológicas que no pueden ser capturadas causalmente. (...). Esto es el resto de la crítica de Searle contra la ciencia cognitiva. Es simplemente más dogma cartesiano. (Devitt 1990, p. 98).

Debemos hacer énfasis, una vez más, en que el problema de basar una teoría del lenguaje en elementos privados, como un contenido Intencional, no es tanto de carga ontológica como de inutilidad. Aunque esta inutilidad debe hacernos extraer también conclusiones de tipo ontológico. Utilizamos ahora palabras de Føllesdal:

Mi objetivo es llevar las cosas hacia lo público, y no considerar los significados en la cabeza, o estados intencionales (o como deba llamárseles) porque no tienen ningún papel. Aun cuando hubiera tales cosas, no ayudarían lo más mínimo a explicar la comunicación. Ésta es mi principal objeción a Searle sobre la cuestión. Y si no prestan ayuda alguna para explicar la comunicación, lo que fue la única razón por lo que los filósofos las introdujeron, deberíamos abstenernos de creer que existen tales cosas. No importa, realmente, si hay intensiones protolingüísticas. La cuestión principal es que en la comunicación no serían de ninguna ayuda. (Føllesdal 1995, p. 66).

Dos son las críticas que hallamos en esta cuestión hacia Searle:

- a) *Dualismo*. Dado que la mente tiene el poder de referirse de manera unívoca a la realidad, y la apelación al acceso privilegiado de cada hablante a los significados de sus palabras – que se hallan en su mente – , la mente posee una propiedad que no poseen ninguna entidad material. El rechazo a la «parafernalia cartesiana» no supone, dentro de la teoría searleana sobre el contenido Intencional, más que un añadido sin justificar; es más, contradicho por el resto de la teoría. Las teorías externistas ofrecen una explicación – en términos de nexos causales – de cómo algo llega a ser una representación de algo, pero Searle rechaza esta idea y postula la capacidad intrínseca del contenido Intencional para referirse a algo:

Searle es perfectamente consciente, por supuesto, de que una explicación de contenido Intencional es una parte central de las teorías causales de Putnam y otros. Él desprecia tales explicaciones. Su objeción es que son desde «un punto de vista de la tercera persona» mientras que el problema es sobre uno «interno de la primera persona». Necesitamos saber «qué hay en la experiencia» que hace un objeto y no otro el objeto Intencional (1983, p. 63). Necesitamos una explicación de las condiciones en las que la persona «se considera estar viendo» [*takes himself to be seeing*] o «pretende referirse a» [*means to refer to*] una cosa y no otra (p. 64). (Devitt 1990, p. 98).

Searle sustituye el nexo causal de las teorías externistas por el poder mágico de la mente de referirse unívocamente a la realidad. Y es aquí donde se hace patente su dualismo:

El elemento dualista se encuentra en la actitud de Searle hacia la cuestión wittgensteiniana: ¿en virtud de qué tiene un estado mental su contenido Intencional particular, un contenido que especifica ciertas condiciones de satisfacción? Rechazar esta pregunta es tomar simplemente como un hecho bruto del mundo que el estado mental tiene ese contenido. Esto es un tipo de dualismo. No necesita ser un dualismo de entidades, pero debe ser al menos un dualismo sobre propiedades – la teoría del «aspecto dual». Supone que la propiedad de un estado mental de tener un cierto contenido no puede ser explicada en términos físicos más básicos. (*Ibid.*, p. 90).

- b) *Vitalismo*. Pero Searle no está dispuesto a aceptar el dualismo; sin embargo, su único argumento para rechazarlo es apelar al cerebro. Así, defiende que «los fenómenos mentales tienen una base biológica: están causados por las operaciones del cerebro y suceden en la estructura del cerebro» (Searle 1983, p. ix). De hecho, llega a considerar su tesis «una especie de naturalismo biológico» (*ibid.* p. 230). Devitt opina que esta respuesta aún empeora la cuestión, dado que ahora el poder mágico pasa de la mente al cerebro, con lo que se cae en el vitalismo: «La apelación de Searle a la biología parece hacer de la Intencionalidad una propiedad misteriosa e inexplicable de las cosas vivientes. Escapa del dualismo para caer en el neovitalismo.» (Devitt 1990, p. 101).

Una teoría semántica que ejemplifica, tan bien o mejor que la propia de Searle, las cuestiones criticadas arriba es la de N. Georgalis. Éste, en vena muy searleana, confecciona una semántica basada en los siguientes conceptos:

- *Referencia intencionada (intended reference)*: «el objeto o el tipo (conjunto) de objetos asignados bajo alguna interpretación.»²⁴ (N. Georgalis 2005, p. 235). Así, por ejemplo, que ‘conejo’ se refiera a los conejos, y no a partes no separadas de conejo. Ésta constituye «el concepto semántico básico» (p. 238).

- *Referencia objetiva (objective reference)*: se trata de la referencia que la comunidad lingüística asocia con un término. El «objetiva» aquí indica «acuerdo intersubjetivo» acerca del tipo de entidades que existen. Al aprender una lengua debemos inferir cuál es la referencia objetiva de los términos de la misma e interiorizarla (v. pág. 237-238). Una noción estrechamente emparentada con ésta es la de *interpretación intencionada (intended interpretation)*, que se refiere «má ampliamente a una oración o un lenguaje y no sólo a sus partes referenciales.» (p. 235).

- *Referencia óptica*. Cuando una expresión se usa para intentar decir qué existe realmente – si conejos o partes de conejos no separadas – . Aquí «estamos hablando, o intentando hablar, en el nivel trascendental.» (p. 238). (v. 2.7).

La idea ahora es que la apelación quineana al valor literal de las palabras (v. 2.1.1). sólo tiene sentido si los individuos asocian a éstas una referencia determinada, dado que «*las palabras mismas no tienen valor literal*» (p. 247). La paradoja pragmática del

²⁴ La expresión está tomada del propio Quine: «La referencia intencionada de los nombres y predicados tiene que aprenderse o por ostensión o por paráfrasis en algún vocabulario antecedente familiar» (1969a, p. 54). Quine señala después que con ambos mecanismos la indeterminación es inevitable.

argumento demostraría, pues, que se hace necesaria la noción de referencia intencionada. Ésta, además, implica la autoridad de la primera persona respecto a la semántica de sus palabras:

Las expresiones mismas no pueden fijar la semántica; ellas mismas no tienen valor literal. La única opción para fijar una interpretación particular parece(n) ser la(s) intención(es) de los agentes que usan esas inscripciones o sonidos. Esto, a su vez, requiere que el agente conozca su interpretación intencionada. Nada de esto es posible sin aumentar la metodología quineana con una de la primera persona. (*Ibid.*, p. 252).

Como se ve, esta teoría *à la* Searle plantea las mismas cuestiones que ya hemos apuntado. Especialmente: ¿cómo consigue un hablante determinar la referencia de sus términos? Esto es, ¿cómo concebir la referencia intencionada? Aquí se apela a que el hablante es directamente *consciente* de ella: «Nuestra conciencia [*awareness*] en tales casos es directa, pero desde una perspectiva de la tercera persona es invisible.» (p. 242).

¿De qué, exactamente, somos conscientes?

El «algo» del que digo que somos directamente conscientes no es del objeto mismo, si hay tal cosa; eso sería referencia óptica[.] (...) El «algo» puede que ni tan siquiera sea la referencia objetiva, ya que mi referencia intencionada puede no coincidir con la extensión de la noción de la comunidad. Es, más bien, la conciencia [*awareness*] del hablante de qué objeto(s) considera que hay y de los que habla cuando realiza su emisión – esto es, el objeto intencionado de su expresión referencial, dejando a un lado la referencia objetiva y la óptica. (p. 243).

Volvemos, pues, a la magia: el hablante considera que está hablando de conejos y no de partes no separadas de conejo, pero lo consigue mediante un acto privado de su conciencia. Y esta apelación a la conciencia supone, a su vez, dualismo. De hecho, se nos dice abrazar «entusiastamente el compromiso de mi teoría con el «mentalismo» muy limitado requerido por los conceptos de referencia e interpretación intencionada[.]» (p. 260). Pero que un término o un lenguaje posean *una* interpretación correcta mediante un acto de conciencia de sus hablantes, no parece constituir un mentalismo «muy limitado», sino más bien todo el mentalismo del que podría pecar una teoría semántica. De manera análoga, se afirma que esta teoría supone «una muy limitada pero inocua “privacidad de lenguaje”» (p. 265), sin embargo, no parece que concebir la referencia como «un acto mental individual» (p. 262) o defender que «el significado se basa, en última instancia, en lo subjetivo» (p. 266) supongan una privacidad ni muy limitada ni inocua. Y, por tanto, resulta sencillo aplicar las críticas que hemos dirigido a Searle: magia, dualismo, vitalismo.

Por otro lado, ¿cómo conseguiríamos inferir a partir de la conducta verbal de los demás hablantes la referencia objetiva de las palabras en una lengua? Se lee que «cualquier mentalismo en mi teoría es un resultado directo del hecho de que la referencia intencionada y la interpretación intencionada son conceptos de la primera persona.» (p. 260). Y, efectivamente, cuando hacemos del significado y la referencia «conceptos de la primera persona», esa pregunta queda sin respuesta.

Si conseguimos aprender el uso de los términos a partir de nuestra observación de la conducta de los otros hablantes, como, de hecho sucede, esto indica que el externismo es inevitable (v. 3.3). «Si el externismo no fuera cierto, entonces, no tendríamos manera de determinar los objetos del pensamiento en los casos más simples, “metodológicamente básicos”» (A. Carpenter 1998, p. 39). Toda teoría semántica mentalista sume, pues, el proceso de aprendizaje de la lengua en un misterio insondable.

2.5.2. Mentalismo y uso

Por otro lado, Føllesdal atribuía a Searle una interpretación errónea de la tesis. Searle insiste en que yo sé que con ‘conejo’ me refiero a un conejo, no a una parte no separada de conejo, a un estadio de conejo, etc. A esto, Føllesdal contestaba con la importancia de, una vez escogido un manual de interpretación, mantenerse en él. Las diferentes posibilidades de traducción – las funciones vicarias – no sólo varían la referencia de los términos, sino también de todo el aparato de individuación, de manera tal, siempre, que todas las oraciones mantengan las mismas condiciones de verdad.²⁵ Así, si Searle continúa insistiendo en que con ‘conejo’ me refiero a conejos, ha malentendido la tesis, ya que esto dependerá de cómo se interprete, por ejemplo, mi signo de igualdad. Es crucial, en orden a entender correctamente la tesis, considerar traducciones de oraciones enteras, o incluso de un manual de traducción completo, no de partes extraídas del contexto. Ya Quine señalaba que

²⁵ De hecho, no todas las interpretaciones deben coincidir en qué palabra considerar un término referencial. Así, ‘conejo’ puede ser considerado un verbo, y ser interpretado como ‘conejea’, al estilo de los fenómenos atmosféricos.

dos sistemas de hipótesis analíticas son, como un todo, equivalentes mientras no impliquen diferencia alguna en la conducta verbal; y, si ofrecen traducciones castellanas claramente discrepantes, se puede afirmar que el conflicto aparente es sólo un conflicto de partes vistas fuera de contexto. (Quine 1960, p. 78).

Aquí, Føllesdal subraya que aún admitiendo entidades mentales, la indeterminación continuaría en pie, ya que también afectaría a éstas. (V. 1.2 y 2.5). De nuevo, podemos preguntar a Searle de qué manera un contenido intencional determina que mi deseo es un deseo de acariciar un conejo, no una parte de conejo no separada.

Sin embargo, dada la indeterminación de la traducción, ¿cuál de las diversas expresiones de otra lengua es apropiada para expresar mi estado mental actual? Cuando me hallo en el estado mental que considero correctamente expresado por ‘conejo’ en castellano, ¿es este estado correctamente expresado por *lapin* en francés? ¿O por *stade de lapin* o *partie non-détachée d’un lapin*? ¿Es *lapin* la traducción correcta porque el nativo francohabla se halla justo en el mismo estado mental cuando pronuncia *lapin* que el castellano cuando pronuncia ‘conejo’? ¿Y se hallan dos hablantes competentes de castellano en el mismo estado mental cuando usan la palabra ‘conejo’? ¿Pero qué quiere decir que se hallan en el mismo estado mental? (Føllesdal 1990, p. 106).

Ya hemos avanzado la respuesta a estas preguntas: «La conclusión que extraeré de esto, es que, dada la indeterminación de las expresiones lingüísticas, apelar a estados mentales no es de ayuda alguna para evitar la indeterminación. La indeterminación se aplica también a los estados mentales.» (*Ibid.*, p. 107). ¿Qué haría que una

representación mental fuera sólo correctamente descriptible como una representación de conejo y no de cualquiera de las alternativas posibles?

Así, Føllesdal avisa de que ninguna entidad mental, por privada, podría determinar el significado de las palabras – Putnam habla de «la futilidad de un acercamiento fenomenológico al problema de la comprensión» (Putnam 1981, p. 20) –, y apela aquí a lo que sería su formulación de lo que Alston bautizó como explicitismo semántico; la denomina “la tesis MMM” – de *man-made meaning* – :

Tesis MMM: El significado de una expresión lingüística es el producto global [*the joint product*] de toda la evidencia que ayuda a los aprendices y usuarios del lenguaje a determinar ese significado. (*Ibid.*, p. 103).

Y se apela, también, a la «evidencia quineano-davidsoniana»:

Por «evidencia» quiero decir, como en la ciencia en general, todo lo que ayuda a eliminar hipótesis. Dado que somos empiristas, la evidencia tiene que ser sensorial, no incluirá telepatía, lectura de pensamientos, etc. Incluso aquellos que apelan a fuentes extrasensoriales de evidencia admitirán que no tienen ninguna función significativa en el establecimiento, aprendizaje y uso del lenguaje. (...) Estos tres procesos se basan en evidencia públicamente accesible, esto es, la tesis de que en particular, el aspecto semántico del lenguaje, lo que llamamos el significado de expresiones lingüísticas, es un producto de esta evidencia, esto es, observación de la conducta. (*Ibid.*, p. 103).

Resulta interesante hacer notar aquí la diferencia en el tratamiento del escepticismo semántico entre Wittgenstein y Quine.²⁶ Kripke indica que Quine basa su argumento en premisas conductistas, mientras que Wittgenstein subraya el hecho de que el escepticismo semántico resulta inevitable aun tomando en consideración toda la evidencia disponible por la introspección. La forma en la que se presenta el argumento escéptico en Wittgenstein, dice Kripke, no es conductista. Se presenta desde «dentro». Así pues, obtendríamos dos líneas de argumentación:

1. *Línea quineana*. Ésta se apoya en el monismo fisicalista – «quizá debemos consentir [*acquiesce in*] el dualismo psicofísico de predicados pero aferrándonos a nuestro sencillo monismo de sustancia» (Quine 1995a, p. 87) – y, dado que los estados microfísicos no determinan unívocamente el significado, obtenemos la indeterminación.

Robert Nozick ha ilustrado ingeniosamente la tesis quineana (ilustración aprobada por el propio Quine (en Quine 1986, p. 365). Imaginemos dos barcos navegando en direcciones contrarias; ambos barcos transportan a niños con sus maestros. Los maestros de un barco enseñan castellano a los niños, mientras que los de otro barco enseñan castellano de estadios temporales – esto es, no se dice

²⁶ En la segunda nota al pie de *Palabra y Objeto*, Quine apunta que «Quizá la doctrina de la indeterminación de la traducción tendrá poco aire de paradoja para los lectores familiarizados con los comentarios sobre el significado de la última época de Wittgenstein». Esta alusión a Wittgenstein ha llamado la atención a algunos autores – «inesperada nota al pie» dice R. Kirk (2004, p. 175), de «sorprendente» la califica D. Koppelberg (1995, p. 337) – , sin embargo, en las *Investigaciones Filosóficas* aparecen varias ideas que prefiguran la teoría de la indeterminación, v., por ejemplo, el párrafo 504, parte I, donde Wittgenstein parece atacar la asimetría entre la primera y la tercera persona respecto a la semántica: «Cuando alguien dice: «¿Cómo voy yo a saber lo que él quiere decir? Yo sólo veo sus signos», así, digo yo: «¿Cómo va él a saber lo que quiere decir? Él también tiene sólo sus signos.»

‘conejo’, sino ‘estadio de conejo’ – . Para abreviar, estos últimos, aun manteniendo la misma referencia de sus términos, pasan de decir ‘estadio de conejo’ a decir ‘conejo’ y así para el resto de términos. Nozick señala que dado que así se han eliminado las diferencias fonéticas, no se podría hallar diferencia alguna entre los niños de ambos barcos. Supongamos que los dos grupos de niños y maestros arribaran a la misma isla y comenzaran a hablar. ¿Cómo sería posible decir quién está hablando castellano y quién castellano de estadios? ¿Habría a este respecto materia objetiva? Nozick cree que no habría diferencia de ningún tipo entre los niños de los barcos, ni neurológica, ni en su conducta verbal; la parábola, pues, ilustra el que no es una cuestión de hecho (*fact of the matter*) qué lenguaje se habla.²⁷

2. *Línea wittgensteiniana.* Esta línea de argumentación resulta más fuerte, dado que no se parte del fisicalismo; la cuestión ahora es hacer ver que aun admitiendo la existencia de entidades mentales o actitudes proposicionales irreductibles no se eliminaría la indeterminación.²⁸ Se subraya ahora, por tanto,

²⁷ Quienes piensan que sí habría una diferencia en el cerebro de los niños deberían explicar cómo es esto posible, dado que sus experiencias (tipo) son idénticas y las palabras que emiten (tipo) también. Pensar, entonces, que «dada una neurociencia completamente desarrollada, podríamos determinar que [las referencias de los términos] eran diferentes examinando sus cerebros» no denota fe en la neurociencia, sino esperar de ella algo similar a un milagro (N. Georgalis 2005, p. 258).

²⁸ Es obvio que no analizaremos aquí la relación entre las doctrinas de Quine y las de Wittgenstein, que ha generado una cuantiosa bibliografía. Con todo, podemos decir con D. Koppelberg que la cuestión en la que más decididamente se separan ambos autores es en la normatividad del lenguaje: para Wittgenstein el lenguaje es cuestión de uso y éste viene regido por normas, en Quine el elemento normativo se ha difuminado. «Es importante», sin embargo, «subrayar que algunas de las críticas más intrigantes expuestas tanto por Quine como por Wittgenstein se dirigen al mismo punto: la aún extendida posición de que tenemos un acceso privado y privilegiado a cuestiones semánticas tales como el significado y la comprensión. (Actualmente, el más conocido representante de tal posición es, quizá, John Searle). tanto Quine como Wittgenstein consideran el atractivo intuitivo de esa doctrina una fuente de gran desconcierto y engaño; ambos han intentado criticarlo a fondo.» (P. 342).

que no puede ser algo interno al sujeto – como el contenido intencional de Searle – lo que le otorga significado a las palabras. Esto es, aun admitiendo la existencia de entidades mentales – de conceptos como esquemas en la mente, por ejemplo – y la posibilidad de la introspección, esto no podría determinar el significado de las palabras. Como hemos visto, ésta es la línea seguida en la crítica de Føllesdal y Devitt a Searle, y la seguida por Davidson, quien apela a intenciones y creencias (v. 1.4). Además, el propio Quine ha dado el visto bueno a esta línea, y dice, en aprobación del planteamiento de Føllesdal:

Dagfinn ha iluminado la tesis de la indeterminación haciendo a un lado lo que no le pertenece. Lo que importa es sólo que el significado lingüístico es una función de la conducta observable en circunstancias observables. (...)

Un conductismo más amplio es irrelevante; el fisicalismo es irrelevante; el monismo es irrelevante. Uno puede regodearse en la más flagrante ontología mentalista sin que ello afecte a la indeterminación de la traducción. (Quine 1990, p. 110)

Aquí son otras las cuestiones que se plantean acerca de la parábola de Nozick relatada arriba:

Parece improbable que surjan conclusiones interesantes sobre la determinación de la traducción de consideraciones neurológicas. ¿Pero surgiría algo interesante sobre la determinación de la traducción de consideraciones sobre estados psicológicos? Se podría pensar en usar la forma en la que estos estados psicológicos fijan el significado para resolver las indeterminaciones de Quine. Pero, como Quine argumentará, ¿cómo pueden estar tan seguros de lo que quieren decir? ¿Cómo pueden estar tan seguros de si hablan

castellano o castellano de estadios? Así que tenemos un quineano que enfatiza los experimentos mentales introspectivos a la manera en la que lo hacen Nozick y Quine, quien no piensa en posiciones que postulan un mundo interno privado que requeriría una refutación elaborada: esta *edición aumentada* de la parábola de Nozick de las dos ramas del castellano muestra que no se puede hallar determinación *ni* en consideraciones psicológicas *ni* en hechos neurológicos. Al contrario que en el cuento de Nozick, la forma en la que se presenta esta edición aumentada no es conductista. Se presenta desde «dentro». (K-M. Cheng, p. 24).²⁹

Resulta de gran relevancia, pues, subrayar que la indeterminación y su rechazo de una autoridad de la primera persona en cuestiones semánticas basada en entidades privadas no se basa necesariamente en el antimentalismo o el fisicalismo. Føllesdal dice a este respecto:

En mi visión, tenemos estados mentales incluso antes de adquirir el lenguaje. Éstos pueden ser estados de nuestro cuerpo, según el monista, o pueden ser estados de algo no físico, según el dualista. Si uno es un monista o un dualista no importa para las cuestiones que discuto en este artículo. (Føllesdal 1990, p. 107).

Algún autor ha hecho, incluso, un repaso de las entidades mentales candidatas a determinar el significado, y la tarea se revela de nula utilidad:

²⁹ La parábola de Nozick intenta mostrar que ni nada dentro de la cabeza de los niños ni nada fuera de ella – en su conducta – podría indicar que se hallan comprometidos con un marco ontológico y no otro; esta argumentación iría pues contra la misma idea de «marco ontológico». Davidson desarrolló esta profunda crítica en un artículo ya clásico de la filosofía contemporánea, «On The Very Idea of a Conceptual Scheme» (Davidson 1974a). La idea básica defendida por Davidson aquí es, como hemos apuntado, que, dado que todos los hablantes comparten un mismo medio físico (recordemos la apelación davidsoniana al estímulo distal) y en referencia a él formamos las creencias y los significados, la noción de marcos conceptuales incompatibles pierde su sentido; es más, la misma noción de marco conceptual – en un sentido ontológico o metafísico profundo – se hace ininteligible.

- a) Creencias. La cuestión aquí sería entonces qué diferencia una creencia sobre los conejos de una sobre partes de conejo no separadas: «La evidencia que apoya una creencia sobre conejos apoyará igualmente la creencia correspondiente sobre partes de conejo, y se llegará a la creencia mediante las mismas inferencias, así que ambas hipótesis acerca de sobre qué son los pensamientos del hablante explicarán igualmente bien sus creencias.» (V. McGee 2005, p. 399). Conviene subrayar este punto: «Somos incapaces de identificar ningún rasgo de los pensamientos del hablante que sería diferente dependiendo de si los pensamientos versan sobre conejos o sobre partes de conejo no separadas.»
- b) Deseos. La cuestión es análoga a la anterior: «Las condiciones de cumplimiento de un deseo serán las mismas para deseos acerca de conejos y los deseos correspondientes acerca de partes de conejo no separadas.» (*Ibid.*)
- c) Lenguaje interno. Tampoco ofrece esperanzas «atender al “lenguaje interno” que Platón indentifica con el pensamiento [nota al pie: *Sofista* 263e]; lo que el hablante dice para sí no será más decisivo de lo que dice en voz alta.» (*Ibid.*) Podemos añadir nosotros aquí que incluso una orden explícita dada por el hablante a sí mismo no serviría de mucho; aun cuando alguien se diera a sí mismo la orden «con ‘conejo’ refiérete a conejos, y no a partes de conejo no separadas, a fusiones de las partes de un conejo, etc.», no se ve cómo podría hacerlo: el término, una vez

utilizado, podría ser interpretado de esas diferentes maneras. Incluso por el propio hablante. Posiblemente, son consideraciones de este tipo las que hacen a Davidson afirmar, como ya vimos en 2.3, que «ya que cada hablante debe saber esto, al menos en algún vago sentido, no puede ni tan siquiera intentar usar sus palabras con una referencia única, ya que sabe que no hay manera de que sus palabras le comuniquen a otro la referencia.» (Davidson 1979, p. 235)

- d) Historia evolutiva. Podríamos pensar que el pasado evolutivo de una especie la ha hecho más proclive a pensar sobre conejos que sobre sus partes no separadas. Putnam ha señalado que la evolución como puede constituir un argumento contra la indeterminación del significado o la referencia, aunque sí lo es a favor de la idea de que una gran cantidad de nuestras creencias han de ser verdaderas. Ciertamente, cuesta trabajo creer que los elegidos por la selección natural no tengamos una mayoría de creencias verdaderas, especialmente cuando consideramos que nuestras acciones se apoyan, indefectiblemente, en nuestras creencias, luego no podríamos tener acciones exitosas sin creencias verdaderas.

Así, podríamos considerar también altamente probable el que la evolución «presione» a cada cultura hacia la formación de conceptos que les serán de gran ayuda en su lucha por la existencia. Pensemos en la gran ayuda que prestarían a nuestros antepasados términos como ‘agua’, ‘fuego’ o ‘mamut’, pero Putnam

vuelve a insistir, como Davidson, en lo abierta que se halla la interpretación de los elementos suboracionales:

La evolución puede producir en mí una tendencia a tener creencias *verdaderas* (de ciertos tipos); pero esto sólo significa que la evolución afecta la supervivencia lingüística o conceptualmente mediada *via* su tendencia a producir en nosotros sistemas de representación cuyas oraciones o análogos a las oraciones tienen ciertas *condiciones de verdad* (y ciertas condiciones de *acción*, o ‘reglas de salida del lenguaje’). Pero *ya se mostró que las condiciones de verdad para las oraciones enteras no determinan la referencia de las partes de las oraciones* (ni ayuda el añadir las ‘reglas de salida del lenguaje’ (...)). Se sigue que es simplemente un error pensar que la evolución determina una correspondencia *única* (o incluso un conjunto razonablemente reducido de correspondencias) entre las expresiones referentes y los objetos externos. (Putnam 1981, p. 41).

- e) Sentimiento. Una respuesta a la desesperada puede consistir en apelar a un sentimiento peculiar que acompaña el pensamiento de un hablante en un conejo, diferente al que acompañaría a uno sobre sus partes no separadas: «Quizá se siente de forma diferente al pensar sobre conejos de lo que lo haría al pensar sobre partes de conejo no separadas. (*Ibid.*) Pero, de nuevo, ¿qué podría hacer intrínsecamente diferente un «sentimiento de conejo» de un «sentimiento de partes de conejo no separadas»? Es más, ¿podemos darle algún sentido a «sentimiento de conejo»?

Así pues, parece que, a pesar de que la tesis de la indeterminación va en su versión quineana – especialmente en su *locus classicus*, Quine 1960, cap. II – unido al antimentalismo, esto no es necesariamente así. No lo es, por ejemplo, en Davidson o Føllesdal.

Recordemos aquí, por otro lado, los aspectos no mentales que podrían aducirse como determinantes de la referencia:

- a) Los *referata*. Se trata aquí de tomar en consideración la historia causal que conecta nuestras palabras con los objetos referidos, pero esto no resulta aquí de ayuda, dado que «aquello que esté causado por un conejo lo está también, ipso facto, por sus partes contiguas.» (*Ibid.*)
- b) Consideraciones neurofisiológicas. Remitimos aquí a la historia de Nozick. En todo caso, nada hace de un estado neuronal un estado de conejo y no de partes no separadas de conejo.

Con todo, hasta qué punto en nuestra discusión la cuestión parece resumirse a si se considera como un hecho primario para la semántica el externismo o la autoridad de la primera persona, queda patente en el hecho de que tras realizar el repaso expuesto arriba, el mismo autor permanece inalterado en su rechazo de la indeterminación: «Pero es simplemente una locura. [*But that's just crazy.*] Nada podría ser más obvio para mí

que cuando uso la palabra ‘conejo’ me estoy refiriendo a conejos.» (V. McGee 2005, p. 401).

2.6. La paradoja pragmática según Soames

Hemos visto que la crítica de Searle al argumento de la indeterminación consiste, básicamente, en que ésta supone una reducción al absurdo de sí misma. La reducción es debida a que la tesis precisa, para su exposición, de nociones que luego pretende declarar inexistentes. Esto es lo que Searle denominó la paradoja pragmática del argumento. Hemos visto en la sección anterior que la apelación de Searle a la autoridad de la primera persona no resulta satisfactoria, sin embargo, esto no implica por sí que el argumento quineano no resulte contradictorio; expondremos, pues, en esta sección el análisis más detallado de la supuesta paradoja, el de Soames en su artículo «The Indeterminacy of Translation and the Inscrutability of Reference» de 1999.

2.6.1. Infradeterminación y referencia

Según Quine, toda la evidencia posible no determina un único manual de traducción, sino que siempre será posible hallar varios manuales incompatibles entre sí que, sin embargo, encajen con toda esa evidencia. El paso de la infradeterminación a la indeterminación resulta automático: si no hay una traducción que podamos llamar *la* traducción correcta, esto implica que la oración o el término no poseen un significado o una referencia determinada. Las teorías lingüísticas, pues, se hallan infra- e indeterminadas.

Quine, como ya apuntamos (v. nota 20; se vuelve sobre ello en 3.1.3), ha hallado serios problemas para dar una formulación satisfactoria de la tesis de la infradeterminación (incluso cuando se aplica a las teorías físicas). La tesis necesita explicar en qué sentido dos o más teorías son incompatibles entre sí y, sin embargo, compatibles con toda la evidencia posible. Tengamos en cuenta que un manual de traducción que tradujera ‘rabbit’ por ‘conejo’ y otro que lo hiciera por ‘estadio espacio-temporal de conejo’ son incompatibles entre sí *para nuestras intuiciones lingüísticas*, pero si aceptamos la tesis de la indeterminación no estaríamos justificados para tratar estas traducciones como incompatibles.

Según Soames (1999, p. 327), Quine parece estar pensando en una incompatibilidad lógica³⁰, pero la incompatibilidad entre los significados de ‘conejo’ y ‘parte no separada de conejo’ no parece ser lógica. Así, dice Soames, las siguientes afirmaciones no serían lógicamente incompatibles:

- (1) El término ‘conejo’ tal y como lo usábamos en el pasado significa lo mismo que el término ‘conejo’ tal y como lo utilizamos ahora.

- (2) El término ‘conejo’ tal y como lo utilizábamos en el pasado significa lo mismo que la oración ‘parte de conejo no separada’ tal y como la utilizamos ahora.

³⁰ En Quine 1970 o en Quine 1975b.

- (3) El término ‘conejo’ tal y como lo utilizábamos en el pasado significa lo mismo que la oración ‘estadio temporal de conejo’ tal y como la utilizamos ahora.

Teorías que contuvieran estas afirmaciones no resultarían lógicamente incompatibles, ya que para ello tendríamos que añadir, entre otras cosas, que ‘conejo’, ‘parte no separada de conejo’ y ‘estadio temporal de conejo’, tal y como las utilizamos ahora, se refieren a entidades diferentes. Soames especifica las premisas que estamos obligados a añadir para hacer lógicamente incompatibles las afirmaciones anteriores:

- (a) Los conejos no son partes no separadas de conejo; las partes no separadas de conejo no son estadios temporales de conejo; los conejos no son estadios temporales de conejo.
- (b) ‘Conejo’ (tal y como lo utilizamos ahora) se refiere a un objeto si y sólo si es un conejo; ‘parte no separada de conejo’ (tal y como lo utilizamos ahora) se refiere a un objeto si y sólo si es una parte no separada de conejo; ‘estadio temporal de conejo’ (tal y como lo utilizamos ahora) se refiere a un objeto si y sólo si es un estadio temporal de conejo.
- (c) Si dos palabras se refieren a cosas diferentes no significan lo mismo.
- (d) Si a significa lo mismo que b y b significa lo mismo que c , entonces a significa lo mismo que c .

Así pues, parece que hemos logrado una clarificación del concepto de incompatibilidad entre manuales de traducción:

A la luz de esto, podemos tomar las afirmaciones de Quine acerca de teorías de traducción incompatibles como afirmaciones acerca de teorías que cuando se les añaden premisas generales [*background claims*] análogas a (a-d) son lógicamente incompatibles. La justificación para apelar a estas premisas auxiliares es que (a) afirma un hecho obvio sobre el mundo, y (b-d) resultan axiomáticas para cualquier teoría que haga un uso significativo de los conceptos de significado y referencia. (Soames 1999, p. 327-328).

Soames afirma que (a) supone un hecho obvio acerca del mundo: los conejos no son estadios temporales de conejos. De hecho, ‘conejo’ y ‘estadio de conejo’ no son tan siquiera expresiones coextensivas: un conejo es duradero en el espacio-tiempo, un estadio de conejo, no. En todo caso, (a) es una afirmación que Quine desea aceptar, pero a la que no encuentra fácil acomodo en su visión del significado. Recordemos unas palabras del propio Quine ya citadas (en 2.1.1):

Parece que nos estamos dirigiendo a la absurda posición de que no hay diferencia alguna, interlingüística o intralingüística, objetiva o subjetiva, entre referirse a conejos y referirse a partes o estados de conejos; o entre referirse a fórmulas y referirse a sus números de Gödel. Con toda seguridad, esto es absurdo, ya que implicaría que no hay diferencia entre el conejo y cada una de sus partes o estados, y que no hay diferencia entre una fórmula y su número de Gödel. La referencia parecería ahora haberse vuelto

una absurdidad no solo en la traducción radical sino en casa. (Quine 1969a, pp. 47, 48).

La premisa (a), pues, no resulta en el marco de las ideas quineanas un hecho tan evidente como debería; se convierte, más bien, en un factor que no puede explicar. Pero tampoco el resto de premisas que hemos de añadir podrían resultar aceptables para Quine. Respecto a (b), la inescrutabilidad de la referencia nos bloquea el afirmar que el término ‘conejo’ se refiere a un objeto si y sólo si éste es un conejo. ¿Qué dato, preguntaría Quine, dentro o fuera de la cabeza de un hablante, puede determinar eso? Y (c) y (d) hacen uso de significados que no parecen ser, meramente, significados estimulativos, sino hallarse más cerca de nuestras nociones intuitivas y prequineanas de significado.

Soames propone ahora una lectura de la tesis de Quine que podría salvarla de la crítica de la reducción al absurdo. Recordemos que dicha reducción se basa en que precisamos, por mor del argumento, ver clara la diferencia entre hacer referencia a conejos o a partes no separadas de conejos, diferencia que el argumento acaba difuminando. Soames ha esquematizado así esta línea de argumentación:

- (1) Está indeterminado (ya que no hay *fact of the matter*) si el nativo usa ‘gavagai’ para referirse a conejos, a estadios temporales de conejos, a partes no separadas de conejos, etcétera.

- (2) Si (1), entonces está indeterminado (ya que no hay *fact of the matter*) si mi vecino usa ‘conejo’ para referirse a conejos, a estadios temporales de conejos, a partes no separadas de conejos, etcétera.
- (3) Por tanto, está indeterminado (ya que no hay *fact of the matter*) si mi vecino usa ‘conejo’ para referirse a conejos o a estadios temporales de conejos, partes no separadas de conejos, etcétera.
- (4) Si (3), entonces está indeterminado (ya que no hay *fact of the matter*) si yo uso ‘conejo’ para referirme a conejos o a estadios temporales de conejos, partes no separadas de conejos, etcétera.
- (5) Por tanto, está indeterminado (ya que no hay *fact of the matter*), si yo uso ‘conejo’ para referirme a conejos, estadios temporales de conejos, partes no separadas de conejos, etcétera.
- (6) Si (1), (3) y (5), entonces no hay diferencia entre referirse a conejos y referirse a estadios temporales de conejos, partes no separadas de conejos, etcétera.
- (7) Si no hay diferencia entre referirse a conejos y referirse a estadios temporales de conejos, partes no separadas de conejos, etcétera, entonces no hay diferencia entre conejos, estadios temporales de conejos, partes no separadas de conejos, etcétera.

(8) Por tanto, no hay diferencia entre conejos, estadios temporales de conejos, partes no separadas de conejos, etcétera

Como afirma Soames, ya que la conclusión, (8), es claramente falsa, y el argumento es lógicamente válido, al menos una de las premisas (1), (2), (4), (6) o (7) ha de ser incorrecta. La primera premisa es la que expone, simplemente, la idea que sustenta todas las demás tesis quineanas: al no haber más hechos relevantes para la traducción que las verdades de la física y éstas no determinan unívocamente una traducción, ésta se halla indeterminada. La premisa (1) ha de ser considerada, simplemente, innegociable. Por otro lado, es seguro que el conjunto de todas las verdades de la física no distingue entre el nativo de la jungla, mi vecino y yo, luego si éstas no determinan la traducción de 'gavagai', tampoco pueden determinar mi interpretación del término 'conejo' en boca de mi vecino o en la mía propia. No podemos salvar el argumento, pues, rechazando las premisas (2) y (4). Llegados a este punto sólo podemos bloquear la conclusión (8) rechazando (6) o (7). La idea de Soames ahora es interpretar el consecuente del condicional que aparece en (6) de tal manera que tanto ésta como (7) resulten aceptables y, sin embargo, no estemos justificados para extraer la conclusión (8). El antecedente de (6) venía a decir que para cualquier hablante h y cualquier expresión e , no está determinado si e , tal y como es usada por h , se refiere a conejos. Ahora se propone tomar el consecuente como afirmando que para ningún hablante h y expresión e es verdad que e , tal y como es usada por h , se refiere a un objeto o si y sólo si o es un conejo *sin ser igualmente verdad* que e , tal y como es usada por x , se refiere a

un objeto o si y sólo si o es un estadio temporal de conejo (o cualquier otro de los referentes potenciales). Dicho de otra manera, el consecuente de (6) vendría a decir que para cualquier hablante h y para cualquier expresión e , e , tal y como es usada por h , se refiere a todos y sólo a los conejos sólo si también se refiere a todos y sólo a los estadios temporales de conejos, etcétera. Así, el condicional entero, esto es, la premisa (6), afirmaría que si no está nunca determinado si una expresión e , tal y como es usada por un hablante h , se refiere a todos y sólo a los conejos (como algo opuesto a referirse a todos y sólo los estadios temporales de conejo o las partes de conejo no separadas, etcétera), entonces, e , tal y como es usada por h , se refiere a todos y sólo a los conejos sólo si también se refiere simultáneamente a todos y sólo los estadios temporales de conejos (y similares).

Así, bajo esta interpretación (7) afirma que si para cualquier hablante h y cualquier expresión e , e , tal y como es usada por h , se refiere a todos y sólo a los conejos sólo si también se refiere a todos y sólo los estadios temporales de conejos, entonces algo es un conejo si y sólo si es también un estadio temporal de conejo (y lo mismo para los demás referentes posibles). Pero de aquí no podemos pasar a la conclusión, obviamente falsa, de (8), que afirma que para cualquier objeto o , no hay diferencia entre ser un conejo o un estadio temporal de conejo (o similar).

Así, Soames propone reformular el argumento a partir de la premisa (5) de una manera, digamos, más directa:

- a) Si (5), entonces está indeterminado si uso ‘conejo’ para referirme a conejos.
- b) Así que está indeterminado si uso ‘conejo’ para referirme a conejos.
- c) Por tanto, no es cierto que use ‘conejo’ para referirme a conejos.
- d) Asimismo, no lo uso para referirme a ninguna otra cosa. Ya que podemos afirmar lo mismo de los demás hablantes y las demás expresiones, las palabras de ningún hablante se refieren a nada (en el sentido ordinario en el que entendemos la noción de referencia).

Ahora, por tanto, no «nos estamos dirigiendo a la absurda posición de que no hay diferencia alguna, interlingüística o intralingüística, objetiva o subjetiva, entre referirse a conejos y referirse a partes o estados de conejos». Si existe una diferencia evidente entre un conejo y una parte de conejo no separada o un estadio temporal de conejo, el argumento sólo afirma que una expresión se refiere a conejos si y sólo si se refiere *también* a partes de conejos o a estadios temporales de conejos. Ya que nunca puede estar determinado por la totalidad de las verdades de la física si una expresión *e* se refiere sólo a conejos – y no al resto de posibles referentes – , entonces resulta igual de cierto que se refiere a conejos o que se refiere a estadios de conejos. Es más, la expresión *e* se refiere a conejos *si y sólo si* se refiere también a estadios de conejos.

2.6.2. Referencia tarskiana y referencia intuitiva

Soames propone contemplar su interpretación del argumento de la indeterminación basándose, a su vez, en su interpretación de la noción quineana de la referencia. Para Quine, es nuestra noción ordinaria de referencia, que toma a ésta en sentido absoluto, la que no tiene sentido. La referencia sólo tiene sentido cuando se relativiza a un lenguaje de fondo o manual de traducción. (V. 2.1.2). Citando a Soames:

Cuando yo digo, utilizando el sustituto propuesto por Quine para nuestra noción ordinaria de referencia, que la expresión α de x se refiere a F 's, estoy diciendo que de acuerdo con algún sistema de traducción T infradeterminado que estoy utilizando, la expresión de x se refiere a F 's. Así, cuando afirmo que la expresión α de x se refiere a F 's relativamente a un manual de traducción T , estoy diciendo que hay alguna palabra o frase p en mi lenguaje presente que satisface dos condiciones: (i) según T , α significa lo mismo que p (T mapea la expresión α en la expresión p), y (ii) p , tal y como la uso ahora, se refiere a F 's. (Soames 1999, p. 348).

Así, Quine estaría proponiendo una eliminación de la noción ordinaria de referencia a partir de su diferenciación en dos tipos: (i) referencia como desentrecomillado [*disquotational reference*] para palabras y oraciones de mi lenguaje actual, y (ii) traducción más desentrecomillado para palabras y oraciones que no forman parte de mi lenguaje actual. Soames basa esta interpretación de la nueva noción de referencia propuesta por Quine en los argumentos que éste ofrece intentado clarificar su exposición de 1969a:

Ahora puedo decir a qué es relativa la relatividad ontológica de forma más sucinta de lo que lo hice en las conferencias, artículos y libro de ese nombre. Es relativa a un manual de traducción. Decir que ‘gavagai’ denota conejos es optar por un manual de traducción en el que ‘gavagai’ se traduce como ‘conejo’, en vez de optar por cualquiera de los manuales alternativos... ¿Y se extiende la indeterminación o la relatividad también, de alguna manera, al idioma propio [*home language*]? En «Relatividad Ontológica» dije que sí, ya que el idioma propio puede ser traducido a sí mismo mediante permutaciones que difieren materialmente de la mera transformación por identidad... Pero si escogemos como nuestro manual de traducción la transformación por identidad, esto es, tomamos el idioma propio con su valor literal [*taking the home language at face value*], la relatividad queda resuelta. La referencia se explica, entonces, en paradigmas desentrecomilladores análogos al paradigma de la verdad de Tarski (...); así, ‘conejo’ denota a los conejos, sean éstos lo que sean, y ‘Boston’ designa a Boston. (Quine 1999, pp. 51-52).

Así pues, Soames considera que Quine está proponiendo reemplazar nuestra noción habitual de referencia por dos nociones afines: la de la referencia como desentrecomillado para el lenguaje actual de uno y la de la traducción más el desentrecomillado para cualquier otro lenguaje. Sin entrar en demasiados tecnicismos, Soames expresa la noción de la referencia como desentrecomillado de manera tarskiana:

Para cualquier nombre n de mi lenguaje actual y para cualquier objeto o , n se refiere a o si y sólo si $n =$ ‘Alfred’ y o es Alfred, o $n =$ ‘Willard’ y o es Willard, o...

Una vez que la referencia se define de esta manera, Soames afirma que hemos de entender la exposición del desentrecomillado

- (i) Para cualquier objeto o , ‘conejo’ se refiere en mi lenguaje actual a un objeto o si y sólo si o es un conejo

como afirmando que

- (ii) Para cualquier objeto o , o bien ‘conejo’ = ‘conejo’ y o es un conejo, o bien ‘conejo’ = ‘perro’ y o es un perro, o ... si y sólo si o es un conejo.

Ya que el conjunto de verdades físicas incluye afirmaciones del tipo ‘conejo’ = ‘conejo’ y ‘conejo’ \neq ‘perro’, entonces (ii) se halla determinado por el conjunto de verdades físicas. Por otro lado, bajo esta interpretación, las afirmaciones acerca de la referencia de mis términos resultan necesariamente verdaderas (veremos que esta línea de argumentación es la seguida por Davidson para defender simultáneamente la indeterminación del significado y la autoridad de la primera persona sobre las adscripciones hechas a sí misma), pero ¿qué sucede con las adscripciones de referencia a una segunda o tercera persona? Cuando afirmo que la expresión e de otro hablante x se refiere a (todos y sólo a los) conejos, mi afirmación puede ser esquematizada de la siguiente manera:

La expresión e , tal y como es usada por el hablante x , significa lo mismo que una expresión de mi lenguaje que se refiere (por desentrecomillado) a (todos y sólo a los) conejos.

Soames subraya que ahora hemos apelado a igualdad de significado y nuestra afirmación, por tanto, queda indeterminada. O bien, Quine podría concebir nuestra afirmación de una manera un tanto diferente:

Hay un manual de traducción T , que he adoptado para traducir las palabras de x a las mías, y hay una expresión r en mi lenguaje actual, tal que (a) T mapea la expresión e en la expresión r , y (b) r se refiere (por desentrecomillado) a conejos.

Soames expone la ventaja de esta interpretación de la siguiente manera:

Aquí, decir que he adoptado un manual de traducción T no necesita implicar que realizo afirmaciones de la forma A significa lo mismo que B que surgen del mapeo que realiza T de A en B . Más bien, quine podría mantener que uno puede adoptar un manual de traducción como una herramienta práctica para tratar con otros hablantes, sin comprometerse con la verdad de las afirmaciones (implícitas) del manual acerca del significado. (Soames 1999, pp. 350-351).

Así, Quine, en su intento de fundar una teoría semántica basada estrictamente en las verdades de la física, había propuesto sustituir nuestra noción intuitiva de significado

por la de significado estimulativo y propone ahora sustituir nuestra noción ordinaria de referencia por la de referencia tarskiana. Con todo, Soames continúa percibiendo problemas:

Normalmente pensamos en los lenguajes como un conjunto de expresiones con ciertas propiedades semánticas. Identificamos los lenguajes en parte por las condiciones de verdad de, y las proposiciones expresadas por, sus oraciones. Si estos hechos semánticos ordinarios se eliminan, entonces no está claro que pueda dar ninguna especificación informativa de qué lenguaje hablo, de cómo difiere de otros lenguajes, o de qué son los lenguajes. Puedo dar las condiciones bajo las que las oraciones son *verdaderas-tarskianamente-en-mi-lenguaje-actual*, pero éstas no son condiciones de verdad, en el sentido ordinario, y no ofrecen información sobre lo que significan mis palabras o sobre qué lenguaje hablo. (Soames 1999, p. 354).

Y aparece de nuevo la crítica de la paradoja pragmática:

¿Puede Quine evitar este tipo de autocontradicción. Bueno, es tentador pensar que podría hacerlo alegando que él no afirmó, aseveró, creyó o desarrolló sus tesis. Pero incluso esto no serviría, ya que si se defiende que no afirmó, creyó o desarrolló nada, entonces aún habrá afirmado algo, y ello es suficiente para refutar sus tesis. El problema es éste: Quine escribe un libro que contiene muchas oraciones. Vemos el libro; queremos saber lo que dice, y lo que Quine estaba intentado decirnos. Sin embargo, todos los conceptos que usaríamos habitualmente para describirlo nos están prohibidos por sus tesis eliminativistas. ¿Podríamos inventar algunas nociones técnicas para reemplazar éstas eliminadas que nos permitieran contestar nuestras preguntas sin reintroducir los rasgos presuntamente objetables de nuestras opciones ordinarias? No hay

razón para pensar que sí. Así, parece que la posición de Quine es realmente autocontradictoria en este sentido. (*Ibid.*, p. 357).

Así pues, a pesar de haber ofrecido una satisfactoria interpretación de la tesis de la indeterminación donde no se genera la paradoja pragmática, Soames continúa considerando autorefutante el argumento quineano; y aunque no apela explícitamente, como Searle, a la asimetría entre la primera y la tercera persona, sí que se nos habla de las proposiciones expresadas por las oraciones o de «lo que significan mis palabras o sobre qué lenguaje hablo». La argumentación de Soames y su posterior rechazo de la indeterminación hacen patentes las dos cuestiones principales aquí:

- a) ¿En qué sentido son incompatibles expresiones como ‘conejo’ y ‘parte no separada de conejo’?
- b) Si la tesis de la indeterminación defiende que no hay diferencias semánticas objetivas entre ellas, no puede formularse sin caer en contradicción.

Sobre a) debemos remitirnos a la especificación del propio Soames expuesta al comienzo de esta sección acerca de por qué las diferentes alternativas quineanas de traducción constituirían interpretaciones incompatibles. Ahora bien, debemos subrayar que la tan repetida cita quineana acerca de la «absurda posición de que no hay diferencia alguna, interlingüística o intralingüística, objetiva o subjetiva, entre referirse a conejos y referirse a partes o estados de conejos» ha llevado a muchos autores a dar por sentado

que un conejo no es lo mismo que una parte de conejo no separada del resto del conejo.

El argumento antiquineano sería entonces el siguiente:

- a) Un conejo y una parte no separada del resto del conejo no son la misma entidad.
- b) ‘Conejo’ denota los conejos y ‘parte no separada de conejo’ denota las partes no separadas de los conejos.
- c) ‘Conejo’ y ‘parte no separada de conejo’ no tienen la misma referencia. Es decir, hay una diferencia entre referirse a conejos y referirse a partes no separadas.

Como se aprecia, el argumento se asienta en a), ¿debe esta premisa ser aceptada sin más? En una línea muy diferente a la anterior, Quine afirma:

Señala un conejo y has señalado un estadio de conejo, una parte integral de conejo, la fusión del conejo, y donde se manifiesta la conejeidad. Señala una parte integral de conejo y has señalado de nuevo los otros cuatro tipos de cosas; y así para el resto.

(Quine 1960, pp. 52-53).

La premisa a) no debe ser admitida como un hecho obvio; más bien debemos preguntar qué distingue un conejo de una parte no separada de conejo o de una epifanía de la conejeidad. Y, por tanto, qué diferencia hay entre hacer referencia a una u otra entidad. (Por otra parte, reparemos en que si negamos la premisa a), invalidamos la explicación de Soames de por qué los diferentes manuales de traducción resultan incompatibles. (Sobre esta cuestión v. 3.2). La solución ofrecida por Davidson será reconocer que no lo son: diferencia no implica incompatibilidad (v. 3.3.2). En todo

caso, debemos insistir en nuestro rechazo de la conclusión c), y la razón es porque «en materia de significado... se puede decir que una palabra está determinada hasta el punto en que lo esté la verdad o falsedad de sus contextos» (Quine 1976, p. 82).

La cuestión b), en todo caso, continuaría en pie: si no hay diferencia entre un conejo y una de sus partes no separadas, ni entre referirse a uno u otra, ¿cómo es posible formular el argumento? Aunque esta cuestión será tratada en 3.2, adelantemos alguna idea. La tesis de la indeterminación puede ser vista como una reducción al absurdo de las intuiciones lingüísticas que nos hacen considerar que expresiones como ‘conejo’ y ‘parte no separada de conejo’ poseen referencias diferentes. Esto es, el argumento no constituye una paradoja pragmática, puesto que no se compromete con la diferencia entre la referencia de ambas expresiones, sino que muestra que la intuición de partida se halla equivocada. El argumento no precisa de una diferencia objetiva entre referirse a conejos o a partes no separadas, sino de que pensemos que hay dicha diferencia.

El propio Quine ha justificado esta lectura de la tesis al afirmar que algunas oraciones sólo difieren de significado de manera intuitiva, pero no estimulativamente. Éste es el lugar:

Las oraciones permanentes pueden diferir entre ellas en «significado», en cualquier explicación intuitiva [*by any intuitive account*], tanto como las oraciones ocasionales; pero, cuanto menos susceptibles son al asentimiento y disentimiento [causado únicamente por la situación perceptiva], menos pistas se pueden hallar en el significado estimulativo. (Quine 1960, p. 36).

Y se dice en nota al pie:

Dos veces me he asombrado al hallar mi uso de «intuitivo» malinterpretado como aludiendo a algún tipo especial y misterioso de conocimiento. Por una explicación intuitiva quiero decir una en la que los términos se usan de la manera habitual, sin reflexionar sobre cómo deberían ser definidos o qué presuposiciones podrían ocultar. (*Ibid.*).

El que la tesis se dirija, pues, a refutar semejantes explicaciones intuitivas del significado, o a sustituir éstas por un paradigma tarskiano de la referencia, como decía Soames, ha abierto una línea de ataque a la tesis siguiendo la vía abierta por Searle. La cuestión ahora sería qué lugar deben ocupar dichas intuiciones en una teoría del lenguaje.

2.7. Nivel inmanente y nivel trascendente

Algunos autores han ofrecido una lectura de la tesis de la indeterminación en la que ésta no resulta autorefutante en el sentido en el que Searle defiende distinguiendo dos niveles de indeterminación, el inmanente y el trascendente:

Exponiendo la tesis trascendentalmente, se dice que la traducción entre lenguajes se halla indeterminada porque no tiene sentido hablar de la referencia única que una expresión dada, digamos ‘conejo’, *realmente* posee, independientemente de un lenguaje de fondo. Exponiéndola interlingüísticamente, por otra parte, la traducción se halla indeterminada en el sentido de que podemos hablar significativamente de lo que

queremos decir sólo relativamente a un lenguaje de fondo. Llamemos al primero el *nivel trascendente*, y llamemos al segundo el *nivel inmanente*. (K-M. Cheng 2005, p. 21).

Encontramos esta misma idea también aquí:

Y el relativista construiría los juicios normativos sobre las traducciones como juicios relativos a un manual que se toma implícita pero erróneamente como fijado. Esto requeriría trazar una frontera entre cómo son las cosas realmente, desde un punto de vista objetivo o teórico (esto es, no hay hechos semánticos determinados, no-relativos-a-un-manual), y cómo nos parecen a nosotros invariablemente desde nuestro limitado y práctico punto de vista. (D. Bar-On 1993, p. 801).

La distinción, pues, es la siguiente:

- (1) *Nivel inmanente*. Relativamente a un lenguaje de fondo podemos hablar significativamente de emparejar expresiones. «En el nivel inmanente, podemos y de hecho hablamos significativamente de la referencia de una expresión dada, tal como ‘conejo’ o ‘estadio de conejo’, hablando sobre su homófono. Hay, por tanto, diferencias entre el significado de ‘conejo’ y el de ‘estadio de conejo’.» (K-M. Cheng 2005, p. 29).
- (2) *Nivel trascendente*. No tiene sentido hablar de la referencia única que un término realmente tiene, independientemente de un lenguaje de

fondo. «Por otra parte, es absurdo hablar de la referencia única de ‘conejo’ y ‘estadio de conejo’ en el nivel trascendente. Por tanto, no hay diferencia entre el significado de estas expresiones diferentes» (*ibid.*, p. 29).

Como vemos, esta idea se basa en la ya estudiada apelación a un lenguaje de fondo (v. 2.1.2). Se trata aquí, para evitar la paradoja pragmática detectada por Searle, de hacer compatible la indeterminación con el hecho de que percibamos diferencias entre expresiones como ‘conejo’ y ‘parte no separada de conejo’, y se busca dicha compatibilidad en la diferenciación de niveles, en uno hay diferencia, en otro no. Así, tenemos por una parte que dicha diferencia se da porque consideramos fijado un manual de traducción, pero el significado se halla indeterminado porque varios manuales son posibles:

Ahora, cuando Searle aprendió a traducir *papillon* de esta manera [como ‘mariposa’], probablemente no cogió la traducción de su propia mente, sino que se adhirió a un manual de traducción establecido que correlacionaba *papillon* con ‘mariposa’. Dado que toda la evidencia que condiciona nuestros manuales de traducción es evidencia conductual, debemos admitir con Quine que otros manuales son posibles. Con todo, una vez que hemos diseñado un manual que ha sido utilizado por millones de personas durante mucho tiempo, no hay necesidad de desecharlo y comenzar a buscar otro. Continuamos utilizando el manual tradicional sin reflexionar sobre la posibilidad de que haya otros. (Føllesdal 1990, p. 105).

Partiendo de esta distinción obtenemos la siguiente lectura del argumento de Searle:

- (a) Como hablante, sé lo que quiero decir.

- (b) Saber lo que quiero decir implica saber qué referencia asocio a los términos.

- (c) Dado que un conejo no es lo mismo que una parte de conejo no separada, yo sé que con ‘conejo’ me refiero a conejos y no a partes de conejo no separadas.

Podemos reescribir (a) como

(a’) como hablante, sé qué referencia (única) poseen mis términos – términos como ‘conejo’ o ‘parte no separada’ – .

Ahora bien, atendiendo a la diferenciación expuesta anteriormente, Searle habría confundido los dos niveles:

El punto crucial aquí es que podemos admitir por una parte que realmente sabemos lo que queremos decir en el nivel inmanente, pero por otra parte negar que haya una referencia única que ‘conejo’ o ‘estadio de conejo’ “realmente” tienen. Por decirlo crudamente: podemos aceptar (a) y rechazar (a’). Por tanto, es un error suponer que (a) implica (a’), y es también un error suponer que la tesis de la indeterminación de Quine debe ser falsa si (a) es verdadera. (K-M Cheng, p. 30).

El error de Searle consistiría, pues, en no diferenciar ambos niveles, lo que lleva a pensar que la tesis de la indeterminación no diferencia entre conejos y partes de conejos – algo absurdo según el propio Quine –, y, además, utiliza diferencias semánticas para después negarlas; pero

podemos explicar las diferencias entre el significado de ‘conejo’ y de ‘estadio de conejo’ sin asumir la existencia de una referencia única. Estas diferencias se explican perfectamente bien en términos de las diferentes maneras de correlacionar las palabras objeto con las del lenguaje de fondo. Sería circular y no persuasivo en absoluto refutar la tesis de la indeterminación de Quine simplemente en términos de la asunción de que ciertas expresiones lingüísticas como ‘conejo’ y ‘estadio de conejo’ deben tener todas una referencia única. (*Ibid.*, p. 31).

Ahora bien, como vimos en 2.1.1 y 2.1.2, la respuesta de Searle ante exposiciones de la indeterminación que apelen al significado literal de las palabras o a lenguajes de fondo será que «no lo podemos tener de las dos formas», esto significaría aquí que es contradictorio afirmar que las expresiones no poseen un valor determinado de manera absoluta pero sí en un nivel inmanente o relativamente a un manual de traducción. El problema es, pues, el siguiente:

Suponemos que la cuestión de qué conjunto de hipótesis analíticas considera fijado un traductor es una cuestión que se puede responder de manera determinada. Si es así, entonces se le da sentido a la noción de un agente lingüístico empleando un conjunto determinado de hipótesis analíticas. Podríamos apelar entonces a esta noción en el caso de la traducción radical, de la siguiente manera. Primero, podemos percibir que tiene

sentido (con la presente presuposición) preguntar qué conjunto de hipótesis analíticas usa un hablante nativo dado para *su propio* lenguaje. Recordemos que Quine considera que la traducción radical comienza en casa: para él, el aprendizaje del lenguaje es en buena parte un proceso de construcción tácita de hipótesis analíticas acerca del lenguaje de los adultos de alrededor (v., p. ej., Quine 1976, p. 58) Se puede afirmar entonces que la tarea del traductor radical es descubrir la hipótesis analíticas que un hablante nativo usa tácitamente, como materia de hecho, al entender su propio lenguaje. Pero si esto es así, entonces la cuestión de qué conjunto de hipótesis analíticas *debería* escoger un traductor radical al traducir el lenguaje de un hablante dado *tendría* (contrariamente a lo que dice Quine) una respuesta determinada. (D. Bar-On 1993, p. 802).

La respuesta quineana aquí, como vimos en 2.1.2, consiste en apelar a una teoría tarskiana de la referencia, es decir, a su idea de la referencia como desentrecomillado. Ésta, como vimos también con Soames en 2.6, puede salvar la idea de la dualidad de niveles. La indeterminación en el nivel trascendente defiende que los términos no poseen una referencia determinada en sentido absoluto; en el nivel inmanente, sin embargo, hallamos que los hablantes utilizan un manual de traducción concreto. Así, si se trata de lenguas diferentes, existe un manual ya fijado, digamos el que traduce ‘rabbit’ o ‘lapin’ por ‘conejo’, y cuando hablamos de la propia lengua los hablantes utilizan la traducción homofónica. En ésta, se aplica el desentrecomillado, y decimos que ‘conejo’ se refiere a los conejos y ‘parte no separada de conejo’ a las partes no separadas de conejos.

Con todo, hemos de reconocer que la apelación a manuales de traducción que se dan por fijados o lenguajes de fondo plantea una dificultad para la tesis de la

indeterminación. Ciertamente, si tiene sentido afirmar que un traductor ha escogido un conjunto determinado de hipótesis analíticas – un manual – , entonces debe de tener sentido afirmar igualmente que los hablantes también escogen uno determinado; a menos que defendamos que sólo los traductores, no los hablantes, pueden utilizar manuales de traducción, pero como sabemos, Quine concibe a todo hablante como traductor, realizando en su lengua materna la traducción de idiolecto a idiolecto. Así, la elección de manual se hallaría infradeterminada, pero no indeterminada; o, dicho de otra manera, la indeterminación aparecería en el nivel trascendente, pero no en el inmanente.

El problema es que Quine dice que la indeterminación comienza en casa. Esto es, tras la apelación a la indeterminación en el nivel trascendente se halla otro proyecto con el que Quine se comprometió desde el principio. Se trata de la indeterminación en el nivel inmanente. Parecería entonces que las diferencias entre el significado de expresiones diferentes como ‘conejo’ y ‘estadio de conejo’ en el nivel inmanente, que pueden ser explicadas perfectamente en términos de las diferencias entre sus homófonos respectivos, son meramente las que aparecen por las *diferentes formas de hablar*. Pero apelar sólo a diferentes formas de hablar no es suficiente para explicar las diferencias entre los significados de expresiones diferentes en el nivel inmanente de forma aceptable. Igualmente, saber lo que significa una expresión no puede ser simplemente conocer de la relación entre ciertas expresiones mencionadas. ¿Cómo, pues, se debe formular la indeterminación en el nivel inmanente, y qué exactamente pretende implicar sobre las diferencias entre el significado de diferentes expresiones en el nivel inmanente? (K-M. Cheng 2005, pp. 32-33).

La noción de inmanencia tiene un papel central en la filosofía de Quine.³¹ Hemos visto (1.1) su rechazo de una supuesta filosofía primera, de una fundamentación del conocimiento científico desde fuera de dicho conocimiento – de la posibilidad, como él la ha llamado, de un «exilio cósmico» (1960, p. 275) – . Podemos variar nuestras teorías, pero no mirarlas a vista de pájaro y, así, juzgarlas desde esa situación privilegiada y escoger entre una u otra. No es posible ofrecer un fundamento para la ciencia que, en su radicalidad, no utilice los datos de la ciencia: «No hay investigación posible sin algún marco conceptual, podemos, así, retener y usar lo mejor de lo que sepamos – yendo hasta el último detalle de la mecánica cuántica, si lo conocemos e importa.» (*Ibid.*, p. 4). La gran candidata para constituirse como filosofía primera, la epistemología, forma parte, también, de la ciencia, es ciencia de la ciencia. Y la posibilidad del exilio cósmico rige para todas las disciplinas, incluida la ontología, la filosofía de la mente o del lenguaje. Y aquí es donde podemos percibir una tensión en las ideas quineanas: ¿no se está formulando la tesis de la indeterminación desde ese exilio cósmico?

Hay, creo, una resistencia discernible por parte de Quine a una comprensión meramente «inmanente» de nuestra posesión de nuestro propio lenguaje, y un deseo de mirar desde

³¹ En realidad, la propia diferenciación entre un nivel inmanente y uno trascendente es obra de Quine, aunque no en referencia al significado sino a la verdad: «En mi pequeño *Filosofía de la Lógica* reactivé la engañosa vieja dicotomía de lo inmanente y lo trascendente, y me atreví a catalogar la verdad como inmanente.» (1995c, p. 353). Se ofrecen para ello tres razones: a) la oracionalidad, «lo que son verdaderas son oraciones, no proposiciones, esto es, significados.» (*Ibid.*); b) el desentrecomillado, «[d]ecir de una oración que es verdadera es, simplemente, incluirla en nuestra propia teoría del mundo.» (*Ibid.*); c) naturalismo, «rechazar cualquier tribunal por encima de nuestra mejor teoría científica del momento.» (*Ibid.*) Ahora bien, dado que la ciencia no fija la verdad, sino que la descubre – una teoría puede estar equivocada – , queda un lugar para la trascendencia: «La verdad aparece como un lugar que continuamos buscando y al que pretendemos dirigirnos. Es un ideal de la razón pura, en la expresión de Kant. Muy bien: inmanente e esos otros aspectos, trascendente en éste.» (*Ibid.*)

arriba [*to stand aloof*] y describir nuestra posición en tesis o doctrinas que se pueden formular y son inteligibles, desde fuera de él. Tengo en mente las doctrinas de Quine de la indeterminación de la traducción y de la inescrutabilidad de la referencia. ¿En qué consiste entender esas tesis como verdaderas de nosotros y de nuestro lenguaje? ¿Es algo que realmente podemos hacer «desde dentro»? ¿Y tendrán la significancia que se supone que tienen si podemos? (B. Stroud 1995, p. 40).

Éste es, pues, el problema de apelar a un nivel inmanente y uno trascendente: que Quine ha rechazado la posibilidad de teorizar en este último. Pero, dado que determinamos la referencia en el nivel inmanente, la tesis debe estar formulada en el trascendente. Pero, entonces, debemos darle la razón a Searle: no podemos tenerlo de ambas maneras, referencia determinada y no determinada a la vez. Quine ha hablado de que hallamos «una cierta sombra de referencia penetrando la misma lengua materna.» (1969a, p. 46). De nuevo, se suscitan varias cuestiones: «¿Se puede decir que sabemos a qué se refieren nuestros términos, o no? Si no, ¿en qué consiste para nosotros hallarse «en casa» en nuestro propio lenguaje? ¿Qué encontramos si intentamos pensar la tesis como verdadera del lenguaje que hablamos?» (B. Stroud 1995, p. 41). Estas cuestiones se resumen en una: ¿por qué permitirle a la lingüística un nivel trascendente que no le permitimos a las demás ciencias si consideramos toda investigación una investigación científica? Así, si existe un nivel inmanente – nuestra lengua, p. ej. – en el que sí es posible ofrecer la referencia y exponer la tesis de la inescrutabilidad, y en la que hay, al menos, «una sombra de referencia», ¿por qué hablar de indeterminación?

Podemos hacer ahora, sobre la base de la diferenciación entre un nivel inmanente y uno trascendente, una nueva lectura de la respuesta quineana al problema que nos ocupa. Esta respuesta se basa en que la referencia no es relativa a un lenguaje de fondo, sino a un manual de traducción. Y, aceptado un manual, no tiene sentido preguntar por la referencia *objetiva* de esos términos. Tampoco es legítimo para la lingüística, pues, la apelación a un nivel trascendente. Y es así como debemos concebir la diferencia entre un nivel inmanente – la posibilidad de diferentes manuales de traducción – y un nivel trascendente – la idea de que pueda existir una referencia unívocamente determinada –. Así, como ya hemos visto en la sección anterior, decir que ‘gavagai’ significa ‘conejo’ «es optar por un manual de traducción en el que ‘gavagai’ es traducido por ‘conejo’, en vez de optar por cualquiera de los manuales alternativos.» (Quine 1999, p. 54). Y preguntar en qué sentido de ‘conejo’ supone querer ascender al nivel trascendental, y ahí es, precisamente, donde se aplica la inescrutabilidad de la referencia: «La referencia se vuelve inescrutable si, con el barco en dique seco, contemplamos un mapeo permutacional de nuestro lenguaje a sí mismo, o si emprendemos una traducción.» (Quine 1981, p. 20). Así, obtenemos una respuesta a qué significa tomar las palabras en su sentido literal: se trata de escoger como manual de traducción la «transformación de identidad» (Quine 1999, p. 52)³²; esto es, la manera de ofrecer la referencia en la propia lengua es el desentrecomillado:

³² Repárese en las fechas: Quine apela a un valor literal y un lenguaje de fondo en 1969 (1969a, concretamente), mientras que la idea de que la traducción sólo es relativa a un manual de traducción (no a un lenguaje de fondo ni similar) aparece en 1981, y en 1999 ya se considera que esta respuesta soluciona el problema: «[S]i escogemos como manual de traducción la transformación de identidad, tomando así la lengua propia [*home language*] en su valor literal [*face value*], la relatividad está resuelta.» (1999, p. 52).

En casa la referencia es captada mediante paradigmas triviales al estilo de Tarski: ‘César’ designa a César y ‘conejo’ denota a los conejos, sean éstos lo que sean. Esto es el valor literal, y en mi naturalismo no pido más. (Quine 1986, p. 367).

Con todo, hay a quien esta respuesta le parece «desesperada», y cree que apelar a un manual de traducción sólo tendría sentido si las palabras de éste

tuvieran algún «valor», algo que pudiéramos «tomar» y hacer uso de ello. Lo que Quine dice sobre detenernos en nuestra lengua materna respecto a la referencia de la palabra ‘conejo’ suena como si fuera, simplemente, una cuestión de emitir una *oración* como ‘La palabra ‘conejo’ se refiere a los conejos’ cuando nos preguntan a qué se refiere esa palabra, y confiar en que a nadie se le ocurra formular la embarazosa pregunta de a qué se refieren esas mismas palabras. Pero eso suena a buena suerte y nada más, o a decir algo con nuestros dedos cruzados. No es algo que pueda detener un regreso. (B. Stroud 1995, p. 42)

La idea aquí es que el desentrecomillado parece no significar más que interpretación homofónica. Presuponemos que cuando otro hablante dice ‘César’ se refiere a lo mismo que yo cuando digo ‘César’, pero tanto en boca de él como en la mía, al término se le ha de atribuir una referencia para su interpretación; la respuesta, pues, parece dejar la cuestión sin tocar. El paradigma tarskiano de la referencia, esto es, la teoría quineana de la referencia como desentrecomillado suele generar insatisfacción entre los filósofos del lenguaje; no resulta extraño que se la considere «trivial e ininformativa» (D. Beisecker 2003, p. 91). Esta sensación suele suscitarse por la idea de que se desea explicar la referencia mediante otras palabras, de las cuales no se explica la referencia; el

problema, pues, vendría a ser el mismo que plantea la apelación a lenguajes de fondo (v. 2.1.2). Así, si se desea explicar la referencia del término ‘conejo’ mediante el siguiente esquema:

‘Conejo’ hace referencia a los conejos,

tenemos que la palabra ‘conejo’ en su segunda aparición debe poseer una referencia para que se haya explicado algo, y, sin embargo, esta referencia queda sin explicar. Aquí es donde se debe tener en mente la forma en la que se aprende el lenguaje – el proceso de triangulación descrito por Davidson, v. 3.3.1 – , y tener presente que un conejo *es* a la vez una fusión de partes de conejo y una parte de conejo no separada. Las ya citadas palabras de Quine acerca de que la referencia en la propia lengua «es captada mediante paradigmas triviales al estilo de Tarski: ‘César’ designa a César y ‘conejo’ denota a los conejos, sean *éstos* lo que sean» apuntan en esta dirección, una dirección que no es, pues, «trivial e ininformativa». Se pueden apreciar en esta crítica a la utilización del desentrecomillado como explicación última de la referencia, como la de B. Stroud arriba expuesta, dos aspectos: por una parte, la idea de que el desentrecomillado no dice nada porque generamos un nuevo regreso, al utilizar palabras cuya referencia no queda explicada, y por otra, la idea de que las diferentes formas en que describimos la referencia de un término al formular la tesis de la inescrutabilidad son percibidas como eso, diferentes. Respecto al primer punto debemos percatarnos de que se basa, precisamente, en la idea en cuestión: que las palabras puedan tener una referencia unívocamente determinada. ¿Después de utilizar el desentrecomillado

podríamos ir más lejos, para explicar la referencia de ‘conejo’, que señalar un conejo? Y esto nos lleva a la segunda cuestión. Como Quine ha señalado, ¿es posible señalar un conejo sin señalar las otras alternativas? Así, la interpretación de Soames de la aplicación por parte de Quine de un paradigma tarskiano a la referencia de los términos, resulta satisfactoria – esto es, una expresión *e* se refiere a conejos si y sólo si también lo hace a partes de conejos no separadas, etc. – .³³

Recordemos que los términos que se utilizan para ofrecer la referencia en un paradigma de tipo tarskiano tienen como función describir las condiciones de verdad del término descrito estructuralmente; el que esta descripción se haga también con palabras – en «‘conejo’ se refiere a los conejos» sólo hay palabras, no hay conejos – no implica que el significado de éstas se halle más determinado de lo que lo están sus condiciones de verdad. Afirmar esto no sería ofrecer argumentos contra la indeterminación, sino asumir su falsedad. La situación es similar en el caso de las actitudes proposicionales: en la lectura propuesta por Quine, afirmar que Juan cree que *p* equivale a afirmar que Juan cree ‘*p*’ – esto es, el verbo de actitud proposicional correlaciona a gente con oraciones, no con pensamientos o similares – . Ahora bien, ‘*p*’ es una oración de nuestra lengua, por lo que somos nosotros quienes debemos interpretarla:

³³ Recordemos en el debate acerca de la referencia que lo relevante es que las oraciones, holofrásticamente consideradas, mantengan su valor de verdad. De ahí que no sólo sea posible sustituir ‘conejo’ por términos coextensivos como ‘parte de conejo no separada’ o ‘fusión de partes de conejo’, sino por otros no coextensivos como ‘conejea’ o ‘estadio de conejo’, siempre y cuando se hagan las variaciones pertinentes en la interpretación de las demás partes de la oración para que ésta mantenga su valor de verdad.

Al ascribir actitudes proposicionales a los hombres y otros animales mediante el entrecomillado, no ascribo dominio de la lengua entrecomillada ni de ninguna otra. Un gato puede creer 'Hay un ratón ahí'. El lenguaje es el lenguaje de quien ascribe la actitud, aunque la ascribe empáticamente a la criatura que tiene la actitud. El gato está, supuestamente, en un estado mental en el que quien ascribe *diría* 'Hay un ratón ahí'. (Quine 1995c, p. 355).

Ahora bien, es el propio Quine quien advierte de que debemos ser precavidos en nuestro análisis de 'p', ya que los valores de nuestras variables «son las cosas de nuestro mundo real, y podrían no coincidir con la ontología de quien tiene la actitud.» (*Ibid.*).

Son consideraciones de este tipo las que motivan la insistencia en la crítica de que, con un paradigma tarskiano, Quine sólo ha explicado algo si las palabras tenían una referencia concreta:

Si las denotaciones de los términos estuvieran también flotando o indeterminadas, [Quine] no habría conseguido especificar la referencia de ese término emitiendo esa oración. De hecho, no habría dicho nada; sólo habría emitido algunos sonidos. Tanto las oraciones V de desentrecomillado como las oraciones desentrecomilladoras sobre la referencia deben poseer un cierto significado, o deben ser entendidas de alguna forma determinada para afirmar o especificar la condición de verdad o la referencia de las expresiones que mencionan. Es esta idea de entender, o conocer a qué se refieren nuestras palabras, y no meramente emitir o estar dispuesto a emitir palabras, a lo que Quine aún, me parece, no ha hecho sitio. (B. Stroud 1995, p. 45).

Así, de nuevo hallamos la idea de que los hablantes deben de conocer el significado de sus palabras si son hablantes competentes, aun cuando no sean muy duchos en explicarlo: «Decir que conocemos las referencias de nuestras palabras no es decir que siempre podemos ser muy útiles o incluso muy ordenados al decir lo que sabemos.» (*ibid.*, p. 49). En todo caso, debemos analizar ahora esta apelación a lo que los hablantes saben acerca del significado de sus palabras y que podría acabar, si fuera tenido en consideración, con la indeterminación del significado.

Ahora bien, no debemos considerar la asignación de significados a las palabras de manera similar a la ascripción de actitudes proposicionales. Algunos críticos de Quine han considerado que la situación es similar sobre la base de que no tendría sentido asignarle a un hablante una ontología de estadios, por ejemplo, si ni tan siquiera *sabe* que esto es posible:

Así, incluso admitiendo la equivalencia de las alternativas, en algún sentido muy fuerte, aún tendría *algún* sentido mantener que un hablante pretendía [*intended*] una y no la otra, a condición de que no sepa que son equivalentes. (...) Así, si intenciono [*intend*] hablar sobre cosas duraderas y no sé que podría estar hablando, equivalentemente, si es equivalente, sobre estadios, entonces aún tiene algún sentido decir que *intenciono* una alternativa y no la otra. (Nos encontramos aquí con dificultades acerca de sustituir en contextos opacos: S intenciona a, y a es equivalente a b, aunque S no sabe esto último, y, por tanto, no es el caso que S intenciona b). (N. Georgalis 2005, p. 333).

Este acercamiento al problema se basa en la idea de que, dado que la mayoría de hablantes ni tan siquiera son conscientes de que se los podría interpretar como

refiriéndose a estadios de cosas, en vez de a cosas, no tendría sentido asignarles ese tipo de referencia a sus términos. La respuesta es que esto supondría atenerse a las intuiciones de los hablantes cuando confeccionamos una teoría semántica, algo que no hacemos cuando confeccionamos una teoría física o biológica. Esto es, si las expresiones ‘conejo’ y ‘estadio de conejo’ son empíricamente equivalente para cualquier estimulación posible, ¿por qué deberíamos prestar atención al hecho de que la interpretación de estadios suscite extrañeza en los legos en semántica?

PARTE 3. INTUICIÓN LINGÜÍSTICA Y SUBJETIVISMO

Este capítulo parte de la forma más actual que ha tomado la apelación a la primera persona en cuestiones semánticas: la apelación a las intuiciones lingüísticas de los hablantes competentes – siendo D. Bar-On quien más ha insistido en este punto – . Se examina la idoneidad de apelar a semejantes intuiciones, examinando cuestiones como si tales juicios resultan de valor científico o si se cae aquí, inevitablemente, en un cartesianismo semántico. Se presta especial importancia a la cuestión de hasta qué punto dicha apelación respeta un elemento básico en el aprendizaje y uso del lenguaje, como parece ser el carácter público de la semántica.

Finaliza el análisis con una defensa de la concepción davidsoniana de la autoridad de la primera persona respecto al significado, y se analiza, especialmente, cómo es posible conjugar una defensa de ésta con la indeterminación del significado.

3.1. Intuiciones lingüísticas

Como hemos visto, la tesis de la indeterminación niega que exista materia objetiva sobre los significados más allá de lo que se pueda manifestar en la conducta verbal de los sujetos – o, en la versión más estrictamente naturalizada de Quine, de las estimulaciones nerviosas – . Recordemos su idea de que la teoría física se halla infradeterminada, mientras que la semántica se halla indeterminada (v. 1.8). La idea de Quine – no de Davidson (v. Davidson 1999b) – es que la indeterminación de la traducción difiere de la infradeterminación de la teoría científica en que la única evidencia que se debe tomar en consideración para determinar la corrección o incorrección de los manuales de traducción es la conducta verbal de los nativos y, por tanto, toda hipótesis que vaya más allá de esta materia, se halla indeterminada. Quine ve claro que «cuando reflexionamos sobre los límites de los datos posibles para la traducción radical, no puede haber duda sobre la indeterminación» (Quine 1987, p. 9). En el caso de la ciencia natural, dado que todas las observaciones posibles resultan, también aquí, insuficientes para determinar una teoría unívocamente, las teorías se hallarán en este ámbito infradeterminadas, ahora bien, la ciencia natural sí posee una «materia de hecho», al contrario que las hipótesis de la «semántica tradicional» y, por tanto, no se da aquí indeterminación. Es decir, aun si consideráramos unívocamente fijada una teoría física, con una descripción exhaustiva de la realidad material, las teorías semánticas continuarían infradeterminadas. El comentario quineano acerca del carácter «paralelo pero adicional» de la indeterminación respecto a la infradeterminación es así glosado por un comentarista:

«Paralelo» se refiere a las similitudes metodológicas entre los dos casos [el del traductor radical y el del físico], mientras que «adicional» indica la persistencia de la infradeterminación de las nociones semánticas aunque no hubiera irresolubilidad empírica en las ciencias naturales, y todos los datos físicos fueran conocidos. (T. Bonk 2008, p. 221).

Al final de la anterior sección hemos visto aparecer una importante tensión en la idea de los diferentes conjuntos de hipótesis analíticas: si tiene sentido afirmar que un traductor escoge un conjunto es que éstos son «objetivos». Si los traductores pueden escoger qué referencia van a asignar a los términos, entonces es que los términos poseen una referencia, con lo que la teoría semántica se hallaría infradeterminada, como toda teoría científica, pero no indeterminada. La respuesta de que los traductores pueden escoger un conjunto de hipótesis analíticas pero que los hablantes no utilizan conjunto alguno en su práctica lingüística habitual, no parece muy persuasiva: la indeterminación se daría en el uso cotidiano del lenguaje, pero no en la traducción o en la reflexión sobre éste – en la exposición de la tesis de la indeterminación, por ejemplo –. Además, si los diferentes conjuntos de hipótesis analíticas son incompatibles, como afirma Quine, pero generan las mismas predicciones respecto a la conducta verbal de los hablantes, ¿no estamos apelando a diferencias “objetivas” de significado más allá de las se hacen manifiestas en la conducta verbal? Estas consideraciones han llevado a algunos autores a defender la existencia de una materia objetiva para la semántica más allá de la que se puede discriminar sobre la base de la conducta verbal.

3.1.1. Intuiciones e infradeterminación

Volvamos al debate acerca de la infradeterminación de las teorías de la ciencia natural (v. 1. 8). Señalamos allí sucintamente que Quine había considerado errada esta idea (v. nota 20). Como dijimos, la razón del rechazo de la tesis es la imposibilidad de hallar un sentido satisfactorio a la idea de la incompatibilidad entre teorías que generan las mismas predicciones (categóricas observacionales, en la terminología quineana, v. Quine 1987, p. 27, y en este trabajo 1. 3). Ciertamente, no hay fácil acomodo entre (a) una visión radicalmente instrumentalista de la ciencia, donde la parte teórica de la ciencia constituye sólo un esquema para sistematizar observaciones y generar categóricas observacionales, y (b) la infradeterminación de las teorías, donde dos teorías de las que se deducen las mismas categóricas observacionales resultan lógicamente incompatibles. Esto es, en la visión quineana – donde, subrayemos, la parte teórica de la ciencia no posee un significado concreto – no puede darse simultáneamente incompatibilidad lógica y equivalencia empírica de dos teorías. El argumento, pues, sería el siguiente:³⁴

(1) Lo esencial de las teorías científicas, en el sentido de que constituyen sus «puntos de control», son las oraciones observacionales categóricas.

(2) Si las observacionales categóricas coinciden las teorías no son incompatibles entre sí. (Las observacionales categóricas son oraciones puramente

³⁴ Se sigue tanto en esta esquematización como en la siguiente el artículo ya citado de Soames.

observacionales y debems pensar, por tanto, que su traducción no se halla indeterminada. Así, dos categóricas observacionales coinciden cuando una es traducción de la otra).

(3) Si dos teorías son compatibles con toda la experiencia posible, de ellas han de seguirse las misas categóricas observacionales.

(4) Hay teorías que son compatibles con toda la experiencia posible y son mutuamente incompatibles.

El argumento resulta claramente incorrecto. (1) es la idea quineana básica acerca de las teorías científicas, y de aquí se sigue, de manera natural, (2) y (3). Si las teorías físicas tienen como función servirnos para deducir oraciones categóricas observacionales (verdaderas), entonces, si de dos teorías se extrae el mismo conjunto de éstas, no tiene sentido considerarlas incompatibles; (3), por su parte, simplemente expone la idea de compatibilidad empírica: una teoría es compatible con la experiencia cuando de ella se derivan únicamente oraciones categóricas verdaderas. (4) se sigue de (1-3). La reducción al absurdo se evitaría si se permitiera apelar a teorías que, por ejemplo, utilizaran ontologías diferentes, pero esto supondría dar por sentado que tiene sentido hablar de ontología en términos absolutos y, por tanto, de la referencia de los términos.

Y este mismo problema se plantea respecto a los manuales de traducción – debido a su carácter de teoría científica – , es decir, al argumento desde arriba. Esquemmatizando también este problema:

- (1) Las únicas diferencias genuinas de significado son las que quedan reflejadas en el significado estimulativo.
- (2) ‘Parte no separada de conejo’ y ‘conejo’ tienen el mismo significado estimulativo.
- (3) Luego, ‘conejo’ y ‘parte no separada de conejo’ no difieren en significado.
- (4) Un manual de traducción MT1 traduce ‘gavagai’ por ‘conejo’ y otro manual MT2 traduce ‘gavagai’ por ‘parte no separada de conejo’.
- (5) MT1 y MT2 son manuales de traducción incompatibles.

Se ve, de nuevo, que el argumento no es correcto. ¿Cuál es la premisa que Quine debería rechazar? Dejando aparte (4), que es nuestra suposición para exponer el argumento, (1) constituye una tesis quineana básica; (2) es obviamente cierta; de (1) y (2) se sigue (3); y los ejemplos de Quine acerca de las diversas ontologías posibles, como la de conejos o partes de conejos, pretenden ser ejemplos de manuales de traducción incompatibles entre sí, luego de (1-4), obtenemos (5). Aquí es donde acecha

la reducción al absurdo denunciada por Searle, como ya veíamos en el caso de la infradeterminación de las teorías. Por tanto, si no se acepta que la tesis sea autorefutante, se debe ofrecer una correcta explicación de la incompatibilidad de la que habla la premisa (5).

Una respuesta consiste en ver la incompatibilidad entre teorías empíricamente equivalentes como una incompatibilidad entre los significados de sus expresiones, y apelar, de nuevo, a dos niveles en el significado.³⁵

Nuestro intento de superar las dificultades para reconciliar las posiciones de Quine sobre la indeterminación y la infradeterminación condujeron a una cierta construcción de la argumentación de Quine. Esa construcción se apoyaba en una apelación a un contraste general entre lo que se puede decir desde un punto de vista intuitivo (la ‘perspectiva del hablante’) y lo que se debería decir desde un punto de vista objetivo (la ‘perspectiva teórica’). (D. Bar-On 1986, p. 257).

3.1.2. Nivel teórico y nivel básico

Matizando más esta idea, se apela a la existencia de un nivel teórico (*theoretical level*) y uno básico (*ground level*).

1. Nivel teórico. Se trata aquí de las teorías físicas y de las semánticas confeccionadas utilizando «evidencia quineana». Aquí hallamos la habitual infradeterminación de las teorías científicas.

³⁵ Se sigue ahora la terminología utilizada por D. Bar-On, por ser quien más ha profundizado en dicha diferenciación, aunque ésta es similar a la hecha en 2.7 entre un nivel inmanente y uno trascendente.

2. Nivel básico. Se apela ahora a los juicios que los propios hablantes realizan sobre el significado de sus palabras. Algunas consideraciones se hacen necesarias aquí. Quine concibe el proceso de aprendizaje de la lengua de forma similar a como concibe la tarea del traductor radical; y ambos, niño y lingüista radical no tienen más datos «que las concomitancias de la emisión y la situación estimulativa observable» (Quine 1969a, p. 81). Ahora bien, este método de asociar significados con expresiones – Quine se refiere a él a como condicionamiento, inducción u ostensión – no nos lleva muy lejos. Tanto el lingüista como el niño han de recurrir a hipótesis analíticas para poder interpretar la mayor parte de las emisiones a las que tienen acceso. De ahí que Quine afirme que la idea de que «el condicionamiento es insuficiente para explicar el aprendizaje del lenguaje... va de la mano [*is of a piece*] de mi doctrina de la indeterminación de la traducción» (Quine 1976, p. 58). Así pues, la idea es que así como el lenguaje científico abandona pronto la inducción para aplicar el método hipotético-deductivo (v. nota 18), de manera similar, el niño abandona pronto la ostensión para aplicar la analogía y otros mecanismos. Así, dos niños «pueden alcanzar un idéntico dominio del castellano por medio de muy diferentes procesos de asociación y ajuste tentativos de los varios adjetivos y partículas interdependientes» (Quine 1960, p. 94). Quine insiste en esta idea afirmando que «dos hombres podrían ser exactamente iguales en todas sus disposiciones a la conducta lingüística bajo todas las estimulaciones sensoriales posibles, y, sin embargo, las ideas expresadas en sus idénticamente causadas y

homófonas emisiones [*identically triggered and identically sounded utterances*] podrían divergir radicalmente, para los dos hombres, en un amplio espectro de casos» (*ibíd.*, p. 26). Esas «ideas expresadas» constituirían el nivel básico, esto es, los juicios semánticos de los hablantes:

Pensando en términos de adquisición del lenguaje, la afirmación es que dos hablantes que convergen plenamente en sus disposiciones verbales pueden, no obstante, divergir radicalmente en sus *juicios semánticos* adquiridos de manera no inductiva (o sin condicionamiento). Éstos son juicios acerca de, por ejemplo, ‘sobre qué’ son las expresiones de la lengua que uno ha adquirido, sobre cómo una expresión dada ‘divide su referencia’ (si denota conejos o estadios de conejo o...), sobre si las expresiones son sinónimas o no (como algo opuesto a ser meramente estimulativamente sinónimo), así como juicios acerca de la estructura correcta (semánticamente relevante) de las oraciones. (D. Bar-On 1992, p. 238).

Así pues, la infradeterminación de las teorías lingüísticas tendría una doble vertiente: (a) en el nivel teórico; ésta sería la indeterminación que hallamos en las teorías de la ciencia natural: «Quine nos insta a reconocer, en particular, la infradeterminación teórica de la semántica, que consiste en el hecho de que podría haber teorías incompatibles sobre la especificación correcta de significados en cualquier lengua bajo estudio (todas compatibles con los datos relevantes)» (*ibíd.*, p. 237); (b) en el nivel básico; teniendo en cuenta lo dicho anteriormente sobre la concepción quineana del proceso de aprendizaje, se interpreta la traducción radical de Quine como un argumento a favor de la infradeterminación en este nivel. Y lo que más nos interesa aquí, este nivel se ve como constituido por los «juicios semánticos» que los hablantes realizan sobre su

propia lengua. Estos juicios, aun estando infradeterminados, no se rechazan como evidencia admisible para la confección de las teorías lingüísticas. Es más, podrían constituir la evidencia principal:

Es muy tentador pensar en una teoría semántica empírica como intentado sistematizar y explicar los juicios semánticos de nivel básico hechos por los hablantes de un lenguaje dado. Lo que distingue la semántica empírica así concebida de las ciencias especiales como la química o la biología es que la primera se ocupa, en última instancia, de los *juicios* de nivel básico de los hablantes. Sin embargo, una teoría química se ocupa, si se quiere decir así, de *fenómenos* de nivel básico. El hombre corriente puede tener, por supuesto, todo tipo de creencias ingenuas sobre los fenómenos relevantes; pero la teoría no se ocupa en absoluto de esas creencias. Así que en ciencias como la química o la biología no habría analogía relevante a la infradeterminación semántica en el nivel básico. (*Ibid.*, p. 239).

Dos críticas suscitadas a menudo por los críticos de Quine están aquí en discusión: a) la infradeterminación de la traducción no implica indeterminación, y b) no hay infradeterminación en caso del lenguaje. Esta segunda es la defendida por Searle, con el argumento de que «yo sé lo que quiero decir». La primera es la escogida por Chomsky. Vimos en la sección 1.8 la polémica surgida acerca del rechazo quineano de una materia objetiva para la semántica pero no así para la ciencia natural. Veamos ahora esta crítica con la apelación a los juicios de nivel básico.

La respuesta de Quine aquí consiste en insistir en que la indeterminación es un elemento diferente de la mera infradeterminación y en admitir que, efectivamente, la

igualdad o no de significado depende de las intuiciones de los hablantes, intuiciones que pone en conexión con la introspección, a la que se niega valor objetivo alguno:

Mi conjetura de la indeterminación de la traducción es un tipo de cosa diferente [a la infradeterminación]. Se trata de que, en el caso interlingüístico general, la noción de igualdad de significado es una cuestión objetivamente indefinible de intuición. Esto implica que la noción de los significados como entidades, aun abstractas, es insostenible, ya que no hay entidad sin identidad. Rechazo la introspección como criterio objetivo, aunque sea de gran valor heurísticamente. (Quine 2000, p. 418).

3.1.3. Nivel básico: infra- e indeterminación

Se trata ahora de admitir el paso de la infradeterminación a la indeterminación sólo en aquellos casos en los que los propios hablantes no consiguen ponerse de acuerdo acerca del significado, ya que, recordemos, serían ahora los juicios de los hablantes sobre su lengua lo que conforma la evidencia principal para una teoría semántica:

¿Por qué la infradeterminación semántica de nivel básico implicaría indeterminación del significado? Brevemente, parece que las maneras en las que los hablantes adquieren y entienden su lenguaje, sus prácticas y juicios lingüísticos nos dan todo lo que hay para la corrección o incorrección en cuestiones lingüísticas. Si, tras un entrenamiento lingüístico normal, los mismos hablantes son incapaces de determinar (en el sentido en que sus juicios podrían divergir irreconciablemente), p. ej., si dos expresiones son o no sinónimas, o sobre lo que era una expresión dada, entonces parece que no podría haber ni siquiera en principio más hechos para decidir la cuestión. (D. Bar-On 1992, p. 239).

Así pues, la infradeterminación en el nivel básico sería inevitable por las diferencias entre los hablantes en el aprendizaje del lenguaje, pero que haya infradeterminación o indeterminación dependerá de si los hablantes convergen o no en sus juicios lingüísticos. Por otro lado, debemos considerar que los juicios semánticos de los hablantes – «juicios intuitivos acerca de lo que significan las expresiones» (como dice D. Bar-On en, por ejemplo, 2004) – , en cuanto tales, suponen ya una teoría, por lo que si una semántica empírica los toma en consideración, obtenemos dos niveles de infradeterminación: en el nivel teórico y en el básico. Esta diferenciación se apoya, además, en otra: la de ciencias de primer y de segundo orden. Así, se nos invita a concebir la semántica empírica como intentando sistematizar y explicar los juicios semánticos de nivel básico hechos por los hablantes de un lenguaje dado. La semántica empírica así concebida diferiría de ciencias especiales como la química o la biología en que trata sobre los *juicios* de nivel básico de los hablantes. En este sentido, es de «segundo orden», mientras que la química, la biología, o la teoría de la fisiología humana, son de «primer orden». Así, cuando se trata de ciencias naturales de «primer orden», la distinción nivel básico-nivel teórico no hallaría aplicación. La distinción serviría, se nos dice, para señalar una cierta diferencia metodológica entre las ciencias naturales, por un lado, y las ciencias cognitivas o sociales por otro. Como la lingüística empírica, las teorías que estudian fenómenos cognitivos o sociales, como el estudio de la percepción visual, la psicología moral, la economía, y varias ciencias sociales, suponen la sistematización,

explicación, y predicción de juicios de nivel básico hechos por sujetos. (v. D. Bar-On 2004).

Esquematizando la argumentación expuesta hasta aquí:

- a) Los hablantes realizan juicios intuitivos acerca del significado de sus palabras.
(Se trata del nivel básico).

- b) Por la forma en que se adquiere el lenguaje, estos juicios se hallan infradeterminados. (Hay una infradeterminación en el nivel básico).

- c) En cuanto juicios semánticos, éstos han de ser considerados una teoría del lenguaje. (El nivel básico supone una teoría infradeterminada).

- d) Una semántica empírica debe tomar en consideración dichos juicios intuitivos.
(Una semántica empírica es un nivel teórico).

- e) En cuanto teoría científica, una semántica empírica se halla infradeterminada.
(Hay una infradeterminación en el nivel teórico).

- f) Una semántica empírica se halla afectada, pues, de una infradeterminación de segundo grado: se trata de una teoría infradeterminada sobre una teoría infradeterminada.

La idea ahora (v. D. Bar-On 2004) es, pues, que en el caso de la semántica empírica, la tesis quineana es que hay infradeterminación semántica de nivel básico – infradeterminación de los juicios intuitivos que los hablantes ordinarios llegan a hacer en el curso de la adquisición de su lenguaje. Para que esta afirmación tenga sentido, necesitamos considerar los juicios semánticos de nivel básico como, en algún sentido, juicios *teóricos* – juicios que representan el resultado de teorizar (aunque sea tácitamente), y que, además, forman una teoría, por vaga e imprecisa que sea. La semántica empírica – concebida como el estudio de los juicios semánticos – constituiría entonces una teoría infradeterminada de juicios de nivel básico infradeterminados; exhibiría infradeterminación de segundo grado.

Y, como se ve, si la semántica empírica se dedica al estudio de los juicios de nivel básico que se hallan infradeterminados, se queda sin materia objetiva. La semántica empírica, por tanto y siempre según D. Bar-On, debe tratar sobre una materia no infradeterminada, so pena de la indeterminación.

A la anterior esquematización debemos añadir, pues, una conclusión:

- g) La infradeterminación en el nivel básico genera la indeterminación del significado.

Se trata aquí de una reconstrucción del argumento quineano acerca de la infradeterminación de la traducción de una teoría física visto en 1.8. Recordemos que Quine hablaba también de una infradeterminación «en segunda intensión». Se trataba allí de que el conjunto de observaciones posibles no determinan unívocamente una teoría, por lo que la fijación de la traducción de las oraciones observacionales no determina unívocamente la traducción de las oraciones más teóricas. En palabras de Quine ya citadas:

Si nuestra teoría física puede variar aunque todas las observaciones posibles estén determinadas, entonces nuestra traducción de su [de un nativo del país del «gavagaiense»] teoría física puede variar aunque nuestra traducción de todas sus oraciones observacionales esté determinada. Nuestra traducción de sus oraciones observacionales no determina más nuestra traducción de su teoría física que lo que nuestras observaciones posibles determinan nuestra teoría física. (Quine 1970a, p. 181).

El argumento anterior se presenta como una reconstrucción de este argumento quineano. Como vimos, la infradeterminación en segunda intensión de la que habla Quine se basa en la posibilidad de traducir de diferente manera el vocabulario menos observacional de una teoría – ya que «los criterios observacionales de los términos teóricos son normalmente muy flexibles y fragmentarios» – , no en la infradeterminación de la teoría lingüística que desarrolla el nativo. Con todo, dado que esta teoría contendría sus opiniones sobre el significado de los términos teóricos, constituye una interpretación posible.

3.2. La teoría «de la teoría» del lenguaje

Como vemos, el argumento se asienta en la consideración de los juicios de los hablantes como una teoría lingüística – lo que se ha bautizado como la teoría «de la teoría» [*'Theory'-Theory*] (v., p. ej., D. Bar-On 2004) –. Esto es, en considerar que el proceso de aprendizaje consiste en la elaboración de una teoría del significado.

Éste es un punto sobre el que han coincidido tanto defensores de la indeterminación, como Quine y Davidson, como oponentes, como Chomsky. Este último presenta «el problema para el lingüista, así como para el niño que aprende el lenguaje» como el de «determinar el sistema subyacente de reglas a partir de los datos de la actuación [*data of performance*]» (Chomsky 1975b, p. 25). Y Quine apoya explícitamente esta idea:

Estoy de acuerdo con Chomsky en ver al aprendiz como un teórico. Ésa es la manera de ver tanto al lingüista de campo como al niño que aprende. Es instructivo intelectualizar la conducta del niño. Está recogiendo datos y conjeturando, aplicando la inducción. La inducción deliberada es más de lo mismo; simplemente que somos más conscientes de lo que estamos haciendo y más refinados en cuanto a cómo lo hacemos. (Quine 1990, p. 291).

Así pues, en cuanto teoría lingüística – y parece haber consenso en esto³⁶ –, las afirmaciones de los hablantes se hallan infradeterminadas, debido a las diferencias en

³⁶ En los autores que nos interesan ahora, ya que el debate – en el que no podemos profundizar en el presente trabajo – ha dado de sí: «Se ha escrito más, mucho de ello exasperantemente vacío, sobre las

el proceso de aprendizaje de la lengua, lo que ocasiona que una teoría empírica no disponga de una materia objetiva, con lo que se genera la indeterminación del significado. El movimiento de los críticos de la indeterminación será negar el carácter de teoría lingüística de esas afirmaciones del nivel básico; no rechazando su idoneidad como datos lingüísticos, sino al contrario: considerando que no constituyen una teoría sobre el lenguaje, sino que conforman más bien la materia de estudio del lenguaje.³⁷

Así, se nos conmina a abandonar la analogía entre el niño y el lingüista, es decir, a dejar de pensar en los juicios lingüísticos que llegan a hacer los hablantes en el curso del aprendizaje de su lengua como creencias o hipótesis formadas a la luz de la evidencia y seleccionadas en preferencia a las alternativas (aun de manera inconsciente) – remitimos, nuevamente, a D. Bar-On 2004 – . El niño y el teórico del lenguaje realizan, pues, tareas diferentes, mientras que la visión del lenguaje de Quine y Davidson

supone que el niño que aprende una lengua y el teórico del lenguaje se hallan en la misma situación, afrontan la misma tarea y se hallan condicionados por la misma evidencia. Pero esto es falso. El lingüista, al contrario que el niño, no está intentado adquirir un lenguaje mediante la observación de los datos, sino intentando, más bien,

confusiones acerca del concepto de competencia y conocimiento-como-competencia que casi sobre cualquier otro tema de la filosofía actual» (Nelson 1978, p. 339).

³⁷ La estrategia de Chomsky es otra; Chomsky niega que la infradeterminación conduzca en lingüística, al contrario que en las ciencias naturales, a indeterminación por su apelación al mentalismo. «Hechos mentales» constituirían la auténtica realidad semántica, éstos determinarían el significado. Las teorías dedicadas a describirlos se hallarían infradeterminadas, pero ahora no hallaríamos más infradeterminación que la habitual en las teorías de las ciencias naturales. Sobre la relación entre el mentalismo y la indeterminación, v. 2.5.

explicar cómo puede el niño haber adquirido su competencia lingüística sobre la base de la evidencia de la que *él* [el niño] disponía. (B. C. Smith 2006, p. 394).

Es preciso subrayar aquí que la teoría semántica davidsoniano-tarskiana (v. 1. 4) debe ser vista como una descripción teórica de lo que se aprende al aprender una lengua, no como la teoría que *consciente y explícitamente* se diseña y aprende. Davidson afirma que la habilidad de alguien para interpretar o hablar con otra persona consiste en «la habilidad que le permite construir una teoría correcta, esto es, convergente, para las transacciones lingüísticas con esa persona» (Davidson 1986, p. 445), ahora bien, vimos en 1.6 la distinción davidsoniana entre la «teoría de paso» y «teoría de entrada». Si recordamos, la teoría de entrada del intérprete supone la descripción de «cómo se halla preparado en un principio [*in advance*] para interpretar una emisión del hablante» (Davidson 1986, p. 442), y la teoría de entrada del hablante describe «lo que él *crea* que es la teoría de entrada del intérprete» (Davidson *ibid.*, p. 442); esta distinción, en especial la teoría de paso, hace hincapié en la capacidad interpretativa más como una habilidad que como el conocimiento de una teoría, dado que la supuesta teoría es modificada para cada hablante en cada momento. Así, Davidson afirma:

Es sólo empleando conceptos como palabra y oración que podemos ofrecer una descripción sistemática de aspectos lingüísticos de la conducta y aptitudes lingüísticas. (...) El punto principal del concepto de un lenguaje (...) es permitirnos dar una descripción coherente de los hablantes, y de lo que los hablantes y sus intérpretes saben que les permite comunicarse. (Davidson 1992a, p. 109).

Ésta parece ser también la idea de Quine cuando explica la analogía entre niño y teórico del lenguaje como una intelectualización de la conducta del niño. Recordemos su explicación conductista de qué implica entender un lenguaje (en 1.3). Asimismo, Quine ha distinguido entre «acomodarse» (*fit*) a una regla y seguirla (*follow*). Así, un planeta «se acomoda» a las leyes de Kepler, pero no tiene sentido decir de él que las obedece o sigue:

Acomodarse es una cuestión de descripción verdadera; guiarse es una cuestión de causa y efecto. La conducta se *acomoda* a una regla siempre que se conforma a ella, siempre que la regla describe correctamente la conducta. Pero la conducta no es *guiada* por la regla a menos que el sujeto conozca la regla y la pueda aseverar. Este sujeto obedece la regla. (Quine 1970b, p. 215).

Consideraciones semejantes hace Devitt cuando habla de objetos que se comportan *como si* su conducta se hallara gobernada por alguna regla, sin que ello implique que éstos conocen la regla o que ésta se halla «implementada» en ellos de alguna manera (v. Devitt 2006, p. 45). Ésta parece ser la situación en el caso de la conducta verbal: el lenguaje consiste en una habilidad que puede ser descrita mediante teorías lingüísticas que el hablante no conoce ni se hallan implementadas en él. (Volvemos sobre ello en 3.4.3).

En todo caso, al rechazar la analogía entre niño y lingüista se busca enfatizar que los juicios lingüísticos de los hablantes no deben ser considerados una *folk linguistic*

theory. Ahora bien, si estos juicios no constituyen hipótesis de una teoría lingüística de legos, ¿qué constituyen?

D. Bar-On (2004) los considera a medio camino entre las meras disposiciones conductuales o aptitudes prácticas, por un lado, y creencias o hipótesis teóricas, por otro. Vendrían a ser, afirma, como los juicios perceptuales que hacemos, por ejemplo, cuando vemos una cierta configuración de líneas en una hoja de papel como un pato una vez y un conejo otra. Mientras que tales juicios se hacen sistemáticamente en respuesta a tipos particulares de estimulaciones o emisiones, aprender a hacerlos es, se afirma, más que adquirir una disposición conductual discriminatoria, o una capacidad práctica para realizar alguna tarea.

Tras negar el carácter de pseudo-hipótesis lingüísticas, el siguiente paso consiste en considerar dichos juicios la materia de estudio de la semántica. Así, se insiste en que es equivocado considerar los juicios de los hablantes acerca de la semántica, por asistemáticos e irreflexivos que resulten, como formando una teoría semántica popular. Al contrario que la química popular, la biología popular, o incluso la así llamada psicología popular, estos juicios no representarían colectivamente un conjunto de creencias y semiexplicaciones *sobre* un dominio dado. Más bien, constituirían parcialmente un dominio: el dominio para la semántica empírica. Lo que Quine llama semántica «intuitiva» o «empírica», pues, no sería una teoría ingenua y errada, como la teoría del flogisto, que debería ser reemplazada por una teoría del lenguaje apropiadamente purgada, puesto que no debería recibir la categoría de «teoría».

Dos son las cuestiones de fondo aquí. En primer lugar, ya que se niega la analogía entre el lingüista y el niño y se pretende ubicar los juicios semánticos de los hablantes a medio camino entre «las meras disposiciones conductuales» y las «creencias o hipótesis teóricas», se debería ofrecer una buena caracterización de ellos. En especial, se debería especificar cuál es el papel que dichos juicios poseen en el proceso de aprendizaje de la lengua. En segundo lugar, debemos plantearnos por qué una teoría lingüística debe tratar sobre las consideraciones de los sujetos no expertos en lingüística mientras una teoría química o biológica, obviamente, no tienen por qué tomar en consideración las creencias de los no expertos en la materia. Es decir, ¿qué implica exactamente «la familiar idea de que los hechos semánticos dependen de nosotros de una forma en la que los hechos de la química o la biología no»? (D. Bar-On 1997, p. 101). Las dos siguientes secciones discuten respectivamente ambos puntos.

Resulta útil, no obstante, hacer alguna consideración antes al respecto. Debemos tener en cuenta que la concepción del niño como lingüista tiene un papel completamente diferente en el ámbito davidsoniano-quineano y en el chomskiano. Respecto al primero, como ya se ha dicho, se subraya el que una teoría tarskiana supone una buena descripción de lo que el niño aprende – de la habilidad que el niño desarrolla, sería mejor decir –, mientras que en Chomsky, la analogía del niño con el lingüista pretende reflejar la idea de que los hablantes poseemos, de manera innata, un conjunto de reglas gramaticales en la mente. Y esta idea chomskiana conduciría a una concepción

cartesiana de la autoridad de la primera persona: ya que las reglas se hallan en la mente, resulta lógico pensar que la primera persona posee un acceso privilegiado a las mismas. Como dice Devitt:

Ésta [la chomskiana] es la asunción de que la competencia lingüística de una persona le confiere algún tipo de acceso privilegiado a hechos sobre el lenguaje, acceso que ejemplifica en sus juicios intuitivos. Este conocimiento intuitivo no depende de investigación empírica del mundo, como lo hace su otro conocimiento. Se halla disponible para él simplemente mediante la introspección de los contenidos de su mente. (Devitt y Sterelny 1999, pp. 179-180).

Con todo, aun cuando no se defiende aquí esta idea chomskiana de la gramática como realidad psicológica innata, conviene recordar lo expuesto en 2.5.2: a pesar de lo que se afirma de manera casi ubicua en la literatura sobre la cuestión, indeterminación y antimentalismo son cuestiones diferentes. Una teoría lingüística innata impresa en la mente no podría tampoco especificar de manera unívoca el significado de nuestras palabras. La gramática innata postulada por Chomsky puede resultar relevante para determinar las reglas sintácticas, pero no la referencia de los términos. Como se señaló en el apartado arriba mencionado, ni consideraciones de tipo «externo» – la historia causal, las disposiciones verbales – ni de tipo «interno» – hechos cerebrales, significados estimulativos, pensamientos, deseos, monólogo interior – pueden determinar el significado. Esto hace que, aun cuando admitamos que el conocimiento de una lengua consiste en saber una teoría, no se vea claro cómo ello puede eliminar la indeterminación del significado.

Fodor, por ejemplo, ha apelado a la noción de representación: conocer un lenguaje es poseer una representación de sus reglas. Y ha intentado ofrecer una visión reduccionista de la misma. Algunos autores piensan que éste es el camino en lingüística:

La distinción entre saber lo que algo significa y no saber lo que significa es una distinción trazada dentro de la psicología, no fuera de ella. Por tanto, debemos buscar, y no resistirnos a tales eliminaciones, como la que pretende Fodor, o mejor, tales formas de integrar nuestras explicaciones abstractas con el funcionamiento de nuestros cuerpos. Concluyo que ninguna de las sugerencias de que el conocimiento del significado o de la referencia es psicológicamente irrelevante se encuentra justificado. (J. Higginbotham 1995, p. 133).

Así, si el significado estuviera determinado por algún tipo de conocimiento que poseen los hablantes, obtendríamos infradeterminación, pero no indeterminación:

En la exposición anterior he asumido una actitud realista hacia el significado, la referencia y las condiciones de verdad. Pero creo que lo que he sugerido es compatible con la idea de que nuestras adscripciones de referencia y conocimiento de la referencia se hallan fijadas sólo dentro de ciertos límites. Es incluso compatible con una tesis más fuerte, relacionada con la tesis de la indeterminación de la traducción radical de Quine, de que podríamos incluso ser incapaces de determinar la verdad y falsedad de todas las oraciones del lenguaje de una persona. La tesis más fuerte mostraría que había una cierta

arbitrariedad en nuestra manera de proceder. Podríamos llegar a percatarnos, según avanza el proceso, de la dependencia de nuestra explicación de una forma provinciana de entender nuestras propias actividades, y las actividades de otros seres humanos. Pero el proceso no sufriría por ello [*the enquiry would be none the worse for that*].» (*Ibid.*)

Dos cuestiones deben ser aquí enfatizadas (v., para ambas, 2.5). Por un lado, el hecho de que tampoco una representación mental ni una determinada estructuración neuronal pueden determinar el significado. Por otro lado, el hecho de que los significados, por definición, no pueden constituirse de manera privada, y esto es ya suficiente razón para la indeterminación, no la mera infradeterminación:

Ésta [la infradeterminación] no es sólo infradeterminación (los datos no son suficientes para determinar unívocamente la teoría, una situación omnipresente en la ciencia), sino que, gracias a la tesis MMM [*man-made meaning*, v. 2.5.], donde los datos dejan un vacío – donde lo que se ha producido mediante la interacción entre los miembros de la sociedad – no hay nada más sobre lo que estar acertado o equivocado, así, obtenemos *indeterminación* y no sólo *infradeterminación*. (Føllesdal 1995, p. 67).

Con todo, se ha intentado conjugar la apelación al conocimiento de los hablantes de los significados con la idea bautizada por Føllesdal como «MMM». Se trata de considerar los juicios lingüísticos de los hablantes – sus intuiciones lingüísticas – como expresando ese conocimiento y, por tanto, como la auténtica materia de estudio de la semántica. Ahora bien, esas afirmaciones, como tales, son públicas; el carácter público del lenguaje y la autoridad de la primera persona entrarían así en contacto. Se defiende,

pues, que «no puede haber más hechos semánticos del castellano que lo que el conjunto de los hablantes del castellano pueden hacer de ellos.» (D. Bar-On 1997, p.102).

Estaríamos aquí ante una nueva versión de la teoría «de la teoría» bautizada como «verificacionismo semántico», una versión que mantiene una difícil relación con el externismo:

El VS [el verificacionismo semántico] puede ser visto como directamente viciado por la concepción externista del significado que es adoptada por una parte sustancial de naturalistas semánticos (Devitt incluido). La idea central del externismo (tal y como se presenta en Putnam en (1975)) es que lo que se halla puramente «en las cabezas de los hablantes» (como contenidos descriptivos o definiciones analíticas) no puede determinar la *extensión* de términos como ‘oro’. Los ejemplos de Putnam de Tierra Gemela intentan mostrar que, *en tanto que pensemos en el significado como lo que determina la referencia, los significados no están en la cabeza*. Pero el verificacionismo semántico insiste en que los hechos sobre los significados de las expresiones dependen de lo que los hablantes de la lengua *juzgan* que son esos hechos. Y esto puede verse como inconsistente con el externismo. (D. Bar-On 1997, p. 104)

3.3. Intuiciones y externismo

Desde una visión davidsoniana, la respuesta que se ofrece a la idea de que los juicios semánticos de los hablantes resultan constitutivos del significado es una exposición del proceso mediante el que se adquieren los significados. La exposición de este proceso

servirá aquí para mostrar cómo *los significados de nuestras palabras se constituyen mediante un proceso en el que no entran en consideración semejantes juicios.*

3.3.1. Interpretación y triangulación

Es, según Davidson, mediante un proceso de *triangulación* como aprendemos el lenguaje – y es también mediante este proceso como nos convertimos en seres pensantes – . Merece la pena citar *in extenso* la descripción davidsoniana de este proceso:

Expuesto en términos enormemente simplificados, un aspecto básico de tal aprendizaje puede ser descrito de esta manera: el aprendiz es recompensado... cuando hace sonidos o responde, de cualquier manera, de una forma que el maestro encuentra apropiada en situaciones que el maestro clasifica juntas. El aprendiz, subsiguientemente, hace ruidos similares a causa de situaciones que el aprendiz clasifica juntas instintivamente... el éxito en el primer nivel se consigue hasta el punto en el que aprendiz responda con sonidos que el maestro encuentra similares a situaciones que el maestro encuentra similares. El maestro responde a dos cosas: la situación externa y las respuestas del aprendiz. El aprendiz responde a dos cosas: la situación externa y las respuestas del maestro. Todas estas relaciones son causales. Así se forma el triángulo esencial que hace posible la comunicación sobre objetos y eventos compartidos. Pero también es este triángulo el que determina el contenido de las palabras y pensamientos del aprendiz cuando éstos se vuelven lo suficientemente complejos para merecer el término... lo que hace el aspecto particular de la causa de las respuestas del aprendiz el aspecto que les da

el contenido que tienen, es el hecho de que este aspecto de la causa es compartido por el maestro y el aprendiz. (Davidson 1991, pp. 197-198).³⁸

Esto es lo que nos interesa subrayar aquí: este proceso no sólo describe *cómo* se aprenden los conceptos, sino en *qué* consisten estos – cuál es su contenido Intencional, en terminología searleana – . El proceso de triangulación «no es sólo una historia sobre cómo aprendemos a usar las palabras: debe de ser también una parte esencial de una explicación adecuada de a qué se refieren las palabras, y qué significan.» (Davidson 1988, p. 44). Y, además, Davidson no limita la dependencia entre la situación externa y el contenido de los conceptos y las creencias a los tipos de conceptos que había tratado Putnam, los de clases naturales:

[H]ay una fuerte tendencia a suponer que la dependencia está limitada a los tipos de expresiones que aparecen en los ejemplos: nombres propios, palabras de clases naturales como ‘agua’ y ‘oro’ y los indéxicos. Pero, de hecho, el fenómeno es ubicuo, ya que es inseparable del carácter social del lenguaje. No es un problema local que ha de ser resuelto mediante algún truco semántico; es un hecho perfectamente general sobre la naturaleza del pensamiento y el lenguaje. (Davidson 1988, p. 48).

³⁸ La concepción davidsoniana de este proceso es claramente deudora de la de Quine, quien ya en *Palabra y Objeto*, a pesar de su apelación a las terminaciones nerviosas, subrayaba la importancia del medio compartido (v., por ejemplo, pp. 1, 31), y de manera más insistente en Quine 1969a (v. p. 28). Sobre esta relación dice Føllesdal: «Lo que me llama la atención es que Davidson ha pasado de hacer caso omiso de la percepción a abrazar enteramente el argumento de la triangulación y explorarlo en detalle, pero Quine parece haberse desplazado hacia la dirección opuesta, alejándose del énfasis en los objetos socialmente observables para dar prioridad a las estimulaciones.» (1995, p. 59).

Comparemos esta visión con una de corte internista, como la de Locke. Según Locke, las palabras se refieren a ideas en la mente y enseñar a hablar a un niño consiste en conseguir que las ideas en la cabeza del niño se sincronicen con las del resto de hablantes. Ahora bien, estas supuestas ideas parecen realizar una escasa labor, dado que el elemento esencial en la comunicación lo constituyen las circunstancias y objetos externos, observables tanto por el maestro de la lengua como por el aprendiz. Y, a este respecto, Davidson señala que

[e]n los casos más simples y más básicos, las palabras y las oraciones derivan su significado de los objetos y circunstancias en cuya presencia fueron aprendidas. Una oración a la que uno ha sido condicionado mediante el proceso de aprendizaje a considerar cierta ante la presencia de fuegos, será (usualmente) verdadera cuando hay un fuego presente; una palabra a la que uno ha sido condicionado a considerar aplicable ante la presencia de serpientes, se referirá a serpientes. (Davidson 1988, pp. 44-45).

La descripción de Davidson del proceso de triangulación, con toda su simplicidad, permite extraer conclusiones de relevancia filosófica, y aquí nos interesa extraer la del antiinternismo: las representaciones, ideas, juicios, etc. de carácter privado no pueden formar parte del significado de las palabras. El significado de una expresión no puede «contener nada que no sea manifiesto en el uso que se hace de ella, hallándose sólo en la mente del individuo que aprehende ese significado» (M. Dummett 1978, p. 216.). Que el lenguaje se aprenda y transmita mediante el proceso triangular descrito – y no mediante el contacto de mentes, con sus contenidos Intencionales – , debe hacernos

desechar, pues, la apelación a toda entidad interna como determinante del significado de nuestras palabras: éste «es determinado por los elementos terminales en el proceso de condicionamiento, y es testado sólo por el producto final» (Davidson 1988, p. 44).

Este «antiinternismo semántico» defendido por Davidson resulta tan decidido que le lleva a rechazar, también, los significados estimulativos de Quine – la versión materialista de las ideas lockeanas – . Para Davidson, Quine, al apelar a entidades internas al sujeto, aun cuando se trate de entidades físicas y no mentales, ha producido una teoría «esencialmente de primera persona y cartesiana» (Davidson 1990c, p. 194). En la visión davidsoniana del lenguaje hallamos, pues, una versión más coherente del carácter esencialmente público de la semántica. Cuando Sam observa que su maestro emite o está dispuesto a emitir la oración *S* sólo cuando hay una serpiente visible para él, puede inferir que su oración *S* significa que ve una serpiente. La causa de que el maestro de Sam emita *S* es la presencia de la serpiente, pero también, señalaría Quine, el conjunto de terminaciones nerviosas activadas en su nervio óptico en ese momento. Esta actividad nerviosa tiene, por consiguiente, derecho a ser considerada la causa de que el maestro de Sam emita *S*. ¿Qué lleva entonces a Davidson a rechazar la apelación al estímulo próximo? La respuesta se halla en el proceso de triangulación: es sólo en presencia de serpientes vistas conjuntamente como el maestro puede premiar la emisión de *S* por parte del aprendiz. Existen intermediarios causales – ondas sonoras emitidas por cuerdas vocales, estimulación del nervio óptico, etc. – , pero no tienen ningún papel en la adquisición del lenguaje por parte de Sam ni para determinar el significado de las emisiones.

Al hilo de lo dicho se pueden hacer algunas consideraciones acerca del supuesto *explicitismo* de Quine y Davidson. Como hemos visto, Alston y Searle fundamentaban su crítica a la indeterminación en el hecho de que parte del explicitismo semántico (v. 2.3). Se criticaba allí que esta idea supone una asunción, «una preferencia metafísica y, creo, una preferencia que no está garantizada por los hechos» – se trata, además de una «metafísica profundamente antimentalista» – . La exposición davidsoniana del proceso de aprendizaje de la lengua como un proceso de triangulación hace patente que el explicitismo no constituye una asunción metafísica, sino la descripción de cómo interiorizamos los significados. Podemos considerar esta exposición una forma más precisa de exponer el hecho de que el lenguaje es esencialmente una institución pública. Podemos señalar dos razones acerca de por qué debemos adoptar una postura explicitista respecto al lenguaje, razones ambas que la triangulación davidsoniana pone de manifiesto:

- a) Razón de hecho. De hecho, aprendemos el lenguaje atendiendo a la conducta de nuestros maestros y al medio que nos rodea. Telepatía o ciencia infusa no entran en juego. Es un hecho que no aprendemos el lenguaje interconectando mentes.
- b) Razón conceptual. Utilizar un lenguaje consiste en comunicarse con otro acerca de cosas. Ambos, hablante y oyente, deben tener acceso a estas cosas para comenzar el proceso comunicativo. Esto es, el hecho de que el proceso de aprendizaje del lenguaje no se lleve a cabo mediante telepatía u otro tipo de

interconexión de mentes, no es contingente, peculiar del ser humano, sino un hecho esencial de cualquier aprendizaje de un lenguaje.³⁹

Estas reflexiones, aun en su brevedad, deben apuntalar lo dicho hasta aquí y, sobre todo, hacer ver que, en todo caso, la carga de la prueba recae sobre los «implicitistas». A falta de argumentos convincentes, la visión explicitista se hace forzosa.

3.3.2. Indeterminación e interpretación

Como hemos visto, los significados – el contenido de nuestras creencias y demás actitudes proposicionales – se determina en un proceso donde no entran en juego las consideraciones de los hablantes acerca de la propia noción de significado o similares. La idea de Davidson, como se ha visto en 1.5, es que las nociones de significado y referencia, o incluso de satisfacción o verdad – aunque ésta constituye el concepto semántico más básico – no son más que postulados de nuestras teorías lingüísticas que resultan útiles para explicar la única realidad objetiva: el uso de las oraciones. Se debe insistir, pues, en que

lo que se halla abierto a la observación es el uso de las oraciones en contexto, y la verdad es el concepto semántico que mejor comprendemos. La referencia y las nociones semánticas relacionadas, como la satisfacción, son, en comparación, conceptos teóricos (como lo son las nociones de término singular, predicado, conectiva oracional y el

³⁹ Esta diferenciación se basa en la crítica de P. Pagin al proceso de triangulación tal y como Davidson lo describe, a saber, que se trata de una mera especulación acerca de cómo se aprende el lenguaje, «pura especulación empírica de sillón», afirma. Davidson responde: «Ya que no me entusiasma la distinción entre analítico/sintético, tampoco hallo atractiva [*neither am I sharp on*] la distinción entre mera-especulación-empírica-de-sillón/especulación-empírica.» (Davidson 2001c, p. 293.)

resto). No hay cuestión acerca de la corrección de estos conceptos teóricos más allá de la cuestión de si ofrecen una explicación satisfactoria del uso de las oraciones. (Davidson 1990a, p. 300).

Así, volviendo al proceso de triangulación, Sam se percató de que su maestro aplica o está dispuesto a aplicar ‘serpiente’ siempre que hay serpientes presentes, Sam ha interpretado entonces el concepto de serpiente y comenzará a aplicar el término de manera similar. Para ello, dos circunstancias se han hecho necesarias: que Sam percibiera como similares las diferentes emisiones de ‘serpiente’ y las diferentes apariciones de serpientes.⁴⁰ La cuestión ahora es que el intérprete puede utilizar diferentes conjuntos de hipótesis analíticas sin que el proceso de triangulación determine unívocamente una de ellas. La evidencia disponible, como hemos visto, se halla constituida por las oraciones – no términos – que el intérprete asocia con situaciones compartidas por el intérprete y el interpretado, la estructura de las oraciones – esto es, nociones como referencia, satisfacción, verdad – es la teoría utilizada por el intérprete para sistematizar la evidencia (las oraciones que el interpretado considera verdaderas ante determinadas circunstancias).⁴¹ Darse cuenta de que la indeterminación surge en este contexto de confeccionar una «teoría» acerca de la conducta verbal de un sujeto, y no considerarla como algo que infecta el aprendizaje de la lengua, evita insistir

⁴⁰ Dado que intérprete e interpretado han de fijarse en los mismos objetos para que el proceso se ponga en marcha, Davidson hace notar que el «trabajo de condicionamiento funciona óptimamente con objetos que interesan al aprendiz y son difíciles de pasar inadvertidos tanto para el maestro como para el alumno.» (2001, p. 44). Davidson ha seguido aquí a Quine en la apelación a estándares de similitud innatos; así, ha hablado de los objetos que el maestro y el aprendiz «encuentran natural agrupar juntos» (2001, p. 118).

⁴¹ Este punto, crucial en la visión davidsoniana del proceso interpretativo (v. 1.5), aparece ya en Quine (v. 1.2). Quine recuerda que la indeterminación se dirige a las oraciones porque, como ya señaló Frege, son éstas las portadoras de significado: «[M]i conjetura de la indeterminación de la traducción no versaba sobre términos como ‘gavagai’ sino sobre oraciones como un todo, ya que sigo a Frege en considerar las oraciones como los vehículos primarios del significado» (Quine 2000, p. 419).

en que el sujeto aprendió a decir ‘conejo’ en presencia de conejos, no ‘parte no separada de conejo’, idea que vemos aparecer, por ejemplo, aquí:

Así, por ejemplo, cuando un niño francés aprende ‘rouge’, aprende que ‘rouge’ se aplica a cualquier objeto x ssi x es rojo. No aprende que ‘rouge’ se aplica a x ssi x es el color de las frambuesas, o que ‘rouge’ se aplica a x ssi [x es rojo \vee [o la nieve es blanca o la nieve no es blanca]]. (G. Segal 2006, p. 207).

Así, podríamos afirmar que saber que el interpretado aplica la oración de una palabra ‘conejo’ a los conejos ya es saber todo lo que se tiene que saber; caracterizar a ‘conejo’ como un término con la referencia ‘conejo’ o la de ‘parte no separada de conejo’ son caracterizaciones posibles de nuestra conducta realizadas desde el metalenguaje – esto es, desde la teoría – .

Para ilustrar este punto, Davidson ha recurrido a la analogía con los sistemas de medición. Reproducimos sus palabras:

Si se tienen los axiomas que determinan algún sistema de medida, sea de peso, temperatura o probabilidad subjetiva, se pueden representar las estructuras así definidas en números de infinitas maneras. (...) Con el peso, un número positivo arbitrariamente escogido se asigna a algún objeto particular; relativamente a esa asignación, los números que miden los pesos de todos los demás objetos se hallan fijados. Se obtiene una manera de determinar los pesos igualmente buena multiplicando las figuras originales por cualquier constante positiva. (...) Así como infinitos conjuntos de números nos permiten determinar las mismas complejas estructuras del mundo, nuestras oraciones puede ser

usadas de infinitas maneras para determinar las actitudes de otros, y lo que quieren decir. (Davidson 1999b, p. 596).

Así, el sistema métrico de pesos se relaciona con el imperial por una simple transformación lineal, y lo mismo rige para las escalas de temperaturas Celsius o Fahrenheit: para una temperatura de n grados Fahrenheit, obtenemos la correspondiente en grados Celsius sustrayendo 32 y dividiendo por 1'8. Estas operaciones de transformación reflejan el hecho de que el número que una escala de temperatura asigna a un objeto constituye únicamente la manera de determinar su posición en un sistema de relaciones físicas. Y lo mismo rige para los significados; un comentarista de Davidson lo expone de la siguiente manera:

Análogamente, un manual de traducción codifica las posiciones relativas de las oraciones en un sistema o estructura de oraciones utilizando las oraciones del intérprete; de una traducción de «¡Gavagai!» por M_1 , por tanto, derivamos una traducción correspondiente según M_2 si correlacionamos las ocurrencias de los términos «conejo» y «partes de conejo no separadas». Así, pues, M_1 y M_2 son esquemas diferentes pero equivalentes para representar una estructura única, invariable de las oraciones nativas; y si según M_1 , la emisión de un hablante de «¡Gavagai!» en el tiempo t es verdadera, entonces, esa emisión es también verdadera según M_2 , así como si algo es n grados Fahrenheit en t , entonces también es $(n - 32)/1'8$ grados Celsius en t . (M. Joseph 2004, p. 73).

Así, cuando Davidson discute los argumentos que se suelen esgrimir contra la tesis de la indeterminación, menciona que alguien podría afirmar que ‘Roma’ no *significa* ‘punto situado a 100 millas de Roma’ (ejemplo visto en 2.3), y según Davidson «la respuesta adecuada es: las palabras individuales no tienen significados. Tienen un rol para determinar las condiciones de verdad de las oraciones (...)». Y dado que este rol es igualmente captado por diferentes adscripciones de significado, no tiene sentido hablar del significado correcto. La idea davidsoniana de comparar la relación de referencia con propiedades como el peso o la temperatura se halla ya, al menos insinuada, en Quine cuando afirma que «[l]a estructura es lo que importa a una teoría, y no la elección de sus objetos» (Quine 1981, p. 20). En efecto, lo llamativo de las diferentes maneras de pesar o medir los cuerpos es que, utilizando diferentes valores numéricos, todas muestran las mismas relaciones entre los cuerpos. Por tanto, a la hora de juzgar la corrección o no de una escala de medición, los valores utilizados resultan irrelevantes, lo esencial es que la estructura de la escala consiga, por ejemplo, que los cuerpos que pesan lo mismo tengan asignado el mismo valor. Mühlhölzer afirma (1995, p. 219) que esta idea no sólo ha sido defendida por Quine o Davidson, podemos hallar formulaciones más o menos explícitas en Frege, Schlick, Russell, Carnap, Weyl y Putnam. Hermann Weyl, por ejemplo, la expresa así:

Una ciencia sólo puede determinar su dominio de investigación respecto a un mapeo isomórfico. En particular, le resulta bastante indiferente la ‘esencia’ de sus objetos. (...) La idea de isomorfismo demarca el autoevidente límite inalcanzable de la cognición. (Citado en Mühlhölzer 1995, p. 219).

Por otro lado, según Davidson, creer que la indeterminación simplemente deja las palabras sin significados, viene a ser como creer que la diferencia entre decir que un lago mide una milla de largo o 1.852 metros es una diferencia del lago. Resulta esencial subrayar aquí dos avances en la explicación davidsoniana de la indeterminación respecto a la quineana:

A) La insatisfactoria apelación quineana a lenguajes de fondo se sustituye por la analogía con las diferentes escalas de medición. Davidson ve claro que relativizar la referencia no ofrece una respuesta adecuada: «La relativización no puede fijar la referencia; nada puede. (...) [L]a tentación de relativizar debe resistirse; no hay nada a lo que relativizar la satisfacción o la referencia.» (Davidson 1995a, p. 14). El lenguaje de fondo plateaba el problema del regreso al infinito, la analogía davidsoniana hace patente que en el proceso interpretativo no se trata tanto de asignar significados a palabras como de sistematizar la conducta verbal de los hablantes:

Visto de esta manera, hablar de cuánto pesan las cosas es relacional: relaciona objetos con números, y también unos objetos con otros. Pero nadie supone que los números son de alguna manera intrínsecos a los objetos que tienen peso, o son, de alguna manera, 'parte' de ellos. Lo que es básico son algunas *relaciones* entre objetos: rastreamos [*keep track*] cómodamente estas relaciones asignando números a los objetos y recordando cómo están reflejadas en los números las relaciones entre los objetos. (Davidson 1989a, p. 59).

Por decirlo de otra manera: la diferencia entre que un campo mida una yarda o 36 pulgadas, no es una diferencia en el campo mismo (Davidson 1989a, p. 65). Y, ahora, la analogía con los significados o las creencias es clara. Cuando decimos que dos personas o dos palabras quieren decir lo mismo no nos comprometemos con significados reificados o con conceptos en o delante de la mente. Sólo estamos caracterizando una relación entre dos hablantes o los criterios de aplicación de dos términos. Caracterización que podría hacerse de diferentes formas:

Al pensar y hablar de los pesos de los objetos físicos no necesitamos suponer que hay tales cosas como los pesos, cosas que tienen los objetos físicos. De manera similar, al pensar y hablar sobre las creencias de la gente no necesitamos suponer que hay tales cosas como las creencias. Ni hemos de inventar objetos que sirvan como los ‘objetos de creencia’ o lo que está ante la mente o en el cerebro. (Davidson 1989a, p. 60).

Y si no hay tales cosas como los significados, la apelación al conocimiento inmediato de éstos por parte de la primera persona resulta más compleja que en teorías como la searleana (v. 3.6):

Hay aún una objeción que alguna gente siente como fatal para la tesis de la indeterminación. La afirmación de que la indeterminación mina la autoridad de la primera persona. Esta gente suele decir «quizá los criterios empíricos de Quine no pueden excluir el uso de métodos alternativos de interpretación para otros, pero yo sé que quiero decir Roma con mi palabra ‘Roma’ y no el área 100 millas al sur de Roma». (Davidson 2001, p. 79)

Davidson cree que esta afirmación «es correcta, pero no entra en conflicto con el indeterminismo». Y la razón es que la única forma que tiene un hablante de hacer explícita la referencia de sus términos es el desentrecomillado; si un hablante se autointerpreta no irá más allá de la afirmación de que con ‘Roma’ quiere decir Roma. Es decir, «las interpretaciones de la primera persona se hallan necesariamente ligadas al manual de traducción homofónico (lo que es lo mismo que decir que aquí no tiene lugar traducción o interpretación)» (Davidson 1997, p. 80). En autoatribuciones de este tipo, los hablantes no hacen más que parafrasearse en el metalenguaje, pero «sabemos que las palabras en el metalenguaje tienen los mismos papeles que estas palabras tienen en el lenguaje objeto, ya que el lenguaje objeto y el metalenguaje son el mismo» (*ibid.*, p. 80).

Es importante subrayar aquí la razón por la que Davidson puede apelar al ascenso semántico sin caer en la regresión al infinito que vimos en Quine; se debe a que, como vemos, concibe el manual de traducción homofónico como sinónimo de que los hablantes no se interpretan a sí mismos. Como hemos visto en 3.1.1, los hablantes simplemente aplican sus términos de la manera en que aprendieron mediante un proceso de triangulación – y no hay más, especialmente no hay «juicios semánticos de nivel básico» gracias a los cuales aprender la lengua – . Y repetimos que la semántica davidsoniano-tarskiana supone sólo una exposición de lo que se aprende en ese proceso de triangulación, no lo que se aprende.

B) Para Davidson – y, en lo que aquí nos interesa, constituye esto el mayor mérito de la analogía entre la indeterminación y los sistemas de medida – las interpretaciones pueden ser diferentes sin ser por ello, necesariamente, incompatibles; desaparece, así, el debate acerca de qué sentido darle a la incompatibilidad de las diferentes interpretaciones (v. 2.6.1 y 3.1.1). En particular, pierde su sentido la apelación a una semántica intuitiva, contra una empírica, en el debate acerca de la infradeterminación. Veamos las palabras exactas de Davidson:

Otra virtud de la analogía es que deja claro por qué Quine está justificado cuando insiste, contra Chomsky y otros, en que la indeterminación de la traducción es diferente de la infradeterminación de las teorías por toda la evidencia posible. Está en cuestión, seguro, si la infradeterminación en el sentido de Quine es posible, ya que requiere que haya teorías empíricamente equivalentes pero incompatibles [Aquí nota omitida donde se avisa de que Quine ha cambiado de opinión al respecto en varias ocasiones]. La indeterminación no es eso: las teorías empíricamente equivalentes que se aceptan como igualmente correctas para entender a un agente no son incompatibles, no más de lo que el peso en libras o en kilos supone teorías incompatibles del peso. (Davidson 1997, p. 76).

Vemos, así, en la exposición davidsoniana del proceso de triangulación, la idea de que aprender un lenguaje es aprender unos patrones de comportamiento lingüístico, es decir, se trata de aplicar palabras de forma coherente, siguiendo el patrón que se ha percibido en el hablante que oficie de maestro, y es este patrón el que le da a las palabras el contenido que éstas adquieren. Estos patrones vienen a constituir, pues, la «materia objetiva» de la semántica, y es el hecho de que sean descriptibles en el lenguaje del

intérprete recurriendo a diferentes conceptos lo que genera la indeterminación. (Se vuelve sobre este punto en 3.4.1).

Esta idea de los patrones de comportamiento lingüístico como las «invariantes» que constituyen la materia que debe conservarse en las diferentes interpretaciones tiene también la virtud de responder a las dudas acerca de respecto a qué se halla indeterminada una traducción/interpretación. Efectivamente, hemos visto en Quine cómo la indeterminación lo era de las teorías semánticas respecto a las teorías físicas, y esto hace surgir ciertas dudas:

¿Qué es para un conjunto de afirmaciones determinar otro? No es que las afirmaciones del segundo conjunto sean consecuencias lógicas de las afirmaciones del primero. Ciertamente, las teorías de traducción no son consecuencias lógicas del conjunto de todas las verdades físicas. Pero eso es trivial, ya que siempre que una teoría empírica de cualquier interés incluye vocabulario que no se encuentre en las verdades de la física no podrá ser una consecuencia lógica de esas verdades. Por ejemplo, no todas las verdades de la biología son consecuencias lógicas del conjunto de oraciones verdaderas del lenguaje de una física ideal. (Soames 1999, p. 332).

Efectivamente, Quine hacía depender la indeterminación de su fisicalismo:

Defendí [respecto a la traducción radical] que las traducciones se hallarían indeterminadas, en el caso de oraciones a una distancia considerable de las oraciones observacionales. Se hallarían indeterminadas en este sentido: dos traductores podrían desarrollar manuales de traducción independientes, ambos compatibles con toda la conducta verbal y todas las

disposiciones de la conducta verbal, y aún así un manual ofrecería traducciones que otro traductor rechazaría. Mi posición era que cualquiera de los manuales podría ser útil, pero respecto a cuál era correcto y cuál equivocado no había materia objetiva.

Mi objetivo ahora no es defender esta doctrina. Mi objetivo es, simplemente, hacer patente que hablo como un fisicalista al decir que no hay materia objetiva. Quiero decir que ambos manuales son compatibles con la disposición [*fulfillment*] de exactamente los mismos estados físicos elementales mediante las regiones de espacio-tiempo. (Quine 1977, p. 284).

La idea de Davidson disipa este tipo de dudas, que surgen con la exposición quineana, ya que no hay ahora un conjunto de verdades que determina, o no, con una determinación lógica o de otro tipo, otro conjunto; se trata, en las teorías semánticas o interpretaciones de descripciones de la conducta verbal de los sujetos, más concretamente, de qué oraciones consideran verdaderas en qué circunstancias. Y, si bien Davidson concede que «la actitud de sostener una oración como verdadera, aceptarla como verdadera (...) es, desde luego, una creencia» (1973, p. 135), tengamos en cuenta que las creencias davidsonianas – y, por tanto, la actitud de sostener una oración como verdadera – se hallan muy lejos de los contenidos Intencionales de Searle o ideas semejantes. Una creencia es, claro, una entidad mental; ahora bien, en la filosofía de Davidson, las entidades mentales son sólo una forma más de describir la realidad física – la única existente – (v. nota 51), y esto, la posibilidad de describir un mismo fenómeno de diferentes maneras es, precisamente, lo que genera la indeterminación: «hay una única ontología, pero más de una manera de describir y explicar los elementos en la ontología» (Davidson 1995b, p. 121).

Dos son las conclusiones principales que debemos extraer de lo dicho: a) El proceso de triangulación hace patente que los «juicios semánticos de nivel básico» no tienen función alguna dentro del mismo. La labor de la semántica no es, pues, la sistematización de semejantes juicios; b) Por otra parte, hemos visto cómo dicho proceso hace inevitable la indeterminación, sin que ésta, no obstante, suponga sumir los significados en una vaguedad inaceptable. (Sobre la utilidad de la analogía con los sistemas de medición se vuelve en 3.4.2).

3.3.3. Interpretación sin convenciones

Tras lo dicho, aún se podría apelar a la idea de traducción o interpretación como el proceso de equiparar significados determinados por convención. Ésta parece ser la idea de Searle cuando afirmaba (v. 2.5):

Si alguien me pregunta cómo sé que ‘lapin’ significa ‘conejo’, sólo puedo decir que he hablado francés durante unos cuarenta años, y sé francés lo suficientemente bien para saber que esto es lo que quiero decir cuando uso esta palabra, y que comparto este significado con otros hablantes de francés. (Searle 1997, p. 3).

Así pues, el argumento de Searle no es sólo que yo sé lo que quiero decir con mis palabras, sino que lo sé porque ése es el significado que las palabras tienen en la lengua que hablo. La misma idea, que las palabras de una lengua poseen un significado determinado que es lo que se capta al aprenderla, aparece, por ejemplo, en esta crítica a la visión davidsoniana de la interpretación:

¿Pero dónde, en esta imagen, hay sitio para la comprensión inmediata de las palabras? La experiencia fenomenológica de oír el significado en las palabras de alguien es innegable. No oímos sonidos a los que luego tenemos que darles un sentido. Normalmente, el proceso es inmediato: simplemente oímos lo que alguien está diciendo. Es incluso involuntario. Si alguien habla de forma audible en una lengua que entiendo, no tengo más opción que oír sus palabras como significando lo que significan. (B. C. Smith 1998, p. 417).

Davidson se ha encargado de mostrar, por un lado, el relativo poco peso que poseen las convenciones lingüísticas en el proceso de interpretación (v. 1.6 acerca de la crítica a la existencia de lenguas naturales) y, por otro, cómo la apelación a éstas no consigue, tampoco, eliminar la indeterminación. Estas ideas han sido desarrolladas por Davidson, especialmente, en el marco de sus reflexiones acerca de la metáfora y el uso incorrecto de palabras (malapropismos). Cuando alguien afirma «el hombre es un lobo», ni el significado literal ni las condiciones de verdad de la oración nos permiten captar lo que quiere decir – no son lo que él quiere decir – .

Puede que sonriamos cuando alguien dice «ve delante y nosotros te precederemos», o, con Archie Bunker, «necesitamos unas risas para romper la monogamia». (...) [L]o que es interesante es ese hecho de que en ninguno de estos casos tiene problemas el oyente para entender al hablante de la forma en la que el hablante quiere. (Davidson 1986 p. 434).

La irrelevancia de significados determinados – en la cabeza de un hablante o donde sea – queda patente en el hecho de que interpretamos las metáforas novedosas o los errores lingüísticos. Lo relevante es, pues, la intención de un hablante de comunicar algo:

El significado (...) obtiene su vida de esas situaciones en las que alguien intenta (...) que sus palabras sean entendidas de una cierta manera, y lo son. En estos casos podemos decir sin dudar: cómo pretendía ser entendido, y fue entendido, es lo que él, y sus palabras, significaban literalmente en esa ocasión (...) Así, para mí el concepto de «el significado» de una palabra u oración da lugar a los conceptos de cómo un hablante desea que sean entendidas sus palabras, y de cómo un oyente las entiende. (Davidson 1994, pp. 11-12).

Davidson advierte de que «[n]o hay nada malo en asignarle significados a las oraciones, pero éste debe ser siempre un significado derivado de las ocasiones concretas en las que las oraciones son usadas» (1993c, p. 298). Y debemos tener presente aquí que la indeterminación surge debido a que lo que el hablante deseaba expresar en una situación particular puede ser siempre descrito en el metalenguaje – el lenguaje del intérprete – de maneras diferentes; a veces, incluso asignando diferentes condiciones de verdad a una expresión. La indeterminación se genera aquí, pues, como ya hemos dicho, debido a que una oración del hablante puede ser expresada de diferentes formas en la lengua del intérprete. Lo habitual en el caso de la primera persona es que no disponga de mejor manera de expresar el significado de sus palabras que acudiendo al desentrecomillado:

El hablante, tras utilizar todo el conocimiento y la pericia que posea en la tarea de decir lo que significan sus palabras, no puede «ir más allá» [*cannot improve on*] del siguiente tipo de afirmación: «mi emisión de ‘Wagner murió feliz’ es verdadera si y sólo si Wagner murió feliz. Un intérprete no tiene razón para suponer que ésta será la mejor manera de expresar las condiciones de verdad de la emisión del hablante. (Davidson 1984b, p. 13).

El desentrecomillado ha sido criticado sobre la idea de que resulta trivial si no exigimos, además, el conocimiento de significados por parte del sujeto que lo utiliza. Tal y como expresa Higginbotham (1998) esta crítica, para generar este tipo de esquema sólo se exige que (1) uno pueda utilizar el lenguaje o la parte del lenguaje que se está mencionando y (2) que uno posea el concepto de significado. «Pero no se exige que uno haya dominado los conceptos», así que el conocimiento de un determinado esquema de desentrecomillado – digamos de «‘la nieve es blanca’ significa que la nieve es blanca» – es posible para una persona que no la comprende plenamente, o cuya concepción del significado de ‘La nieve es blanca’ es, quizás, erróneo (J. Higginbotham *ibíd.*, p. 433). El esquema del desentrecomillado, según esta crítica, es compatible con que un individuo no comprenda plenamente los conceptos que aparecen en la oración mencionada. Precisamos la exigencia del conocimiento exacto de significados o referencias y no la asunción de que saber las circunstancias en las que una oración resulta verdadera es todo lo que hay que saber acerca de su significado.

En primer lugar, se debe tener presente que el desentrecomillado aquí sólo trata de describir el proceso de interpretación de una segunda persona (supone, en ese sentido,

una *hiperintelectualización* del proceso de interpretación, utilizando la expresión quineana). En segundo lugar, la idea de Davidson resulta casi obvia: ¿qué mejor manera puede tener alguien de decir a qué se refiere con ‘conejo’ que decir que a los conejos? Puede no ser la única manera de expresar el significado de su palabra, pero sí es, desde luego, la mejor. Ahora bien, podría haber otras formas diferentes de hacerlo en el lenguaje del intérprete, con lo que se genera la indeterminación.

Quine recuerda que sería erróneo considerar que alguien sabe el significado de una expresión sólo cuando sabe expresarlo mediante otras palabras: «Entender una expresión es, diría uno, saber el significado; y saber el significado es, diría uno, ser capaz de dar el significado[.]», pero relata una breve anécdota para llamar la atención sobre lo confuso de esta idea:

Dije a mi hijo pequeño, ‘ochenta y dos. ¿Sabes lo que quiero decir?’ Dijo, ‘no’. Entonces dije a mi hija pequeña, ‘ottantadue. ¿Sabes lo que quiero decir?’ Ella dijo, ‘sí. Ochenta y dos.’ Dije, ‘mira, Margaret entiende el italiano mejor que Douglas el castellano’. (Quine 1975b, p. 247).

Y Quine admite aquí que, si bien se adhiere a la propuesta semántica davidsoniana – «[n]uestro estándar [para comprobar si entendemos una oración] es, aún, éste: da las condiciones de verdad. De ahí el plan de Davidson para la semántica al estilo de la definición de verdad de Tarski» – , tampoco conviene concebir al hablante como poseyendo un *conocimiento* de las condiciones de verdad:

Pero cuando defino la comprensión de una oración como conocimiento de sus condiciones de verdad, no estoy ofreciendo una definición donde detenerse; mi término ‘conocimiento’ es tan pobre para detenerse en él como el mismo término ‘comprensión’.

(*Ibíd.*, p. 249).

En la visión davidsoniana, pues, los significados son constructos teóricos – de la teoría que ofrece una interpretación de las oraciones de un hablante – . Dado que la teoría se halla infradeterminada (diferentes descripciones del significado de una oración son posibles en el lenguaje del intérprete), el significado se halla indeterminado. Con todo, el hablante sabe lo que quiere decir porque sabe a qué objetos y situaciones aplica sus palabras (v. 3.5).

Como veremos (3.5), Davidson acepta la existencia de una asimetría en el conocimiento semántico entre la primera y la tercera persona. Ahora bien, ésta se basará en cómo se desarrolla necesariamente el aprendizaje de una lengua y el proceso de interpretación, consideraciones acerca de «la experiencia fenomenológica» de la comprensión quedan fuera de lugar.

3.4. Antiintuicionismo lingüístico

Vimos que la crítica a la tesis de la indeterminación basada en la apelación a la autoridad de la primera persona derivó en una reivindicación de la relevancia que una teoría semántica debería otorgarle a las intuiciones lingüísticas de los hablantes. Dichas intuiciones, se sugiere, eliminan la indeterminación – los hablantes afirman que con ‘conejo’ se refieren a conejos y no a partes no separadas – . Ahora bien, hemos visto en

la sección anterior que dichas intuiciones no aparecen en el proceso de aprendizaje de la lengua y que éste, por la manera en que se desarrolla, hace inevitable la indeterminación. Ahora bien, aún se podría insistir en que la descripción davidsoniana del proceso elimina dichos juicios – las intuiciones – arbitrariamente. ¿Por qué no considerarlos la materia principal de una teoría semántica y ver, entonces, a qué queda reducida la indeterminación? ¿No resulta lógico pensar, después de todo, que un hablante sabe mejor que nadie lo que quiere decir con sus palabras?

Dos cuestiones se plantean aquí. En primer lugar, se trata de ofrecer una caracterización general de la labor de la semántica, en especial, de aclarar si la labor de ésta consiste en la sistematización de las intuiciones lingüísticas de los hablantes. En segundo lugar, se suscita la cuestión de qué relevancia teórica otorgarle a semejantes juicios.

3.4.1. Intuiciones, sinonimia y analiticidad

La cuestión aquí, tal y como la plantea Devitt, es: «¿Si la tarea de la semántica no es sistematizar o explicar las intuiciones semánticas, ¿cuál es? ¿Qué está intentado hacer la semántica?» (M. Devitt 1996, p. 54). La diferencia entre la semántica y las ciencias naturales a este respecto consiste en que el objeto de éstas últimas «se halla ya identificado sin demasiada controversia» (*ibid.* p. 55), mientras que el de la semántica sí suscita polémica. En todo caso, la cuestión que nos ocupa en este apartado es por qué no debemos aceptar la idea de que la labor de la semántica es sistematizar las intuiciones lingüísticas de los hablantes. Devitt ofrece la siguiente respuesta:

La confianza en intuiciones semánticas podría también ser apropiada si estas intuiciones fueran elementos de un conocimiento a priori. Algunos las tratan como si lo fueran. Sin embargo, desde la perspectiva naturalista, debemos negar que haya *algún tipo* de conocimiento a priori. (*Ibid.*, p. 49).

El conocimiento a priori toma en semántica la forma de la analiticidad, por lo que no es casualidad que quienes defienden la relevancia para una teoría seántica de los juicios lingüísticos de los hablantes mencionen, en particular, las opiniones de éstos sobre la sinonimia y la analiticidad. Así, Dummett afirma: «Es un rasgo innegable de la noción de significado que el significado es transparente en el sentido de que si alguien le da un significado a dos palabras, debe saber si estos significados son el mismo.» (Dummett 1978, p. 131). Contra esta idea debemos recordar que el aspecto del significado de una palabra que pueda tener que ver con la relación que ésta guarda con otras palabras no es independiente de cuestiones fácticas. Aun cuando aprendamos el significado de una palabra aprendiendo a relacionarla con otras, la aplicación de éstas últimas dependerá de factores empíricos, a riesgo de hacer de la lengua un diccionario que nunca toca la realidad.⁴² Ésta es, básicamente, la razón esgrimida por Quine para rechazar la dicotomía analítico/sintético. Así, le damos a ‘mojado’ el significado que le damos *porque* le damos a ‘líquido’ el significado que le damos y porque los líquidos tienen una determinada propiedad (la de mojar). Así, los juicios de los hablantes acerca de la sinonimia no dejan de ser juicios empíricos, *a posteriori*.

⁴² Se utiliza aquí contra la analiticidad el argumento que utiliza F. Cowie (1987) contra los defensores del holismo, acusando a éstos de no subrayar los suficientes nexos extralingüísticos que constituyen, al menos en parte, el significado, expresándose a veces «como si el significado se constituyera solamente mediante las relaciones entre *tokens*» (en Devitt 1996, p. 14).

El rechazo de la analiticidad se fundamenta en la concepción holista del significado (v. 1.2).⁴³ Aceptada ésta, tanto la indeterminación del significado como la recusación de la dualidad analítico/sintético aparecen como consecuencias naturales. Así, una vez que se acepta que los significados de nuestras palabras dependen de nuestras creencias sobre el mundo (tesis de la inextricabilidad) y que éstas forman una compleja red (holismo), dejamos de hipostasiar significados. Y si éstos dejan de ser concebidos como entidades, dentro o fuera de la mente de los hablantes, se cierra el camino al conocimiento inmediato de éstos para la primera persona mediante la introspección o cualquier otro procedimiento cognoscitivo. Se puede afirmar, por tanto, esquemáticamente que *la semántica tendría como objeto las intuiciones de los hablantes si éstos tuvieran un acceso directo y privilegiado a los significados de las palabras, pero dado que esto no es así, la semántica debe tener otro ámbito de estudio.*

Devitt refina este argumento. El significado de una palabra parece hallarse constituido por dos tipos de relaciones: a) las relaciones que guarda con otras palabras y b) las relaciones, de tipo causal, que guarda con la realidad. Por ejemplo, aplicamos el término ‘soltero’ a las personas a las que se lo aplicamos en parte porque también les aplicaríamos la expresión ‘no casado’ y en parte porque nuestro sistema jurídico distingue entre personas solteras, casadas, separadas, etc.). Así, las expresiones ‘soltero’ y ‘no casado’ guardan una relación entre ellas que resulta relevante para la

⁴³ No podemos entrar aquí en un análisis en profundidad del ataque quineano de la analiticidad, nos centramos en el aspecto de éste que atañe directamente a nuestro tema; por lo demás, nos remitimos a lo expuesto en 1.2. El *locus classicus* de este ataque es Quine 1951, «quizá el artículo más famoso de la filosofía del siglo XX» (R. Creath 2004, p. 47).

determinación del significado de ambas; como lo resulta el hecho de que la persona designada como soltera nunca haya pasado por un tribunal o una vicaría para casarse. Ahora bien, ninguna de esas relaciones se constituye en la mente de un hablante. La cuestión es que ningún hablante puede percibir esas relaciones simplemente analizando su mente. Y si un hablante percibe en su mente dichas relaciones, nada le indicaría que esas relaciones son precisamente las constitutivas del significado. Pero la cuestión aún sería otra: aun suponiendo que pudiéramos percibir esas relaciones como las constitutivas del significado de las palabras, ¿por qué aceptar ese procedimiento introspectivo como el determinante del significado? Devitt expone así estas ideas:

El significado del *token* de alguien se halla constituido, presumiblemente, por propiedades relacionales de algún tipo: «internas», relaciones inferenciales entre *tokens* y «externas», ciertas relaciones causales directas con el mundo. Toma una de esas relaciones. ¿Por qué suponer que simplemente en virtud del hecho de que su *token* tiene esa relación, la reflexión debe conducirlo a *creer que* la tiene? Incluso si la reflexión lo hiciera, ¿por qué suponer que simplemente en virtud del hecho de que la relación constituye en parte el significado de su *token*, la reflexión debe llevarlo a *creer que* lo hace? Lo más importante de todo, aun cuando la reflexión lo llevara a estas creencias, ¿por qué suponer que simplemente en virtud de su competencia, este proceso de formación de creencia *justifica* las creencias, o les da alguna autoridad epistémica especial, y así las convierte en *conocimiento*? Suposiciones de este tipo parecen estar injustificadas. Necesitamos una explicación plausible de estos procesos supuestamente no-empíricos de formación y justificación de creencias y algunas razones para creer en ellos. (M. Devitt 1996. p. 53).

Cuando Quine señala que «hay una tozuda noción de que podemos decir intuitivamente qué idea expresa la oración de alguien, o la nuestra, incluso cuando la intuición es irreducible a criterios conductuales» (Quine 1969b, p. 304), parece estar conectando el recurso a las intuiciones semánticas con el mentalismo semántico. Y no sería injusto. Por lo que hemos visto hasta ahora, parece claro que, a menos que se apele a actos mentales irreducibles de determinación del significado y la referencia («el acto de determinación del agente», se lee en N. Georgalis 2004, p. 265), éstos se hallan afectados de indeterminación.

3.4.2. La labor de la semántica

La conclusión que debemos extraer de aquí es que la labor de la semántica no puede consistir en la sistematización de los juicios lingüísticos de los hablantes. Para ello sería preciso defender una concepción de la sinonimia y la analiticidad que no parece aceptable. ¿Cómo debemos, entonces, concebir la tarea semántica? El objetivo de una teoría semántica no consiste, como hemos visto, en sistematizar nuestras intuiciones lingüísticas, sino en ofrecer una interpretación de los hablantes. Davidson ha hecho hincapié en la idea de que lo que buscamos con la semántica es *racionalizar* la conducta de los demás – de ahí, por ejemplo, la utilización inevitable del principio de caridad por parte del intérprete radical (v. 1.7) – . En esta labor no entra en consideración el análisis de los juicios lingüísticos de los hablantes porque, entre otros factores, nociones como significado o referencia constituyen meros postulados con la única finalidad de explicar conductas, verbales y no verbales. Como dice Devitt: «Es

común definir la labor semántica en términos de *verdad* y *referencia*. Yo creo que la verdad y la referencia son centrales para conseguir la labor semántica pero no tienen lugar en su definición» (1988, p. 63). Ya hemos visto (1.7) que para Davidson la semántica constituiría sólo una parte de la labor interpretativa, ya que interpretar las palabras de alguien no puede ser independiente de la adscripción que realicemos de creencias o intenciones, la teoría interpretativa, pues, debe ser una Teoría Unificada de todos estos elementos:

La Teoría Unificada es holística de principio a fin. Está diseñada para asignar contenidos a creencias, emisiones y valores simultáneamente porque estas actitudes básicas son tan interdependientes que no sería posible determinar cada una por separado, o incluso de dos en dos. Su tratamiento de cada uno de estos dominios es también holístico: las oraciones se interpretan en términos de sus relaciones con otras oraciones, las creencias en términos de sus relaciones con otras creencias, etc. Un holismo así es característico de cualquier sistema de medida: los elementos deben su medida a sus relaciones con otros elementos. Se podría asignar un significado a una única oración suelta tan poco como se puede asignar peso o localización a un único elemento suelto. (D. Davidson 1995b, p. 130).

La concepción de Davidson hace patente la idea de que interpretar a un hablante no supone emparejar palabras con significados, sean éstos lo que sean. La interpretación de las palabras se convierte sólo en un elemento de una interpretación global de la conducta de otra persona, partiendo de su consideración como agente racional, y en esta concepción, la presencia o ausencia de algún tipo de entidad sólo accesible al

hablante, no parece ser un elemento relevante. Hemos adelantado ya que Wiggins hace consideraciones semejantes a este respecto, apelando al «sentido total» que debemos darle a un sujeto; Wiggins, en la misma línea que Davidson, afirma que «los esfuerzos lingüísticos de un intérprete forman parte de un esfuerzo mayor para interactuar con éxito con otros, para coordinar los esfuerzos de uno con los de ellos (donde resulte apropiado), para darles sentido, etc.» (Wiggins 1998, p. 18). Entender las palabras de alguien no supone asignarles un significado o una referencia, sino ver qué función tienen en la conducta, verbal y no verbal, de esa persona.

Devitt ha objetado contra la visión de la labor de la semántica que él considera la davidsoniana: «Los davidsonianos consideran que la labor de la semántica es decir cómo construir una teoría de la «interpretación radical» para un lenguaje.» (1996, p. 66). Pero, como hemos visto, esto no hace justicia a la visión davidsoniana – sea o no la de los davidsonianos – , donde la construcción de una teoría de la interpretación radical se concibe como una parte de la teoría global que pretende racionalizar la conducta de un sujeto, esto es, considerarlo un agente racional. En todo caso, la objeción de Devitt se dirige hacia la visión deflacionaria de los pensamientos en Davidson:

Desde mi perspectiva, esta tarea [construir una teoría de la «interpretación radical»] es valiosa pero no lo suficientemente fundamental. Se basa en nociones semánticas que necesitan explicación independiente. Los davidsonianos estarán en desacuerdo, ya que piensan que en estas nociones no hay más de lo que se revelaría cumpliendo su tarea. Esto refleja una visión de la mente muy diferente de la mía. Yo defiendo un mentalismo

robusto que considera los pensamientos estados objetivos postulados independientemente del lenguaje [...] (*Ibid.*)

Y se pasa a rechazar «la visión generalmente antirealista del significado» y el conductismo davidsonianos: «Los davidsonianos no han ofrecido ningún argumento a favor del conductismo que subyace a su perspectiva. Desde mi perspectiva, ese conductismo parece, simplemente, otro dogma del empirismo.» (*Ibid.*, p. 67). En ambas cuestiones remitimos aquí a lo ya dicho en las secciones anteriores (sobre el antirealismo de Davidson se vuelve en 3.7), con todo, añadamos algún comentario. El antirealismo de Davidson, entendido, al menos, como el rechazo de entidades tales como significados, no va ligado, como en Quine, al antimentalismo o al conductismo: los significados serían admitidos como entidades si hubieran demostrado utilidad explicativa. Repetimos aquí unas palabras de Davidson ya citadas en 1.5: «Mi objeción a los significados en la teoría del significado no es que son abstractos o que sus condiciones de identidad son oscuras, sino que no tienen una utilidad demostrada.» (Davidson 1967, p. 21). Por otro lado, el conductismo que pueda subyacer en la visión semántica de los davidsonianos se debe aquí, coincidiendo aquí con los quineanos, al reconocimiento del carácter público del lenguaje y, por consiguiente, a la consideración de que cualquier factor relevante para la determinación del significado debe de hallarse explícito en la conducta de los hablantes. Las palabras de Davidson que Devitt reproduce aquí de forma crítica: «El significado se halla enteramente determinado por la conducta observable, incluso, por la conducta abiertamente observable.» (Davidson 1990a, p. 314), no suponen más que una forma de enunciar

este hecho. Para Davidson, una vez aceptado este hecho, la indeterminación se convierte en una tesis inofensiva, meramente negativa: contra el platonismo semántico. Así pues, la analogía con los sistemas de medida puede ser vista como: (a) clarificando el concepto de «materia objetiva»:

No importa si se dice que Sam está a la izquierda de Susan o que Susan está a la derecha de Sam. Si se tienen los axiomas que definen algún sistema de medida, sea de peso, temperatura o probabilidad subjetiva, se pueden representar las estructuras así definidas en números de infinitas maneras. Lo que importa es lo que es *invariante*. Con el peso, un número positivo arbitrariamente escogido es asignado a algún objeto particular; relativamente a esa asignación quedan fijados los números que miden los pesos de todos los otros objetos. Se obtiene una forma igualmente buena de determinar los pesos multiplicando los originales por alguna constante positiva; son las *ratios* lo que son invariantes. Las invariantes son la «materia objetiva» («Invariances are «facts of the matter»»).

Así es como yo entiendo la indeterminación de la traducción y la interpretación. Dada la riqueza de todas las lenguas naturales, sería sorprendente si no fuera siempre posible describir los hechos de cualquier disciplina de diversas maneras. Tal indeterminación no amenaza la realidad de lo que se describe. Por supuesto, habrá confusión si no tomamos en consideración la relatividad de alguna manera de describir las cosas a la escala o el modo apropiado de descripción: importa si los números son Fahrenheit o centígrados, tus pesos libras o kilos, tus distancias millas o kilómetros. La analogía con el caso de oraciones o los contenidos de actitudes proposicionales es ésta: cada uno de nosotros puede pensar en sus propias oraciones (o sus contenidos) como si fueran los números; guardan múltiples relaciones entre sí y con el mundo. Manteniendo estas relaciones tanto como sea posible, podemos emparejar nuestras oraciones con las de un hablante, y con las actitudes de ese hablante, de formas diferentes sin cambiar nuestra opinión

acerca de los hechos. Así como infinitos conjuntos de números nos permiten representar [*to keep track*] las mismas complejas estructuras del mundo, así nuestras oraciones pueden ser usadas de infinitas formas diversas para representar las actitudes de otros, y lo que quieren decir. Quine señaló esto para subrayar que no hay más en la identificación de los significados de lo que hay al captar estas complejas relaciones empíricas. (Davidson 1999b, p. 596).

(b) mostrando que la tesis resulta inocua respecto al estatus ontológico de significados y actitudes proposicionales (no cambiamos nuestros compromisos ontológicos por medir en metros o en yardas):

Lo que estoy exponiendo es mi propia visión de la indeterminación y la infradeterminación. Está claramente en contradicción con la manera de Roger Gibson de explicar la visión de Quine de la distinción, que sí la hace parecer teniendo implicaciones ontológicas, y le da un giro peculiar a la expresión «materia objetiva». No sé Quine, pero mi forma de entender la distinción no tiene significancia ontológica en absoluto, y, por tanto, no puede implicar que hay algo ligeramente irreal, inferior o acerca de los fenómenos mentales. (*Ibid.*, p. 597).

(c) clarificando la noción de relativizar a una lengua. Nuestra familiaridad con esquemas de traducción como «‘Rabbit’ = ‘Conejo’» o el manual homofónico para la propia lengua, puede hacernos olvidar la necesidad de hacer explícito el esquema utilizado, uno de los siempre varios posibles:

Si utilizáramos realmente más de las diversas formas posibles de decir lo que alguien quiere decir y piensa, querríamos ser explícitos sobre el sistema que estamos usando. En la práctica, nos aferramos enteramente a una forma; es como si todo el mundo hablara en centígrados o millas o kilos. Pero para objetivos teóricos, está bien saber que nuestras asignaciones son relativas: son relativas a un lenguaje. (*Ibid.*).

(d) ejemplificando la idea de que la tesis no tiene implicaciones respecto a la ontología de los significados, excepto su rechazo de éstos como entidades platónicas:

Esto puede sonar como una tesis negativa, y lo es; es un ataque a la idea de que los significados pueden ser capturados de una manera exclusivamente, adhiriendo significados platónicos a las expresiones. Pero este punto negativo no conlleva que no haya materia objetiva; la materia es las relaciones empíricas entre un hablante, sus oraciones y su medio. Este patrón es invariante. (*Ibid.*, p. 596).

3.4.3. Intuiciones y competencia

Un defensor de la relevancia de las intuiciones semánticas de los hablantes como materia para una teoría semántica empírica puede continuar insistiendo en el aspecto central que deben tener las opiniones de los hablantes sobre lo que quieren decir, al contrario que en otras materias. Esto es, la «idea familiar de que los hechos semánticos dependen de nosotros de una forma en la que los hechos de la química o la biología no». Se trae aquí a colación una tesis bautizada como «verificacionismo semántico» (v. 3.2.):

Es la idea de que los hechos semánticos de cualquier lengua particular deben ser en principio accesibles a cualquier agente lingüístico bien situado; la realidad semántica no podría hallarse para siempre más allá del alcance epistémico de los hablantes. (D. Bar-On 1992, p. 235).

Siguiendo aquí reflexiones de Devitt, debemos comenzar subrayando que las teorías semánticas actuales se hallan enfrentadas de una manera tan inusual que ni tan siquiera su campo de estudio se halla claro. ¿Cuál debe ser la materia de estudio de la semántica: las disposiciones verbales de Quine, los sentidos fregeanos, los contenidos Intencionales de Searle? En todo caso, parece que todas estas teorías constituyen diferentes concepciones del significado, pero todas ellas, como teorías semánticas, estudian los significados, esto es, las emisiones de los hablantes como emisiones significativas. Y aquí es donde, dice Devitt, comienza la labor de la semántica: en localizar emisiones significativas. Y una vez hallados ejemplares inequívocos, podría comenzar su análisis. La labor de la semántica, pues, se divide en dos etapas:

1. Hallar ejemplares incuestionados de emisiones con significado.
2. Analizar los ejemplares para determinar qué tienen en común y ofrecer, así, una explicación de en qué consiste ser una emisión significativa – esto es, qué es el significado – .

Pues bien, la idea ahora es que las intuiciones lingüísticas de los hablantes son de utilidad en la primera fase y relativamente. Ya hemos visto (3.4.1) cómo Devitt

rechazaba la utilidad de las intuiciones lingüísticas sobre cuestiones como la sinonimia argumentado que no se trata de cuestiones comunes de los hablantes, ahora bien:

En contraste, las identificaciones de *tokens* con un cierto significado putativo no tienen por qué contener tales términos y son centrales en la práctica popular de ascribir pensamientos y emisiones. Así que debemos suponer que la gente es menos experta en identificar la sinonimia y similares que lo son en identificar significados putativos. (Devitt 1996, p. 77. Nota a pie de página.)

Es aquí, en este primer estadio de la semántica, donde hallamos un lugar para las intuiciones lingüísticas, ya que los hablantes competentes, por el hecho de serlo, distinguen las emisiones significativas de las que no lo son. La relevancia de estas intuiciones, no obstante, resulta muy limitada, ya que si pensamos que los teóricos semánticos son también hablantes competentes, no necesitan recurrir a los hablantes «de a pie». Así, si la semántica debe considerar las emisiones significativas como su materia, su primera fase consistirá en identificarlas, para lo que podría resultar útil consultar a los hablantes no expertos, sin embargo, «en esta primera fase, los teóricos semánticos tienen dos ventajas: no necesitan consultar a otros y tienen un acceso inmediato a los datos.» (*Ibid.*, p. 78).

Devitt concibe estas intuiciones básicas con las que la semántica comienza su tarea como «intuiciones de identificación», esto es, en la que los hablantes pueden indicar a qué objetos se refieren con sus palabras. Así, pasa a ilustrar el caso de la semántica con el de una teoría cuya materia de estudio fueran los equidnas. Primero, se trataría de

hallar ejemplares de echidna para luego analizarlos. Y una manera de hallar los ejemplares sería preguntar a quien sabe distinguir los echidnas de otros animales. Ahora bien, alguien podría saber decir si algo es un echidna o no y no tener demasiado que decir sobre ello. El caso podría ser aún peor, ya que una vez estudiados los equidnas, podría resultar que el hablante no distinguiera perfectamente a los equidnas de otros animales o que, sometido a experimentos mentales – del tipo de Tierra Gemela – , no pudiera distinguirlos correctamente en determinadas circunstancias. Por tanto

incluso las intuiciones más básicas, expresadas en experimentos de identificación, se hallan sujetas a revisión a la luz del examen científico. *Las intuiciones suelen ser necesarias para identificar la materia de estudio para la tarea descriptiva, y podrían resultar útiles, pero nada depende de ellas en última instancia. (Ibíd., p. 74).*

Insistir en que el dominio del lenguaje debe de implicar conocimiento acerca de elementos como la referencia o la sinonimia significa confundir la pericia en una actividad con el conocimiento de una teoría descriptiva de esa actividad. Se ha dicho ya algo acerca de la teoría «de la teoría» del lenguaje (v. 3.2); se trata aquí de recalcar que el dominio del lenguaje es una habilidad, no un conocimiento de una teoría, por lo que las opiniones lingüísticas de los hablantes no resultan de relevancia para la confección de una teoría lingüística. Ha sido Devitt quien más ha insistido en el hecho de que el dominio del lenguaje constituye más el dominio de una técnica, un *know-how*, que un conocimiento teórico, un *know-that*. De hecho, personas muy hábiles en una actividad podrían reflexionar muy poco sobre ella, con lo que sus opiniones al respecto no

resultarían de gran trascendencia; Devitt introduce aquí la comparación con los deportistas:

Una persona puede ser competente y, sin embargo, reflexionar poco sobre el producto [*output*] de esa competencia; piensa en un ciclista, por ejemplo. Es un tópico en el deporte que los grandes jugadores no siempre son grandes entrenadores. El hecho de que posean una competencia en un grado superlativo no implica que puedan articular y comunicar los elementos de esa competencia. El saber-cómo no es un saber-qué [*Knowledge-how is not knowledge-that*]. Incluso si una persona competente se convierte en un experto, no hay razón para creer que sus opiniones poseen autoridad especial simplemente porque es competente; no hay razón para creer que la competencia le otorga acceso cartesiano a *la verdad*. (Devitt y Sterelny 1999, p. 182).

Resulta aquí sencillo confundir el acceso inmediato a una gran cantidad de datos con las conclusiones que se extraen de estos datos. El hablante, por el hecho de serlo, tiene acceso inmediato a los datos lingüísticos – qué palabras consideraría adecuadas en cada momento, p. ej. – , pero esto no implica que sea capaz de inferir conclusiones correctas a partir de éstos. Estas conclusiones constituyen, más bien, una teoría del significado confeccionada por expertos en el lenguaje, no por meros hablantes. Devitt introduce aquí otra comparación esclarecedora:

La mecanografía ofrece un bonito ejemplo de la reflexión sobre el producto [*output*]. Pregúntale a un mecanógrafo qué dedo debería utilizarse para escribir, por ejemplo, una ‘k’, y muy probablemente no tendrá una respuesta inmediata: es un experto en mecanografía pero no un experto en la «teoría» de la mecanografía. De hecho, se

preguntará: «¿cómo escribiría yo una ‘k’?» (...) El único privilegio de que disfruta esta opinión es el privilegio de estar basada en lo que es, con seguridad, un buen dato: en cómo él, un buen mecanógrafo, escribe una ‘k’. (*Ibid.*, pp. 182-183).

Con todo, el verificacionismo semántico parece una idea sumamente coherente.

Pensemos, si no, en lo siguiente:

Una comunidad de hablantes no puede malinterpretar sistemáticamente expresiones usadas habitualmente – no pueden estar siempre equivocados acerca de lo que significan tales expresiones o acerca de si dos expresiones significan lo mismo. Si los hablantes de la lengua consideran que una expresión significa algo (o que dos expresiones tienen el mismo/diferente significado), entonces no podría haber lugar para el descubrimiento de que todos ellos se hallaban equivocados. (D. Bar-On 1997, p. 102)

Sin embargo, en cuestiones como la sinonimia nos remitimos nuevamente a lo defendido por Devitt: se trata de una cuestión demasiado técnica como para que un hablante no lingüista haya reflexionado sobre ella. Ni, por otro lado, le resulta necesario para poseer un dominio perfecto de la lengua. No podemos esperar que los hablantes no expertos resulten muy útiles a la hora de contestar cuestiones sobre la sinonimia de expresiones como ‘conejo’ y ‘parte de conejo no separada’.

Podemos esquematizar las conclusiones del presente debate:

- a) Se ha apelado (o se podría apelar) a las intuiciones lingüísticas de los hablantes respecto a los significados de sus palabras como un argumento contra la indeterminación (los hablantes, por ejemplo, suelen afirmar que con ‘conejo’ se refieren unívocamente a conejos).
- b) La apelación a dichas intuiciones asume, obviamente, que dichas intuiciones deben ser tomadas en consideración por una teoría lingüística.
- c) Dichas intuiciones deben ser tomadas en consideración porque parece lógico pensar que un hablante competente conoce el significado de sus palabras – o sabe decir cuándo dos expresiones son sinónimas, a qué se refieren sus términos, etc. – .
- d) El error aquí consiste en la concepción de la competencia lingüística como el conocimiento de una teoría – una gramática, por ejemplo – , en vez de como el dominio de una habilidad.
- e) Dominar una habilidad no implica ser un buen teórico de la misma. Las opiniones lingüísticas de los hablantes, pues, no resultan de gran utilidad, puesto que no son expertos en lingüística.
- f) Las intuiciones de los hablantes sólo resultarían de utilidad si nos comprometemos con la tesis internista de que los significados de las palabras se

hallan en nuestra mente y con la idea cartesiana del acceso privilegiado a los contenidos mentales por parte de la primera persona; así, ser un hablante competente implicaría tener un conocimiento inmediato e infalible de éstos. Contra esta idea ya se ha argumentado en las secciones precedentes.

- g) Se debe concluir, pues, que las intuiciones en cuestión de semántica se hallan en las mismas condiciones que en cualquier otra cuestión sujeta a estudio científico:

¿Qué decir sobre las intuiciones lingüísticas? Las cuestiones sobre el estatus de las intuiciones no sólo surgen en lingüística, por supuesto; las intuiciones tienen un papel en la ciencia y parecen dominar la filosofía. Considérese, por ejemplo, un paleontólogo en el campo, buscando fósiles. Ve un trozo de piedra blanca asomando a través de roca gris, y piensa «es la mandíbula de un cerdo». Este juicio intuitivo es rápido, irreflexivo. Puede estar bastante seguro, pero es incapaz de explicar cómo lo sabe. Las intuiciones del paleontólogo, como las intuiciones en general, *están cargadas de teoría, y son casi siempre respuestas incompletas a hechos del mundo* [wordly]. Él puede estar seguro, pero su juicio se halla aún abierto al examen y la revisión. (Devitt y Sterelny 1996, p. 181).

3.5. Davidson: interpretación y primera persona

Hemos analizado hasta ahora los argumentos de quienes han rechazado la tesis de la indeterminación apelando a la autoridad de la primera persona. Dicha autoridad, con todo, constituye un hecho, *prima facie*, tan palmario que parece difícil no reconocerlo. Analizamos ahora la explicación davidsoniana de la autoridad de la primera persona en

lo referente al significado de las propias palabras, explicación que conjuga la defensa de dicha autoridad con la indeterminación del significado. Partiendo, como hemos visto, de una teoría semántica extensionalista y de una postura decididamente externista, la defensa davidsoniana de la autoridad de la primera persona no se basará en el acceso privilegiado a los contenidos mentales – los contenidos Intencionales de Searle, por ejemplo – . La explicación davidsoniana resulta, así, de sumo interés por aunar elementos a primera vista incompatibles: el reconocimiento de la autoridad de la primera persona, el externismo y la indeterminación.

3.5.1. Asimetría epistémica y simetría semántica

Como se ha visto en la exposición del proceso de triangulación (3.3.1.), Davidson defiende que lo que las palabras de un hablante significan depende de los tipos de objetos y situaciones en las que la persona ha aprendido que esas palabras se aplican. Así, el sujeto nunca se pregunta (o casi nunca) si sus palabras significan lo que él cree, ya que simplemente está aplicándolas a los contextos donde las considera aplicables. Lo que implica, también, que nada puede contar como que un hablante está utilizando sus palabras sistemáticamente de forma incorrecta (premisa de la imposibilidad del error masivo). El hablante les puede estar asignando significados «no estándar», esto es, no usuales entre el resto de la comunidad lingüística, pero en todo caso, él aplica las palabras a las situaciones a las que las considera aplicables.

Así, Davidson defiende la autoridad de la primera persona admitiendo una *asimetría interpretativa* o *epistémica* entre la primera y la tercera persona; no obstante, dado que

lo que determina el significado de las palabras es la conducta públicamente accesible respecto a los significados la primera y la tercera personas tiene acceso a lo mismo; hay, pues, *simetría semántica*.

La asimetría interpretativa es la forma en la que Davidson admite la autoridad de la primera persona: ésta conoce el significado de sus palabras de manera diferente a cómo lo hace otra persona. Como ya hemos señalado, esta explicación no se basa en el acceso privilegiado a la mente, sino en el hecho de que yo no necesito interpretarme a mí mismo para saber lo que quiero decir, ya que al utilizar el lenguaje simplemente aplico mis palabras a los objetos y las situaciones a los que aprendí a aplicarlas:

Tu conocimiento de lo que significan mis palabras tiene que estar basado en evidencia e inferencia: tú, probablemente, supones que lo has hecho correctamente y probablemente lo has hecho. Sin embargo, es una hipótesis. (Davidson 1989a, p. 66).

Así pues, Davidson admite una de las ideas tradicionales en la defensa de la perspectiva de la primera persona: que ésta conoce el significado de sus palabras de manera directa, mientras que la tercera persona lo hace de manera inferencial – inferencias realizadas a partir de la conducta observable – . Pero lo interesante ahora es que esta idea se asienta en una concepción netamente anticartesiana de la mente; el acceso no inferencial al significado de las palabras por parte de la primera persona no consiste en un acceso a su mente, donde se hallan los significados, sino en un conocimiento de la historia causal (proceso de triangulación) mediante la cual aprendió a usar las palabras. La asimetría consiste en la diferente posición de alguien que se ve obligado a realizar hipótesis acerca

de a qué situaciones aplico qué palabras, y yo, que simplemente aplico esas palabras en las situaciones en las que lo considero adecuado. Davidson parece acercarse aquí a la concepción de Ryle de la autoridad de la primera persona: yo tengo más información sobre mis estados mentales porque dispongo de más información sobre mí que cualquier otra persona (v. 2. 4.)⁴⁴, pero no porque tenga acceso a nada diferente. De ahí que algunos autores hayan visto la explicación davidsoniana de la autoridad de la primera persona como una explicación desde la perspectiva de la tercera («la diferencia entre la visión de la primera y la tercera persona se hace desde el punto de vista de la tercera», K. Puhl 1994, p. 342), de hecho, el propio Davidson admite que prestar atención «a cómo atribuimos pensamientos y significados a otros explicaría la autoridad de la primera persona» (1987, p. 17).

Tal y como lo expone B. C. Smith (1998), la asimetría lingüística se basa en la diferente respuesta que damos a las siguientes preguntas:

- (a) ¿Cómo sabemos lo que quieren decir los demás hablantes con sus palabras?

- (b) ¿Cómo sabemos lo que queremos decir nosotros con nuestras palabras?

⁴⁴ Nótese que Davidson rechaza la idea de Ryle de que uno se observa a sí mismo para conocer sus estados mentales, esto es, que la asimetría entre la primera y la tercera persona consistiría en que uno puede observarse a sí mismo más (más tiempo) y mejor que los demás, por lo que dispone de más información. Nuestra comparación entre ambos autores pretende subrayar simplemente que tampoco para Davidson hay una diferencia cualitativa entre la información de la que dispone la primera y la tercera persona, se trata sólo de que la primera persona dispone ya de toda la información.

La respuesta a (a) es que lo sabemos interpretándolos. Y la interpretación, como vimos, aun aplicando de forma exhaustiva el principio de caridad y apelando al estímulo distal en vez de al estímulo próximo, no puede eliminar un grado considerable de indeterminación. La respuesta a (b), y de ahí proviene la asimetría entre la primera y la tercera persona, es que lo sabemos sin necesidad de llevar a cabo un proceso de interpretación. El interrogante que suscita (b), por tanto, sería:

(c) ¿Cómo sabemos lo que queremos decir si no nos interpretamos a nosotros mismos?

Ya hemos visto cuál es la respuesta de Davidson a (c): la apelación al proceso de aprendizaje – de triangulación – . Un hablante sabe lo que quiere decir sin necesidad de interpretarse porque lo único que hace es aplicar sus palabras a los objetos y situaciones a los que aprendió a aplicarlas. Así visto, ni tan siquiera podría darse el caso de que un hablante no sepa lo que quiere decir con sus propias palabras. Podemos clarificar la distinción entre la asimetría interpretativa y la simetría semántica diciendo que la primera se basa en que la primera persona no precisa emprender un proceso de interpretación para determinar el significado de sus palabras, y la segunda se basa en el hecho de que lo que determina el significado de las palabras son objetos y eventos públicamente accesibles (estímulo distal). Esto se aplica a los significados tanto como a los pensamientos:

[D]onde el intérprete debe conocer, o realizar conjeturas correctas, los eventos y situaciones que causan una reacción verbal o de otro tipo en otra persona para adivinar sus pensamientos, el que piensa no precisa tal conocimiento nómico para decidir lo que está pensando. (Davidson 1989b, pp. 195-196).

La visión davidsoniana aquí se basa, pues, en cuatro puntos (se sigue un tanto libremente G. Soldati 2002, p. 23):

1. El significado de las palabras se determina, en parte, por su causa. (Premisa externista).
2. La causa de las palabras es públicamente accesible. (Premisa del estímulo distal).
3. El hablante sabe el significado de sus palabras sin realizar hipótesis interpretativas acerca de sí mismo. (Premisa de la autoridad de la primera persona).
4. El intérprete no puede determinar el significado de las palabras del hablante sin formular hipótesis acerca de lo que las causó. (Premisa de la asimetría interpretativa).

Con todo, los críticos de Davidson no han hallado convincente la explicación de la autoridad de la primera persona en base a una asimetría meramente interpretativa. El

problema vendría a ser que, si se admite la indeterminación, un hablante podría ser interpretado de una manera que él mismo no aceptaría, con lo cual indeterminación y autoridad mantienen la tensión habitual. Si se conjuga ambas, como hace Davidson, debe ser al precio de ofrecer una visión de la autoridad de la primera persona tan sumamente mitigada que se parecería poco a la autoridad que buscábamos explicar. Dentro de la semántica davidsoniana, se puede pensar, no parece haber lugar para que un hablante quiera decir algo determinado:

La cuestión clave es si los significados que un teórico asigna a las palabras y oraciones de un hablante son los significados que el hablante asocia con esas palabras y oraciones. ¿Hay una explicación más directa del conocimiento de primera persona del lenguaje que tiene el hablante que pueda explicar cómo los significados se aprehenden conscientemente por los hablantes y se reconocen en las palabras de otros? Con seguridad, deseamos que los significados que la teoría asigna al lenguaje sean los que los hablantes y los oyentes del lenguaje considerarían que son. (B. C. Smith 2006, p. 388).

Hagamos aquí una diferenciación, siguiendo a M. Dummett (2002), entre una indeterminación fuerte y una débil. Haciendo una lectura *débil* de la indeterminación, podríamos permitir diferencias en el significado de las palabras que componen las oraciones de un hablante, pero las condiciones de verdad deben permanecer inalteradas. Interpretar consistiría aquí en determinar las condiciones de verdad de una oración. Dicho de otro modo, toda interpretación que ofrezca estas condiciones de verdad es correcta. Aquí, la caridad – o el mecanismo que se haya escogido para comenzar la

interpretación – determina la *metodología* de la interpretación, no su corrección. La primera persona mantiene aquí su autoridad por su conocimiento de las condiciones de verdad.

Por otro lado, haciendo una lectura *fuerte*, la indeterminación podría implicar también variaciones en las condiciones de verdad de las oraciones. Aquí, la caridad no sólo constituiría una metodología de la interpretación, sino que definiría las condiciones de corrección. Dummett ve en esta segunda lectura una incompatibilidad insalvable entre la indeterminación y la autoridad de la primera persona, ya que si realmente hay asimetría entre la primera y la tercera persona se debe a que el conocimiento de la primera persona de los significados de sus palabras no depende de las reglas que el intérprete haya escogido para interpretar.

Dummett piensa que Davidson «quizá nunca pretendió defender más que la indeterminación débil», y parece sugerirse la idea de que sólo ésta resulta compatible con la autoridad de la primera persona. La razón sería que, admitiendo una indeterminación débil la autoridad queda a salvo bajo la forma de conocimiento de las condiciones de verdad, pero la indeterminación fuerte permitirá no sólo la variación de significados, sino también la de las condiciones de verdad, ¿de qué, entonces, podría tener la primera persona un conocimiento privilegiado? Esto es, si el significado de las palabras de uno se determina por su interpretación correcta y uno no es su propio intérprete – no determina los criterios correctos de interpretación – , «entonces, uno no

puede saber, sólo a sus propias luces, lo que quiere decir con lo que dice.» (G. Soldati 2002, p. 21).⁴⁵

Efectivamente, Dummett tiene razones para dudar de que Davidson se haya comprometido alguna vez con una indeterminación fuerte, ya que el proceso de triangulación, con su énfasis en el estímulo distal y la empatía entre maestro y aprendiz, no parece dejar mucho lugar para la indeterminación de las condiciones de verdad. Quine, sin embargo, sí ha hablado a favor de una indeterminación fuerte en el sentido de Dummett. Más fuerte aún, ya que ha defendido también la indeterminación de los valores de verdad: dos traducciones «podrían ser, incluso, manifiestamente contrarias en valor de verdad, a condición de que no hubiera estimulación que produjera el asentimiento a ninguna de las dos.» (Quine 1960, p. 73-74). Esto es, un manual de traducción que tradujera una oración *o* como *p*, y otro que tradujera *o* como $\neg p$, resultarían ambos empíricamente correctos si el sujeto nunca se manifiesta sobre el valor de verdad de *o*. Davidson no parece objetar a esto, ya que al explicar que la

⁴⁵ Hay quien piensa que la indeterminación débil ya es muy fuerte, ya que no resulta tolerable hacer depender los significados de las palabras de *mi* elección de manual. Así, B. Stroud ha ideado un experimento mental en el que se le provee de varios manuales de traducción urdu-turco, para que utilice uno u otro según el día, pero no entiende ni urdu ni turco:

¿Sobre lo que ellos [hablantes nativos de turco] están hablando puede «ser relativo» a mi elección de manual de traducción? Obviamente no. Qué manual escoja yo ni tan siquiera determina a qué objetos *creo* que se están refiriendo. No tengo opinión sobre a qué se están refiriendo. Carezco de toda comprensión de los términos que estoy traduciendo y de los términos a los que estoy traduciendo. No son términos míos. Eso es por lo que mi habilidad para traducir no equivale a entender. (B. Stroud 1995, p. 46).

Esta crítica se basa en lo que ya hemos criticado a Searle: la presuposición de que porque un término tenga referencia, hay *algo* que es su referencia: «sé que si la palabra turca se refiere a conejos, no se refiere también a estadios de conejo», dado que «conejos y estadios de conejo no son el mismo tipo de cosa» (*ibid.*). Dummett, sin embargo, no considera que esta relatividad al manual de traducción ponga en peligro la autoridad de la primera persona mientras ésta pueda determinar, al menos, las condiciones de verdad de sus oraciones.

indeterminación puede surgir debido a las diferentes maneras de diferentes manuales de traducción de repartir significados y creencias, afirma que esto implica «que en una teoría una oración dada puede ser verdadera para un tiempo y un hablante, y en otra teoría, no.» (Davidson 1995a, p. 16). Y explica que esto «sólo significa que, dentro de límites, podemos asignar diferentes condiciones de verdad a una oración tal y como es usada por un hablante o hablantes dados mientras hagamos los cambios apropiados en otros lugares de la teoría.» (*Ibid.*, p. 17).

Con todo, incluso la indeterminación en sentido fuerte resulta compatible con la explicación de Davidson de la autoridad de la primera persona. Esto es debido a que las críticas a esta explicación, sea defendiendo una indeterminación débil o fuerte, se basan en la petición para la comunicación de *algo más* que la confluencia de hablante e intérprete respecto a los objetos y eventos referidos *en un momento dado*. Ese «algo más» sería en el caso de la indeterminación débil el conocimiento de significados o referencias determinadas, y en el de la fuerte, de condiciones de verdad. Pero ni una cosa ni la otra son necesarias para la interpretación.

Hemos de añadir aquí otra idea fundamental en la concepción davidsoniana de la comunicación: mantener una conducta verbal – en contraste a emitir sonidos sin significado alguno – consiste en aplicar las palabras siguiendo un patrón coherente. La coherencia consiste aquí en un patrón de uso discernible – en contraste a emitir sonidos azarosamente – . La coherencia así entendida se convierte en Davidson, pues, en condición de la interpretabilidad, y ésta constituye el criterio de lingüisticidad.

(Volvemos sobre esta cuestión en la siguiente sección.) Un conjunto de emisiones (o inscripciones) constituye conducta verbal si es interpretable, y será interpretable si existe un patrón coherente de aplicación.

Un hablante sabe lo que quiere decir porque en cada momento dado emplea las palabras con las que mejor se hará entender – que hay un conejo, que se cierre una ventana o preguntar la hora – . La indeterminación aparece como algo inocuo aquí: el intérprete puede expresar lo mismo utilizando otras oraciones.

Quienes demandan «algo más» a esta explicación de la autoridad de la primera persona localizan el problema en que Davidson «no ve diferencia entre alguien queriendo decir algo y ese alguien sabiendo lo que quiere decir» [*he sees no difference between someone meaning something and his knowing what he means*] (B. C. Smith 1998, p. 417). La idea de que los hablantes saben lo que quieren decir por el hecho de aplicar de forma coherente sus palabras parece que da la vuelta a lo que normalmente sucede: si aplicamos de forma coherente nuestras palabras es porque hemos aprendido lo que significan. Nuestras intuiciones nos dicen que un hablante utiliza coherentemente sus palabras porque sabe lo que quieren decir, no al revés, como Davidson afirma. Citando a B. C. Smith:

Los hablantes, se nos dirá, quieren decir lo que quieren decir con sus palabras porque esas palabras son regularmente aplicadas a las mismas cosas. Pero al aplicarlas no actúan ciegamente. No nos limitamos a usar palabras: las entendemos. Los hablantes saben lo que quieren decir con las oraciones que consideran verdaderas. Esto se fija por

las condiciones de verdad que los hablantes consideran que tienen, y es este conocimiento el que les es accesible sin interpretación. Así que ¿en qué consiste el conocimiento de un hablante de las condiciones de verdad? De nuevo, parece haber un nivel de conocimiento ausente. (B. C. Smith 1998, p. 417).

Como vemos, parece que se argumenta aquí tanto contra la indeterminación débil como contra la fuerte, y en ambos casos sobre la premisa de que los hablantes conocen el significado de sus palabras. (a) Por un lado, el problema de la indeterminación débil consistiría en que aprender unas palabras es algo más que aprender las condiciones de verdad o patrones objetivos de uso, es aprender el significado de las palabras («no nos limitamos a usar palabras: las entendemos»), y es este significado el que determina las condiciones de uso (este sería el factor ausente en la teoría davidsoniana), no viceversa; (b) por otro lado, el problema de la indeterminación fuerte vendría a ser que los hablantes conocen las condiciones de verdad de las emisiones (éste sería el conocimiento, se nos dice, «accesible sin interpretación»).

Bilgrami, por ejemplo, abunda en la primera cuestión, y afirma que «el desentrecomillado, si ha de estar al servicio de una explicación del significado no es una idea completamente trivial. Ha de estar anclado en algo que no se halla explícito en la misma oración desentrecomilladora» (citado en Barry C. Smith *op. cit.*, p. 420). Se debe recordar aquí, de nuevo, que la oración a la derecha de la convención V se limita a ofrecer una interpretación – lo hace ofreciendo las condiciones de verdad – de la oración mencionada a la izquierda. Si la lengua de ambas oraciones – la mencionada y la usada – coincide, entonces la mejor manera de ofrecer dicha interpretación es,

meramente, el desentrecomillado. Si lo que se pide es que la convención V ofrezca el significado en un sentido «más fuerte» que, meramente, describir las condiciones de uso, entonces hemos de insistir, una vez más, en el carácter público del lenguaje y, en especial, en la simetría semántica que éste conlleva. Así, Davidson afirma que «lo que un intérprete perfectamente informado puede llegar a saber sobre lo que un hablante quiere decir es todo lo que hay que saber.» (1983, p. 129).

Conviene percatarse de que este tipo de críticas se dirigen, en última instancia, al extensionalismo de la teoría semántica davidsoniana. Se trata de críticas que comparten el descontento por la equiparación en la convención V de significados y condiciones de verdad (v. 1. 4.) El ejemplo paradigmático de esta crítica lo constituye Foster 1976, donde se trata de mostrar que una teoría tarskiana de la interpretación no puede resultar interpretativa debido a su rígido extensionalismo. Allí se lee:

Y a partir de este caso [el ejemplo de indeterminación expuesto aquí en 1.5] ahora podemos extraer la conclusión general que demole la tesis de Davidson, a saber, que las teorías V no consiguen ser genuinamente interpretativas para todos los lenguajes donde la completa significación de los elementos, junto con sus contribuciones a lo que se expresa en las oraciones que los contienen, supera su significación en términos de condiciones de verdad [*exceeds their truth-relevant significance*], junto con lo que contribuyen a la verdad o falsedad de estas oraciones. (Foster 1976, p. 15).

Ante esta crítica no podemos más que mantenernos firmes en la idea de que la significación no puede superar lo que queda manifiesto en la conducta verbal de los

hablantes. Y conviene hacer hincapié aquí en el carácter holístico de una teoría del significado de tipo davidsoniano: el significado de una oración *o* viene dado no sólo por el teorema que se refiere directamente a ella, sino por el conjunto infinito de teoremas que se refieren a todas las oraciones que contienen algún componente de *o*. Así, cada esquema *V* particular dice menos de lo que parece, pero tomados en su conjunto todos los esquemas *V* que una teoría de la verdad de una lengua puede suministrar, dicen todo lo que hay que decir sobre el significado. Por decirlo con palabras de Quine:

Se tome por donde se tome, el hecho es que al aprender el lenguaje aprendemos a distribuir valores veritativos – en esto coincido con Davidson – o sea, aprendemos condiciones veritativas. Al aprender una oración ocasional aprendemos en qué circunstancias hay que considerarla verdadera y en qué circunstancias hay que considerarla falsa. Al aprender la construcción predicativa eterna aprendemos a juzgar si un par dado de términos produce una predicación verdadera – definitivamente verdadera – o una predicación falsa – definitivamente falsa –. (Quine 1974, p. 83.).

Consideremos ahora la indeterminación fuerte. La crítica se basa en la idea de que conocer el significado de una palabra implica saber sus condiciones de verdad, y lo mismo regiría para la interpretación. Si hemos dicho que la crítica anterior surgía por el descontento que suscita entre algunos autores el que la convención *V* ofrezca el significado de una oración ofreciendo sus condiciones de verdad, ahora el descontento surge porque se ofrezcan las condiciones *en un momento dado*, sin apelación a reglas, algoritmos o convenciones. Pero Davidson niega que esto sea necesario. Darle un

significado a una palabra sólo implica aplicarla de forma coherente; o mejor: aplicarla de forma que algún otro hablante pueda hallar coherente – el significado es siempre cosa de dos – . La comunicación se hace posible cuando el intérprete es capaz de hallar ese patrón en las palabras del hablante, e infiere, por tanto, a qué objetos y situaciones se está refiriendo (v. 3.3.1.) En ese sentido, podemos afirmar que los significados se generan en cada conversación *in vivo*, y ésta es la idea de lo que Davidson llama la teoría de paso (v. 1.6). Así, un hablante sabe lo que quiere decir con sus palabras – las aplica en las circunstancias en las que las considera aplicables – y querrá decir algo con éstas siempre y cuando se utilicen de manera coherente – resulten *interpretables* – , y exigir el conocimiento de condiciones de verdad unívocas por el hablante o el intérprete no supone más que una nueva hipóstasis de significados. Hablar de condiciones de verdad, en contraste con la aplicación concreta de las palabras, nos llevaría de nuevo a una discusión sobre las intuiciones de los hablantes acerca de qué dirían en determinados casos, o bien nos conduciría a la descripción de la conducta lingüística como sometida a reglas. Hemos discutido ya tanto la validez de dichas intuiciones (3.4.1) como la conveniencia de apelar a reglas y convenciones (3.3.3 y 3.4.2), con la conclusión de que no abren un camino promisorio en el rechazo de la indeterminación debido a la autoridad de la primera persona.

Como hemos señalado ya, interpretar a un sujeto – y, por tanto, aprender su lenguaje – sólo requiere la *habilidad* de captar el patrón de aplicación de palabras que utiliza. Paa ello, sólo resulta imprescindible que ambos sujetos, interpretado e intérprete, posean unos estándares semejantes con los que juzgar la similitud de los objetos o

circunstancias ante los que se realice la emisión; así debe suceder, al menos, en los casos más básicos – en las llamadas «oraciones observacionales» – . Quine lo explica así:

Adquirir una oración observacional es aprender cuándo se debe esperar de un hablante veterano que apruebe la emisión de ella que uno haga, o que asienta ante ella por sí solo. Esto se puede aprender de instancias simples por inducción: extrapolando a los demás casos sobre las líneas de similitud subjetiva. Estas inducciones lingüísticas tienden a ser altamente exitosas – más aún que el curso general de las inducciones en nuestro bastante familiar mundo. La razón es que, mientras las inducciones de uno respecto a la naturaleza deben su éxito sólo a la congruencia aproximada entre los estándares de similitud de uno y el curso de los eventos en la naturaleza, por otra parte, las inducciones de uno acerca del asentimiento del hablante veterano ante la oración observacional deben su éxito a unos estándares de similitud compartidos entre el hablante y uno mismo. (Quine 1977, p. 274).

Así es como se explica – sin apelación a convenciones ni reglas lingüísticas, o quizá sería adecuado decir naturalizando estos conceptos – la relativa uniformidad en el uso de las palabras dentro de una misma comunidad. En todo caso, en lo que a nuestra cuestión afecta, nos interesa subrayar dos cuestiones:

Por una parte, no tiene sentido pensar que pueda haber algo así como una «regla semántica» o una convención en virtud de la cual el término ‘conejo’ se refiere unívocamente a los conejos. Como hemos visto, en el proceso de aprendizaje no entran en juego juicios semánticos de tipo alguno – sea de una teoría semántica «empírica» o

«popular» – , sino que se trata de interiorizar un patrón de aplicación de palabras o, dicho más quineanamente, de ver en qué situaciones los hablantes veteranos asienten ante cada oración dada. Y, repitamos, este patrón o este comportamiento lingüístico es describible tanto recurriendo a un manual de traducción homofónico como no homofónico.

Por otro lado, la existencia de un conjunto de reglas sintácticas, gramaticales, no eliminaría en absoluto la indeterminación. (Por otro lado, la propia gramática se halla indeterminada, aunque no abundaremos aquí en esta idea, v. nota 7). Supongamos, incluso, que se naciera con un conjunto de reglas lingüísticas ya interiorizadas – una gramática innata – , y que ésta fuera, por ejemplo, la del castellano: ¿qué impediría interpretar a otro hablante como refiriéndose a partes no separadas de conejo con su término ‘conejo’? Y lo mismo sucedería para la adscripción de creencias y demás actitudes proposicionales. Lo que determina, siempre, el significado de nuestras palabras son las condiciones en las que hemos aprendido éstas en el proceso triangulatorio ya descrito, existan reglas unívocamente fijadas o no. El propio Davidson llama la atención sobre que esto, la manera en la que se desarrolla el proceso interpretativo, es lo relevante en el estudio del lenguaje:

Yo quiero saber qué tiene el pensamiento proposicional – nuestras creencias, deseos, intenciones y lenguaje – que lo hace inteligible a otros. Esto es una cuestión acerca de la naturaleza del pensamiento y el significado que no se puede contestar descubriendo mecanismos neuronales, estudiando la evolución del cerebro, o hallando evidencia que explique la increíble facilidad y rapidez con la que llegamos a adquirir una primera lengua.

Incluso si todos nacióramos hablando inglés o polaco, sería una cuestión de cómo entendemos a los otros, y qué determina el contenido cognitivo de nuestras oraciones.

(Davidson 1995b, p. 133).

La conclusión que nos interesa extraer de aquí es que un conjunto de reglas sintácticas, constituyentes de una lengua, por innato que fuera, tampoco diría nada acerca de lo que los hablantes quieren decir con sus palabras en cada momento de emisión: «No importa con cuánta gramática venimos equipados desde la cuna, debemos aprender qué significan las palabras de cualquier lenguaje particular» (*ibid.*, p. 132), y lo que esas palabras significan viene dado por sus condiciones de aplicación y, de nuevo, éstas son susceptibles de ser descritas en el metalenguaje utilizando diferentes expresiones. «En la sintaxis no aparecen cuestiones acerca de la referencia», recordaba ya Davidson en fechas tempranas (1977, p. 221).

Recordemos (v. 1.2) que el hecho de que la relación de significación no tenga lugar en el nivel de los términos u otros elementos suboracionales, sino que lo haga en el nivel de las oraciones, permite atribuir diferentes referencias a los términos mientras que la estructura de la lengua permanece invariable. Reiteramos aquí una afirmación de Quine ya citada con anterioridad: «La estructura es lo que importa a una teoría, y no la elección de sus objetos» (Quine 1981, p. 20)

Además, la idea de que el hablante debe asignar a sus emisiones unas condiciones de verdad determinadas, y que en eso consiste saber el significado de las oraciones, se asienta en una concepción errónea de la labor de la semántica (v. 3.4.2): la idea de una

teoría semántica predecirá la conducta verbal de los hablantes, ya que ésta se rige por reglas. A este respecto dice M. Dummett:

He señalado muchas veces que una teoría del significado no debe ser considerada una sistematización científica de regularidades en fenómenos complejos; ha de ser juzgada en base a si ofrece una descripción inteligible de una práctica de agentes racionales. No estamos buscando una teoría con poder predictivo, sino una descripción que dé sentido a la actividad como una actividad realizada por seres racionales. (M. Dummett 1987, p. 260).

Hay, además, quienes han considerado que la explicación davidsoniana de la autoridad de la primera persona apelando a una asimetría interpretativa no constituye una explicación genuina. La idea de Davidson es, como hemos visto, que «[l]a asimetría se basa en el hecho de que el intérprete debe, mientras que el hablante no, confiar en lo que, si se hiciera explícito, sería una difícil inferencia al interpretar al hablante» (1984b, p. 13). La crítica, entonces, sería la siguiente:

Detenerse aquí, sin embargo, no es explicar la asimetría, ya que todo lo que hemos hecho es invocar otra asimetría, cuya relación con la asimetría de justificación permanece siendo un misterio. Que el intérprete debe guiarse por la evidencia conductual, mientras que el hablante no se guía por evidencia alguna, es, propiamente, parte de la descripción del problema, no su solución. (K. A. Ludwig 1994, p. 378).

Pero dicha crítica no hace justicia a la respuesta davidsoniana. En primer lugar, se debe decir que la exposición tradicional del problema de la asimetría se comprometía con una

imagen cartesiana de la mente (acceso privilegiado a contenidos mentales), al contrario que la davidsoniana. En segundo lugar, Davidson no sólo formula el problema, sino que ofrece una explicación; por un lado, señalar que existe una asimetría interpretativa, como hemos expuesto, no semántica, es ya una parte de la explicación, otra parte es la consideración de la autoridad como precondition de la interpretación. Es la conjunción de ambos elementos la que conforma la concepción davidsoniana de la autoridad de la primera persona sobre el significado de sus palabras.

3.5.2. La autoridad como precondition

Otro aspecto, pues, junto con la asimetría interpretativa expuesta, vertebra la concepción davidsoniana de la autoridad de la primera persona: dicha autoridad supone una presuposición necesaria para interpretar a un hablante. Si no suponemos que un individuo sabe lo que quiere decir, no podríamos continuar con el proceso de interpretación. Davidson introduce esta idea así:

Hay una presuposición – una presuposición inevitable que la interpretación lleva incorporada en su naturaleza – de que el hablante normalmente sabe lo que quiere decir. Así que hay una presuposición de que si sabe que considera verdadera una oración, sabe lo que cree. (Davidson 1984b, p. 14).

Asimetría interpretativa y la autoridad como precondition aparecen conjuntamente en el pensamiento de Davidson:

Tu conocimiento de lo que significan mis palabras tiene que estar basado en evidencia e inferencia: tú, probablemente, supones que lo has hecho correctamente y probablemente lo has hecho. Sin embargo, es una hipótesis. Por supuesto que puede que yo tampoco sepa lo que quiero decir con las palabras. Pero hay una *presunción* de que sí lo sé, ya que no tiene sentido suponer que estoy generalmente equivocado acerca de lo que significan mis palabras; la presunción de que generalmente no estoy equivocado acerca de lo que quiero decir es esencial al hecho de que tengo un lenguaje – para poder ser interpretado. (Davidson 1989a, p. 66).

Son varias las ideas que aparecen condensadas en los fragmentos citados, intentemos desglosarlas:

1. El conocimiento del significado de mis palabras no es inferencial, interpretativo. El conocimiento del significado de las palabras de otra persona, sí lo es.
2. Los hablantes saben normalmente qué oraciones consideran verdaderas.
3. Hay una presuposición necesaria para la interpretación de que un hablante sabe lo que quiere decir con sus palabras. Se trata de una presuposición «esencial» para poder interpretar a otro hablante.
4. La presuposición de que la primera persona sabe lo que quiere decir con sus palabras es «esencial» al hecho mismo de que posea un lenguaje.

5. No tiene sentido pensar que alguien está normalmente equivocado acerca de lo que significan sus palabras.

La premisa 1 es la premisa de la asimetría interpretativa de la que ya hemos hablado. La premisa 2 expresa la autoridad sobre las creencias. Las premisas 3 y 4 exponen la idea de la autoridad como precondition, la primera de ellas considera esta precondition necesaria para la interpretación mientras que la segunda la considera un factor constitutivo del hecho de hablar una lengua. La 5 es la premisa de la imposibilidad del error masivo. Éstas son las ideas, en su relación con la ideterminación del significado, que habrán de ser aclaradas en la presente exposición.

Thöle bautiza 2 como la tesis de la presuposición de verdad (*truth presumption thesis*) y la caracteriza así:

- Hay una presuposición de que las autoascripciones en primera persona son verdaderas en general, aunque no universalmente.

La premisa 3 contendría dos importantes ideas; llamemos a la primera la tesis de la interpretación (*interpretation thesis*):

- la autoridad de la primera persona «es esencial para» o se halla «dentro de la naturaleza de la interpretación».

Esta idea nos llevaría a la tesis de la no contingencia (*non-contingency-thesis*):

- La autoridad de la primera persona no es un hecho contingente.

Se llama ahora la atención sobre el hecho de que 2 y 3 son más diferentes de lo que pueden parecer a simple vista. Se trata, según Thöle, de dos tipos de asimetrías distintos. La premisa 2 habla acerca de la asimetría de la adscripción de creencias: un hablante sabe qué oraciones considera verdaderas, y 3 nos habla de una asimetría de significados: un hablante sabe el significado de las palabras que utiliza. Thöle denomina la primera ‘asimetría de actitud’ y la segunda ‘asimetría de significado’ y ejemplifica la situación así: supongamos que un hablante emite la oración ‘Wager murió feliz’; el hablante (a) quiere decir algo con su oración y (b) cree que ello es verdadero, luego el hablante ha expresado una creencia. Davidson afirma aquí que si tanto el hablante como el oyente saben (i) que el hablante considera verdadera la oración y (ii) lo que el hablante quiere decir con la oración, entonces ambos saben qué creencia se expresó. Davidson admite que hay asimetría entre el hablante y el oyente tanto respecto a la certeza de (i) como de (ii). El problema es que Davidson se centra en (ii), considerando, por alguna razón no expuesta, que solucionando la asimetría del significado queda solucionada la de la actitud. Thöle expresa esta crítica así:

En primer lugar, Davidson considera sólo un tipo de caso muy especial – un caso en el que la atribución de creencia se refiere a una creencia que es expresada verbalmente. Pero, claro está, la autoridad de la primera persona no se halla restringida a las creencias que un hablante expresa verbalmente.

En segundo lugar, Davidson parece comprometerse con la extraña idea de que en el tipo de situación bajo consideración, el conocimiento de un hablante de qué creencia expresó *se deriva* de su conocimiento (anterior) de (i) y (ii) [Nota al pie: Esto es así debido a que su explicación depende de la afirmación de que el hablante sabe lo que cree *en virtud* de su conocimiento de (i) y (ii)]. Pero un hablante sabe normalmente qué creencias tiene antes de que esté dispuesto a expresarlas. [Nota al pie: Davidson subraya este punto obvio en su crítica a Ryle (...), pero parece hacer caso omiso de ello cuando presenta su propia explicación.] (B. Thöle *ibíd.*, p. 239).

Efectivamente, Davidson ha afirmado explícitamente que la forma de solucionar el problema de la autoridad respecto a las creencias es solucionándolo respecto al significado. Davidson «defiende que tenemos algún tipo de autoridad especial sobre nuestras creencias y pensamientos, pero, contrariamente a las explicaciones epistémicas, él pretende explicar esta autoridad lingüísticamente[.]» (K. Puhl 1994, p. 342). Así, dice:

Asumo (...) que si la autoridad de la primera persona en el lenguaje puede ser explicada, hemos hecho mucho, si no todo, de lo que se necesita hacer para caracterizar y explicar el hecho epistemológico[.] (Davidson 1984b, p. 3)

Efectivamente, parece extraño mantener que un hablante sabe qué creencia tiene *porque* sabe qué oración considera verdadera y lo que ésta significa. De aquí se derivaría que no podríamos tener ninguna creencia – o saber que la tenemos – si nunca nos hemos planteado si consideramos verdadera la oración que la expresa. Ya que Davidson admite que la asimetría respecto al conocimiento de los significados se

asienta en que la primera persona no precisa autointerpretarse, no parece coherente pensar que, para saber lo que cree, sí necesita observar su conducta verbal, ver qué oraciones considera verdaderas. Ésta es la postura adoptada por Ryle, postura que niega la existencia de asimetría o autoridad alguna, que Davidson afirmaba rechazar (v. 2. 4.) Así, Thöle insiste en su crítica:

Davidson está, con seguridad, en lo cierto cuando afirma que el conocimiento de (i) y (ii) es suficiente para saber qué creencia expresa un hablante; pero él parece asumir – sin ninguna explicación – que saber (i) y (ii) es *la única manera* de llegar a un conocimiento de qué creencia tiene un hablante; y esto es, con seguridad, falso. (*Ibid.*, p. 240).

No nos detendremos aquí en esta cuestión acerca de la relación entre una autoridad de actitud y una lingüística, ya que nos interesa la segunda, sólo introduciremos alguna consideración que resulta de relevancia para nuestro estudio. Sin entrar, pues, en el debate acerca de la posibilidad de pensamientos preverbales, podemos decir que la crítica se basa en el mismo supuesto que la de Searle, aplicada ahora a las creencias. Searle decía que si yo sé lo que quiero decir con mis palabras y éstas tienen un significado, entonces yo sé el significado de las palabras, pero para atacar la indeterminación, Searle debía partir de que las palabras tienen un significado determinado – o la primera persona lo puede determinar de manera unívoca –. Es decir, Searle asumía que hay *algo* que es el significado o la referencia de las palabras. Ahora hallamos que si yo sé qué oraciones considero verdaderas y sé lo que éstas significan, sé lo que pienso. La apelación a los pensamientos no expresados verbalmente nos vuelve a

comprometer con la idea de que hay algo que es mi pensamiento. Recordamos aquí que, en la visión davidsoniana, la adscripción de creencias tiene el mismo papel que la de significado o referencia. Adscribimos creencias a los demás en el mismo proceso interpretativo en el que adscribimos significado a sus palabras, siempre con el objetivo de entablar comunicación con ellos, para lo que debemos presuponer su racionalidad. Debemos recordar que la evidencia disponible a un intérprete radical es de carácter *no individuativo*, no hay en dicha conducta significados y pensamientos individuados. Esto es tarea del proceso interpretativo. La interpretación asigna significados y creencias a partir de la conducta verbal y no verbal.⁴⁶

Volvamos sobre la cuestión de la autoridad respecto a los significados de las propias palabras. Parece claro que si no se presupone que un hablante sabe lo que quiere decir con sus palabras, resultaría ininterpretable. Así pues, la autoridad de la primera persona constituye una presuposición necesaria dentro del proceso de interpretación. Thöle vuelve a ver problemas en esta idea:

Pero Davidson no nos da ninguna razón de por qué es *una precondition para la interpretación* que un hablante *sabe* lo que quiere decir. Todo lo que dice apoya sólo la idea más débil y trivial de que la interpretación requiere interpretabilidad, esto es, un uso coherente del lenguaje. (*Ibid.*, p. 243).

⁴⁶ Davidson se refiere a menudo a su similitud, en este punto, con los autores de la teoría de la decisión, en especial, F. Ramsey. Éstos ofrecen un modelo para distinguir los deseos y las creencias en la conducta de un individuo. Para Ramsey no debemos conformarnos con ofrecer la intensidad de las creencias en base a sentimientos internos de convencimiento o algo similar. El grado en el que alguien cree algo *es* el punto hasta el que alguien actúa basado en ello.

El propio Thöle esquematiza la que parece ser la argumentación davidsoniana:

- a) Las emisiones de un hablante son (en general) interpretables, luego
- b) Las emisiones de un hablante son (en general) significativas, luego
- c) Un hablante (en general) quiere decir algo con sus emisiones, luego
- d) Un hablante sabe (en general) lo que quiere decir con sus emisiones.

Thöle cree ahora que «El paso de b) a c) no es, ciertamente, evidente. Pero el paso más problemático es de c) a d)»⁴⁷. ¿Por qué se permite dar Davidson el paso del hecho de que un hablante quiera decir algo con sus emisiones a que sepa lo que quiere decir?

⁴⁷ Como se ve, la idea de que el paso b)-c) es controvertido es la misma que hemos visto en la sección anterior en B. C. Smith: la de que Davidson «no ve diferencia entre alguien queriendo decir algo y ese alguien sabiendo lo que dice». A la idea subyace la concepción de que el hablante podría estar utilizando palabras cuyo significado no conoce. La respuesta la hemos visto: las palabras no tienen significado *per se*, éste se constituye en el mismo proceso interpretativo. Lo que Davidson quiere expresar con «saber lo que se quiere decir» es el hecho de que aplicamos nuestras palabras en las situaciones que creemos adecuadas y, dado que esas situaciones son las que determinan el significado de nuestras palabras, es por principio imposible que un hablante utilice sus oraciones de forma equivocada habitualmente. De ahí que afirme que un hablante «no está en posición de preguntarse si está aplicando sus palabras, generalmente, a los objetos y eventos adecuados, ya que sea lo que sea a lo que regularmente las aplica, esto le da a sus palabras el significado que tienen» (Davidson 1987, p. 37). Una vez que se tiene en mente que el proceso de triangulación «no es sólo una historia acerca de cómo aprendemos a usar las palabras: debe ser también una parte esencial de una explicación adecuada de a qué se refieren las palabras y qué significan.» (Davidson 1988, p. 44), no parece posible que alguien quiera decir algo con sus palabras y, sin embargo, no sepa qué quiere decir – puesto que lo que quiere decir es viene determinado, precisamente, por las circunstancias en las que las aplica –. Esta crítica se puede ver, como decimos en Thöle, B. C. Smith y, más recientemente, en M. F. Fricke:

Davidson sólo demuestra que no es posible que el hablante (generalmente) haga emisiones incorrectas para decir lo que quiere decir. Pero eso no equivale a decir que el hablante sabe qué significan sus emisiones. Eso sólo demuestra que el hablante dice lo que quiere decir, no que sabe qué quiere decir o que sabe qué significan sus palabras. (M. F. Fricke 2007, p. 67).

Identificando el hecho de saber lo que se quiere decir con utilizar correctamente el lenguaje:

a menos que haya una presuposición de que el hablante sabe lo que quiere decir, esto es, de que está utilizando correctamente su lenguaje [*is getting her language right*], no habría nada que interpretar para el intérprete. Por decirlo de otra manera, nada podría contar como que alguien utiliza sus propias palabras de forma incorrecta habitualmente [*as someone regularly misapplying her own words*]. (Davidson 1987, p. 38)

Thöle no considera aceptable dicha identificación. Así, pone el ejemplo de un hablante que se mostrara agnóstico acerca de si dos expresiones cuyo significado conoce son sinónimas: no estamos justificados para decir que este hablante no es interpretable. Por otra parte, se dice que no está claro qué significa en esta visión davidsoniana «saber lo que uno quiere decir»: si estamos exigiendo con ello que los hablantes puedan explicar lo que dicen de otra manera, es evidente que muchas veces no podemos y no dejamos por ello de resultar interpretables.

Ofrecemos aquí una lectura de los argumentos davidsonianos que intentan responder las críticas expuestas arriba. Recordemos el proceso de interpretación, y tengamos presente que es mediante este proceso, la triangulación (v. 3.3.1.), como (1) se aprende un lenguaje y (2) se interpreta a otros – se trata de dos aspectos de un mismo proceso –. Desde la perspectiva de la primera persona, durante el proceso se aprende a aplicar

palabras a objetos y situaciones y *así* se generan también creencias.⁴⁸ Es desde este punto de vista desde el que tiene sentido afirmar que sé lo que significan mis palabras porque sé qué oraciones considero verdaderas, ya que en el proceso de triangulación aprendemos lo que significan las palabras viendo qué debe suceder para que sean aplicables, esto es, qué las hace verdaderas. Y es desde este punto de vista, también desde el que resulta lícito el paso de c) a d): aplicar las palabras a las situaciones en las que se cree que se han de aplicar es ya saber lo que hay que saber sobre el significado de éstas. Considerar espurio dicho paso es algo que sólo se puede hacer partiendo de que debe de haber «algo más» para el significado de las oraciones que las condiciones de aplicación en un momento dado (v. 3.5.1). Sin embargo, si consideramos que aprender una lengua es aprender a aplicar oraciones y que, por tanto, un hablante, por el hecho de serlo, sabe aplicar sus palabras, entendemos por qué Davidson considera la autoridad esencial al hecho de tener una lengua.

Como vemos, es el aspecto externista del proceso de triangulación el que ofrece la conexión necesaria entre verdad y significado⁴⁹, y el que abre la puerta a la explicación

⁴⁸ Insistamos una vez más en la inextricabilidad de ambos elementos: no se pueden aprender significados sin formarse a la vez creencias. Ya hemos citado las palabras de Quine (1.7): «No aprendemos primero de qué hablar y luego qué decir sobre ello» (Quine 1960, p. 16). Siendo esto así, no se puede considerar verdadera una oración sin saber su significado, no se puede saber su significado sin formarse una creencia.

En realidad sí se puede considerar verdadera una oración sin saber su significado: podemos confiar en alguien que nos asegura que una determinada oración de una lengua que no conocemos es verdadera, pero eso no tiene nada que ver con la interpretación.

⁴⁹ No podemos detenernos aquí en las ideas davidsonianas acerca del concepto de verdad. Lo relevante para nosotros es la forma en la que el proceso de triangulación pone en contacto la verdad con el significado. Davidson afirma a este respecto: «La buena traducción preserva la verdad, la verdad es lo que se preserva en la buena traducción, si no oración por oración, sí en el conjunto. Lo mejor que podemos hacer para romper este círculo es conectar la verdad o la traducción con las reacciones observables de los hablantes al lenguaje y al mundo. (...) [A]quí [en Quine] la conexión con la verdad es evidente: alguien que asiente sinceramente a una oración la considera verdadera – cree que es verdadera.» (1995, p. 15).

Davidsoniana de la autoridad de la primera persona. El aspecto externista, que consiste en subrayar el hecho de que «no podemos, en general, identificar primero creencias y significados y preguntar después qué los causó» (Davidson 1983, p. 131), ya que eso que los causa es, precisamente, el contenido de las creencias y el significado de las palabras; Davidson lo expresa así: «Si algo está sistemáticamente causando ciertas experiencias (o respuestas verbales), eso es sobre lo que son los pensamientos y las emisiones.» (1982, p. 322).

Y debemos señalar de nuevo que la autoridad no entra en contradicción con la indeterminación: la primera persona sabe lo que quiere decir con sus palabras de manera diferente a como lo sabe la segunda, pero ésta podría siempre describir el significado de las emisiones de aquélla de varias formas diferentes. Ello no afecta a la autoridad. Consideremos qué determina el significado de la emisión de un hablante ‘ahí hay un conejo’. Lo determina, como hemos dicho, la causa de su emisión, ahora bien, ésta es tanto un conejo como una parte no separada de conejo.⁵⁰

Quienes aún puedan insistir en que Davidson, apelando a una asimetría interpretativa, simplemente ha reformulado la cuestión de la autoridad pero no lo ha explicado, pueden

Para Davidson, apelar al estímulo distal en vez de al estímulo próximo tiene como resultado «salvar la relación natural entre significado y verdad.» (*Ibid.*, p. 19).

⁵⁰ Conviene señalar aquí la diferencia esencial entre el externismo de Davidson y los de Putnam y Burge. Para éstos últimos, el significado viene determinado por el uso de la comunidad científica o de la comunidad lingüística en general (al menos, en lo que respecta a los términos que denotan clases naturales), respectivamente; Davidson salva la autoridad de la primera persona apelando únicamente a cada hablante particular en una situación de triangulación. Hemos de ver esto en relación con el rechazo davidsoniano a las convenciones lingüísticas como factores relevantes en el proceso de interpretación (v. 3.3.3): ¿de qué utilidad resulta saber lo que quieren decir otros con sus palabras a la hora de interpretar a un hablante dado?

concebir la estrategia davidsoniana aquí *à la* Wittgenstein: el problema no se resuelve, se disuelve. Y esto porque, al llamar la atención sobre el proceso de triangulación, se ve cómo saber lo que uno quiere decir con sus palabras entra dentro del concepto de tener una lengua. En este sentido... no hay mucho más que decir. Las premisas 1, 2, 4 y 5 van así de la mano en el proceso triangulatorio.

Así concebido el proceso de aprendizaje y uso de la lengua, observamos unidos tres aspectos que parecían difíciles de conjugar:

La autoridad de la primera persona, el carácter social del lenguaje, y los determinantes externos del pensamiento y el significado van juntos de forma natural, una vez que abandonamos el mito de lo subjetivo, la idea de que los pensamientos requieren objetos mentales. (Davidson 1987, p. 38).

Volvamos nuestra vista ahora sobre la premisa 3. La autoridad de la primera persona como precondition de la interpretación plantea una cuestión diferente. Cuando hablamos de las condiciones de la interpretación nos referimos a aquello que debe ser asumido, necesariamente, para que el proceso pueda, siquiera, comenzar. La precondition por antonomasia del proceso interpretativo la constituye el principio de caridad (v. 1. 7.). Es fácil ver por qué sería imposible comenzar a interpretar a alguien si no *presuponemos* que se trata de un agente racional – según nuestros criterios, al menos en los primeros estadios de la interpretación – y sus emisiones son generalmente sinceras (v. nota 19). Así, partiendo de estos supuestos, intentamos traducir sus

emisiones, y la cuestión es similar respecto al conocimiento de los significados de las propias palabras.

Como hemos visto, conocer los significados de las palabras de uno es algo que entra en el mismo concepto de poseer una lengua, y un proceso de interpretación no puede comenzar si no se presupone que el sujeto habla una lengua – es decir, que hay algo que interpretar – . Imaginemos a un intérprete radical comenzando su tarea en el terreno: presta atención a las emisiones de los hablantes y a sus percepciones en el momento de sus emisiones, pero ¿a dónde le llevaría eso si no presupone que se está hablando una lengua, esto es, que se están aplicando palabras de forma coherente?

Interpretar consiste en hallar patrones coherentes en el uso de las palabras de otro hablante, hablar una lengua consiste en aplicar patrones que otro hablante pueda considerar coherentes, y la autoridad consiste en aplicar esos patrones: interpretabilidad, lingüisticidad y autoridad aparecen como tres aspectos ligados en el hecho mismo de la comunicación.

Así, parece que siguiendo el extensionalismo davidsoniano, que nos lleva a concebir el dominio de una lengua como la aplicación coherente (según los criterios de al menos otro hablante) de palabras, hemos hecho depender la autoridad de la primera persona del éxito interpretativo de una tercera persona. Y así es. Debemos recordar, y subrayar con ello la inevitabilidad de la indeterminación, que lo que la concepción davidsoniana muestra aquí es que no hay una noción de significado independiente de una

interpretación. En una línea davidsoniana, podríamos decir: *yo sé lo que quiero decir porque *tú* puedes interpretarme.*

Conclusiones y Perspectivas

1. Conclusiones

Resumimos aquí, sucintamente, las conclusiones de este trabajo. La idea principal que se debe extraer es que la apelación a una supuesta autoridad de la primera persona respecto al significado de sus oraciones o la referencia de sus términos no resulta un argumento satisfactorio contra la tesis de la indeterminación o la inescrutabilidad.

La idea de Searle, y de quienes en esta línea lo han seguido, supone un malentendido de la tesis quineana. Resumamos brevemente lo ya dicho. La insistencia en que yo sé que con mi término ‘conejo’ me refiero a los conejos y no a las partes de éstos no separadas se debe a que se toma el término independientemente de la oración de la que forma parte. La tesis, sin embargo, subraya que la función de los términos consiste en indicar las entidades referidas dentro de una oración, y es aquí donde adquiere su relevancia la posibilidad de aplicar las llamadas funciones vicarias. Lo que éstas *demonstran* es que se puede variar la referencia de los términos, variando también la interpretación de otros elementos suboracionales si es necesario, sin que se alteren las condiciones de uso de las oraciones.

Si es posible interpretar el término ‘conejo’ de las dos formas antes indicadas, por ejemplo, manteniéndose invariables los valores de verdad de todas las oraciones, ¿afirma la tesis que un conejo es lo mismo que una parte no separada del conejo?, se preguntaba Searle y, como hemos visto, temió el propio Quine. Y es aquí donde Searle aprecia el carácter autorrefutante de la inescrutabilidad: percibimos diferencias en la

referencia de los términos que, sin embargo, la tesis niega. Es más, la tesis necesita que percibamos esas diferencias para poder ser formulada. Estamos, de nuevo, ante la confusión anterior. ‘Conejo’ y ‘parte de conejo no separada’ tienen, como oraciones, unas condiciones de uso determinadas (aunque idénticas): ‘Conejo’ se aplica a los conejos y ‘Parte de conejo no separada’ se aplica a una *parte* de conejo no separada del conejo. Ahora bien, la inescrutabilidad se produce sólo cuando pasamos a considerarlos términos, no oraciones de una sola palabra. Es entonces cuando se hace posible aplicar las funciones vicarias. Podemos afirmar, pues, que la crítica searleana comete dos errores fundamentales y conectados: por un lado, no se ha entendido la función de los términos como elementos suboracionales y, por ello, se pretende tomarlos de forma aislada del resto de la oración. Es esto lo que conduce a la idea de que el sujeto debe saber a qué entidades se refiere con sus términos. Y, por otro lado, no se entiende la diferencia entre considerar una palabra como una oración – con unas condiciones de uso – o como un término – que, insistimos, sólo cumple una función en el contexto de la oración –. Y es este malentendido el que lleva a afirmar que la tesis cae en el absurdo de no ver diferencia entre un conejo y sus partes.

Se puede seguir insistiendo en que, aún dentro de una oración, los términos deben tener una referencia determinada. Es decir, negar el principio de oracionalidad, según el cual son las oraciones los vehículos del significado, no sus partes constituyentes, y ver los términos como poseedores de una referencia. Referencia determinada, incluso, por ciertas convenciones del lenguaje; ésta parece ser la idea de Searle cuando afirma que si sabe que ‘lapin’ se refiere a los conejos es, simplemente, porque habla francés. Pero

negar de esta manera que las funciones vicarias suponen una demostración de la inescrutabilidad sólo se puede hacer defendiendo una capacidad mágica de hacer referencia. Hemos defendido aquí la crítica que a este respecto le dirigiera M. Devitt a Searle. Dado que las funciones vicarias nos permiten cambiar la referencia de los términos sin que ello afecte a las condiciones de uso de las oraciones de un sujeto, ¿cómo podría éste determinar la referencia? Es inevitable pensar aquí en teorías mágicas de la referencia, apelando a extraños poderes de la mente. Y a algo muy similar parece abocada la poco desarrollada teoría de Searle del Contenido Intencional.

Hemos dedicado parte de nuestro trabajo al análisis de la que supone la última versión de una crítica *à la* Searle a la tesis de la indeterminación, la que apela, no a contenidos intencionales, sino a las intuiciones de los hablantes respecto al significado de sus palabras. La idea vendría a ser ahora que no tiene sentido que los hablantes competentes no sepan cuál es el significado que le dan a sus palabras. D. Bar-On ha insistido también en la idea searleana de que para entender la tesis de la indeterminación necesitamos de esa «semántica intuitiva», dado que debemos ver una diferencia entre hacer referencia a conejos o a partes de éstos no separadas, diferencia que la tesis niega. Así, se afirma que «en su argumentación de la tesis de la indeterminación, Quine debe apelar a un contraste entre dos perspectivas en el lenguaje – la «perspectiva del usuario» y la «perspectiva del teórico» (D. Bar-On 1993, p. 801). En la segunda perspectiva mencionada, sólo tenemos las oraciones y sus condiciones de uso, pero es la primera la que nos permite ver la diferencia entre referirse a conejos o a partes de éstos no

separadas. Como vemos, se repiten los errores de Searle para entender la tesis que ya hemos comentado arriba.

Nuestra argumentación respecto a este «intuicionismo semántico» ha adoptado dos líneas: por un lado, se rechaza que las supuestas intuiciones resulten de utilidad para la tarea de la semántica y, por otro, se defiende la visión davidsoniana de la autoridad de la primera persona, según la cual los hablantes, efectivamente, saben qué significado le dan a sus palabras, pero esto no entra en contradicción con la indeterminación.

Respecto al primer punto citado, tengamos ahora presente la idea de D. Bar-On de que la tarea de la semántica es sistematizar los juicios sobre el significado de sus palabras que realizan los hablantes competentes. Es una idea coherente si se acepta la anterior premisa acerca de la utilidad de las intuiciones lingüísticas de los hablantes. Hemos contestado a esto en nuestro trabajo con una doble argumentación. Por una parte, hemos señalado que las intuiciones nunca son un elemento de valor científico indiscutible. No lo son en ciencia alguna. Pensemos, incluso, en la psicología; el hecho de que los individuos tengan una mente no los hace *eo ipso* buenos conocedores de ésta. Es de suponer que la situación de las demás ciencias se repite respecto a un mecanismo de tan complicado funcionamiento como un lenguaje natural. ¿Por qué, preguntamos nosotros, sería de más utilidad preguntar a los hablantes si consideran sinónimas dos expresiones que observar si las utilizan de manera intercambiable? ¿No son las nociones de la semántica – como «sinonimia» – nociones técnicas, desconocidas para la mayor parte de los hablantes? Pensemos en la tesis que nos ocupa: ¿qué diría un lego en la materia

acerca de la aplicación de una función vicaria para sustituir el discurso de conejos por uno de estadios de conejo? En principio, no parece que las intuiciones lingüísticas de los hablantes sean de mucha utilidad para confeccionar un manual de traducción.

Respecto al segundo punto arriba mencionado, nuestra defensa de la teoría davidsoniana de la autoridad de la primera persona respecto al significado de sus palabras, hemos subrayado aquí la importancia de la llamada «triangulación». Davidson ha resaltado mediante esa idea el carácter esencialmente intersubjetivo del lenguaje y, con ello, ha arrojado luz sobre la manera en que las palabras adquieren su significado. Atender a este proceso triangulatorio nos ha resultado especialmente valioso para analizar la llamada teoría «de la teoría» del lenguaje. Si bien es cierto que tanto Quine como Davidson – y autores tan distantes de éstos en su concepción del lenguaje como Chomsky – se han referido al aprendiz del lenguaje (un niño o un traductor de una lengua desconocida para él) como alguien que está confeccionando una teoría, y que, además, es en el marco de la confección de dicha teoría donde adquiere sentido hablar de indeterminación, no se refutará la indeterminación por negarle al aprendiz el carácter de teórico del lenguaje. Efectivamente, hablar del aprendiz del lenguaje como de alguien confeccionando una teoría no supone más que hiperintelectualizar la labor de su aprendizaje. Ha sido el propio Davidson quien más ha insistido en que lo que se aprende al aprender un lenguaje es la habilidad práctica de interpretar; nuestra teoría sólo pretende describir ese proceso. El aprendiz no confecciona ni expresa ni tácitamente dicha teoría. Ahora bien, hemos visto también cómo esto no afecta a la indeterminación:

una vez que queremos expresar lo que otro hablante nos intenta comunicar en el proceso triangulatorio, la indeterminación se hace insoslayable.

En la triangulación se hace patente que, al menos en los casos más básicos, las palabras adquieren para un sujeto el contenido que le dan las situaciones en las que ha aprendido de otro hablante que dichas palabras son aplicables. Efectivamente, no se aprende teoría alguna, en sentido estricto; se aprende a realizar ciertas emisiones en ciertas circunstancias. Es planteando así la cuestión cómo, por un lado, se preserva sin dificultad la autoridad de la primera persona: ¿cómo podría el hablante no saber lo que quiere decir si simplemente está aplicando las palabras de la manera en la que aprendió a hacerlo?, y, por otro, la indeterminación resulta inevitable, dado que cuando se desea explicar en qué situaciones un hablante utiliza una determinada expresión, podemos utilizar diferentes expresiones: ¿qué más da decir que un hablante utiliza su expresión ‘hay un conejo’ cuando hay un conejo presente o cuando hay una parte de conejo no separada?

2. Perspectivas

Finalizamos, definitivamente, con unos comentarios acerca de las reflexiones que se abren a partir de lo expuesto en nuestro trabajo. Abrimos en esta breve sección dos cuestiones: ¿desemboca la concepción davidsoniana de los significados y las creencias, a su pesar, en una teoría conductista del tipo de la de Dennett?, ¿resulta satisfactoria la explicación davidsoniana de la primera persona más allá de la aplicada a los significados, como se ha defendido aquí?

Una de las cuestiones que más solivianta a los críticos de la indeterminación es que ésta nos aboca presuntamente a una especie de «nihilismo semántico». Davidson rechaza la reificación de significados y creencias, pero no les niega realidad objetiva. La cuestión aquí es lo que hemos llamado «la teoría de los invariantes»: lo que una teoría semántica trata de describir son las situaciones en que un hablante sostiene una oración como verdadera; esas situaciones – de cuya realidad no tiene sentido dudar – son lo que da contenido a las palabras. Así, por ejemplo, el holismo resulta el gran generador de indeterminación pero no implica la disolución de los significados, sólo la interdependencia de éstos en un sistema de complejas relaciones. Eso sí, no se implica la disolución de significados (o actitudes proposicionales) así entendidos, sin entidades cartesianas:

La indeterminación de la traducción quiere decir que diferentes conjuntos de emisiones (u oraciones, como Quine lo expone) funcionan igualmente bien al interpretar el lenguaje de un hablante (o los pensamientos); esto no sugiere que los estados de la mente del hablante o del sujeto que cree así capturados son, de alguna manera, vagos o irreales. El holismo mantiene que los contenidos del lenguaje y del pensamiento dependen de las relaciones entre los significados y entre los pensamientos. Pero, de nuevo, no hay nada en esta afirmación que amenace la realidad de los estados que están relacionados de estas maneras. La amenaza a la realidad de los pensamientos y los significados que Searle y Fodor creen detectar es, de hecho, una amenaza muy diferente, una amenaza a la asunción de que las entidades usadas para identificar los pensamientos y los significados son, de alguna manera, «captadas» por la mente, y así, si las entidades son diferentes, los pensamientos mismos han de ser diferentes. Es como si la

«diferencia» entre que un campo mida una yarda o 36 pulgadas fuera una diferencia en el campo mismo. (Davidson 1989a, p. 65).

Con todo, podemos preguntarnos si Davidson consigue salvar la realidad de las actitudes proposicionales; podemos plantearnos si, en el marco que Davidson traza para el surgimiento de la comunicación y el pensamiento, el proceso de triangulación, no sería más adecuado concebir las actitudes proposicionales como patrones de conducta, es decir, *à la* Dennett. Davidson nos ha ofrecido una concepción de la indeterminación no unida, como en Quine, al fisicalismo. Davidson, al contrario que éste, no considera nuestro lenguaje mentalista un residuo de una psicología precientífica a la espera de ser reducido a los términos de la física.⁵¹ De hecho, su analogía de la indeterminación con diferentes sistemas de medida ha llevado a D. Dennett a pensar lo contrario de lo que afirman la mayoría de filósofos, esto es, que la exposición davidsoniana hace las actitudes proposicionales demasiado reales. Efectivamente, si hay algo que se puede representar de diferentes maneras... es que hay algo. D. Dennett considera que lo único *real* es la conducta observable, las actitudes proposicionales vendrían a ser patrones distinguibles en la conducta – distinguibles por parte de un observador –.⁵² Es un punto

⁵¹ No es este el lugar para exponer la filosofía de la mente de Davidson, sólo señalar que su postura es la conocida como monismo anómalo. Éste defiende que los sucesos mentales son idénticos a sucesos físicos, pero que no es posible hallar leyes que los conecten. La identidad de lo mental y lo material se daría en los ejemplares, pero no en los tipos. Davidson considera que esta posición reconcilia muchas de las consideraciones de Quine sobre las actitudes proposicionales con el carácter indispensable del lenguaje mentalista: «El monismo anómalo le da sentido a la afirmación de que las actitudes son disposiciones para actuar de determinadas maneras, que son, a su vez, estados fisiológicos, que son, finalmente, estados físicos, tanto como a la afirmación de que las descripciones intencionales no son reducibles a descripciones conductuales o físicas, y, así, no son aptas para su incorporación a ningún sistema estricto de leyes. El vocabulario mental es práctico e indispensable, pero no está hecho para la ciencia más precisa.» (Davidson 1997, p. 72).

⁵² V. Dennett, «Real Patterns», *The Journal of Philosophy* 88, 1, enero, 1991, pp. 27-51.

que merece reflexión – ¿es la teoría de Davidson, simplemente, la teoría de Dennett? – .
En todo caso, aportamos aquí la siguiente reflexión en esta misma línea:

Establecer una analogía con la teoría de la medida no presenta la indeterminación como anodina. Por el contrario, hace incluso más evidente una consecuencia de combinar la afirmación de que la interpretación se halla indeterminada con la idea de que las actitudes proposicionales del interpretado no pueden trascender la evidencia del intérprete: sólo este par de principios es suficiente para implicar que no hay estados proposicionales. Analogía o no, la identidad de actitudes proposicionales entre esquemas falta, y las actitudes proposicionales se pierden. (P. Rawling 2001, p. 249).

Efectivamente, P. Rawling ofrece buenos argumentos a favor de la idea de que la indeterminación combinada con la aceptación como única evidencia legítima de la evidencia observable conlleva la eliminación absoluta de las actitudes proposicionales. La analogía davidsoniana de los significados con sistemas de medición, pues, no resultaría inocua, sino que tendría una enorme relevancia para nuestra concepción de la mente, ya que estaríamos ante un conductismo y un antimentalismo radicales. Así, el rechazo explícito de Davidson a la idea de Dennett resultaría un añadido injustificable dentro de la filosofía davidsoniana. La cuestión es que «si hay indeterminación genuina, entonces, incluso si el uso de oraciones para rastrear [*track*] actitudes proposicionales es análogo al uso de números para describir pesos o longitudes, no habrá, en general,

identidad de actitudes proposicionales entre esquemas, y, por tanto, no habrá actitudes proposicionales.» (*Ibid.*, p. 249).⁵³

Hemos defendido aquí la concepción davidsoniana de ésta respecto a los significados, esto es, basada, por un lado, en que la primera persona no precisa interpretarse a sí misma y, por otro, en que se debe presuponer que otro hablante, para ser tal, sabe lo que quiere decir con sus palabras, ahora bien, la cuestión se vuelve más intrincada cuando pasamos de los significados a las creencias. ¿Cómo sabe un sujeto, sin necesidad de observarse a sí mismo, lo que cree? Davidson ha apelado a la autoridad respecto a los significados para explicar la autoridad respecto a las creencias. La idea es, básicamente, que un sujeto sabe lo que cree porque sabe lo que significan sus palabras: «el hablante sabe normalmente lo que quiere decir. Así que (...) si sabe que mantiene una oración como verdadera, sabe lo que cree.» (Davidson 1984b, p. 14). ¿Se trata, pues, de una inferencia: infiere el sujeto lo que quiere decir a partir de su conocimiento de qué oraciones considera verdaderas y qué significan éstas? Esto es muy plausible respecto a nuestro conocimiento de las creencias de los demás hablantes: «Sabiedo que [el hablante] mantiene como verdadera una oración y sabiedo el significado, podemos inferir su creencia» (Davidson 1973, p 134), pero dicha inferencia no parece darse en el caso de la primera persona. Queda, pues, abierta la cuestión respecto a las creencias.

⁵³ P. Rawling, sin embargo, niega que el holismo, por sí solo, engendre antimentalismo, ya que el que las actitudes proposicionales se presenten siempre en interconexión no resta peso ontológico a éstas: «El holismo, per se, pues, no supone amenaza alguna para las actitudes proposicionales. Su desaparición es una consecuencia de la indeterminación y la dependencia de la evidencia.» (*Ibid.*, pp. 249-250).

REFERENCIAS

- Alston, William P. 1986 «Quine on meaning», en Hahn, L. E. y Schlipp, P. A. (eds.) 1986, pp. 49-72.
- Arrington, R. L y H-J. Glock (eds.) 1996 *Wittgenstein and Quine*, Routledge.
- Bar-on, D. 1986 «Semantic Indeterminacy and Scientific Underdetermination», *Pacific Philosophical Quarterly* 67, 1986, pp. 245-263.
- 1990 «Scepticism: The External World and Meaning», *Philosophical Studies* 60, pp. 207-231.
- 1992 «Semantic Verificationism, Linguistic Behaviourism, and Translation», *Philosophical Studies* 66, pp. 235-259.
- 1993 «Indeterminacy of Translation – Theory and Practice», *Philosophy and Phenomenological Research*, Vol. LIII, nº. 4, Diciembre, pp. 781-810.
- 1997 «‘Natural’ Semantic Facts – Between Eliminativism and Hyper-Realism», en D. Jutronic (ed.) 1997, pp. 99-117.
- 2004 «Semantic Eliminativism and the ‘Theory’-Theory of Linguistic Understanding», en C. Viger *et al.* (eds.) 2004, pp. 159-200.
- Beisecker, D. 2003 «Interpretation and First-Person Authority: Davidson on Self-Knowledge», *Southwest Philosophy Review*, vol. 16, nº. 1, enero, pp.89-96.
- Bonk, T. 2008 *Underdetermination. An Essay on Evidence and the Limits of Natural Knowledge*, Boston Studies in the Philosophy of Science, vol. 261, Springer Netherlands.
- Boolos, G. (ed.) 1990 *Method, Reason and Language: Essays in honour of Hilary Putnam*, Cambridge University Press, Cambridge.

-
- Brandl, J. y Gombocz, W. C. (eds.) 1989 *The Mind of Donald Davidson*, Vol. 36, Grazer Philosophische Studien, Graz.
 - Carnap, R. 1937 *The logical syntax of language*, Routledge y Kegan Paul, Londres.
1950 «Empiricism, Semantics and Ontology», *Revue Internationale de Philosophie* 4, pp. 20-40. Citado del suplemento de *Meaning and Necessity: A study in Semantics and Modal Logic*, University of Chicago Press, 1956.
 - Carpenter, A. 1998 «The Unintelligibility of Massive Error», *Disputatio* 4, pp. 25-45.
 - Cheng, Kuang-Ming 2005 «Must We Know What We Mean?», *Kriterion* 19, pp. 21-33.
 - Chomsky, N. 1975a «Quine's Empirical Assumptions», en Davidson, D. and Hintikka, J. (eds.), 1975, pp. 53-68.
1975b *Reflections on Language*, Fontana/Collins.
 - Creath, R. 2004 «Quine on the Inteligibility and Relevance of Analicity», en Gibson Jr, R. F. (ed.) 2004, pp. 47-64.
 - Davidson, D. 1967 «Truth and Meaning», *Synthese*, 17, pp. 304-323. Citado de Davidson 1984a, pp. 18-37.
1973 «Radical Interpretation», *Dialectica* 27, pp. 313-328. Citado de Davidson 1984a, pp. 125-140.
1974a «On the Very Idea of a Conceptual Scheme», *Proceedings and Addresses of the American Philosophical Association* 47, pp. 5-20. Citado de Davidson 1984a, pp. 183-198.
1974b «Belief and the Basis of Meaning» *Synthese* 27, pp. 309-323. Citado de Davidson 1984a, pp. 141-154.

-
- 1975 «Thought and Talk», in Guttenplan, S. (ed.) 1975, pp. 7-23. Citado de Davidson 1984a, pp. 155-170.
- 1976 «Reply to Foster», en Evans, G. y McDowell, J. (eds.) 1976, pp. 33-41.
- 1977 «Reality without Reference», *Dialectica* 31, pp. 247-258. Citado de Davidson 1984a, pp. 215-225.
- 1979 «The Inscrutability of Reference», *The Southwestern Journal of Philosophy* 10, pp. 7-19. Citado de Davidson 1984a, pp. 227-242.
- 1982 «Empirical Content», en Lepore, E. (ed.) 1986, pp. 320-332.
- 1983 «A Coherence Theory of Truth and Knowledge», en Malachowski, A. (ed.), pp. 120-134. (Recogido también en Davidson 2001a, pp. 137-157).
- 1984a *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford.
- 1984b «First Person Authority», *Dialectica* 38, pp. 101-111. Citado de Davidson 2001a, pp. 3-14.
- 1986 «A Nice Derangement of Epitaphs», en Lepore, E. (ed.) 1986, pp. 433-446.
- 1987 «Knowing One's Own Mind», *Proceedings and Addresses of the American Philosophical Association 1987*, pp. 441-458. Citado de Davidson 2001a, pp. 15-38.
- 1988 «The Myth of the Subjective», actas de la conferencia *Bewusstsein, Sprache und die Kunst*, editadas por M. Benedikt y R. Berger, edición de Verlag der Österreichischen Staatsdruckerei, 1988. Citado de Davidson 2001a, pp. 39-52.
- 1989a «What is Present to the Mind», en el volumen especial de *Grazer Philosophische Studien: The Mind of Donald Davidson*, Brandl, J. y Gombocz, W. (eds.), vol. 36, 1989, Amsterdam: Rodopi. Citado de Davidson 2001a, pp. 53-67.
- 1989b «The Conditions of Thought», en Brandl y Gombocz (eds.) 1989, pp. 193-200.

-
- 1990a «The Structure and Content of Truth», *Journal of Philosophy* 87, pp. 279-328.
- 1990b «Meaning, Truth, and Evidence», en Barrett, R. y Gibson, R. (eds.) 1990, pp. 68-95.
- 1990c «Epistemology Externalized», *Dialectica*, vol. 45, nº. 2-3, pp. 191-202. citado de Davidson 2001a, pp. 193-204.
- 1991 «Epistemology Externalized», *Dialectica* 45(2-3), pp. 191-202. Citado de Davidson 2001a, pp. 193-204.
- 1992a «The Second Person», *Midwest Studies in Philosophy* 17, editado por French, P., Uehling, T. E. y Wettstein, H. (Indianapolis: University of Notre Dame Press), pp. 255-267. Citado de Davidson 2001a, pp. 107-121.
- 1992b *Mente, Mundo y Acción*, Paidós/ICE-UAB.
- 1994 «Exchange between Donald Davidson and W.V. Quine following Davidson's Lecture», *Theoria* 60, 3, pp. 226-231.
- 1995a «Pursuit of the Concept of Truth», en Leonardi, P. y Santambrogio, M. (eds.) 1995, pp. 7-21.
- 1995b «Could There Be a Science of Rationality?», *International Journal of Philosophical Studies*, 3, pp. 1-16. Citado de Davidson, D. (2004), pp. 117-134.
- 1997 «Indeterminism and Antirealism», en C. B. Culp (ed.) *Realism/Antirealism and Epistemology*, pp. 109-122. Citado de Davidson 2001a, pp. 69-84.
- 1998 «The Irreducibility of the Concept of the Self», en Stamm, M. (ed.) *Philosophie in synthetischer Absicht* (Stuttgart: Klett-Cotta). Citado de Davidson 2001a, pp. 85-91.
- 1999a «Reply to John Collins», en Hahn, E. L. (ed.) 1999, pp. 529-530.
- 1999b «Reply to Richard Rorty», en Hahn, E. L. (ed.) 1999, pp. 595-600.

- 2001a *Subjective, Intersubjective, Objective*, Clarendon Press, Oxford.
- 2001b «Externalisms», en Kotatko, P. *et al.*, (eds.) 2001, pp. 1-16.
- 2001c «Comments on Karlovy Vary Papers», en Kotatko, P. *et al.*, (eds.) 2001, pp. 285-307.
- 2004 *Problems of Rationality*, Oxford Clarendon Press, Oxford.
- Davidson, D. y Hintikka, J. (eds.) 1969 *Words and Objections. Essays on the work of Work of W. V. Quine*, Dordrecht, Reidel.
 - Dennett, D. 1991 «Real Patterns», *Journal of Philosophy* 88, pp. 27-51. (Recogido en D. Dennett, *Brainchildren. Essays on Designing Minds*, MIT Press, 1998, pp. 95-120).
 - Devitt, M. 1990 «Meanings Just Ain't in the Head», en Boolos, G. (ed.) 1990, pp. 79-104.
- 2006 *Ignorance of Language*, Oxford University Press, Oxford.
- Devitt, M. y Sterelny, K. 1999 *Language and Reality. An Introduction to the Philosophy of Language*. Oxford, Basil Blackwell.
 - Dreben, B. 2004 «Quine on Quine», en Gibson Jr, R. F. (ed.) 2004, pp. 287-293.
 - Duhem, P. 1906 *La Théorie physique: son objet et sa structure*, Chevalier, París.
 - Dummett, M. 1978 *Truth and Other Enigmas*, London, Duckworth.
- 2002 «Meaning in Terms of Justification», *Topoi* 21, enero, pp. 11-19.
- 1986 «A Nice Derangement of Epitaphs: Some Comments on Davidson and Hacking», en Lepore, E. 1986, pp. 459-476.
- 1987 «Reply to John McDowell», en Taylor, B. M. (ed.), pp. 253-268.
- Eckardt, R., y von Heusinger, K., (eds.) 1999 *Meaning Change – Meaning Variation*, Actas del congreso celebrado en Constanza, febrero.

- Evans, G. 1975 «Identity and Predication», *Journal of Philosophy* 72, pp. 343-63.
- Evans, G. y McDowell, J. (eds.) 1976 *Truth and Meaning. Essays in Semantics*, Oxford University Press.
- Feyerabend, P., «Explanation, Reduction and Empiricism», *Minnesota Studies in the Philosophy of Science*, Vol. III, pp. 28-97.
- Field, H. 1973 «Theory Change and the Indeterminacy of Reference», *Journal of Philosophy* 72, pp. 462-481.
- Føllesdal, D. 1990 «Indeterminacy and Mental States», en Gibson, R., y Barrett, R (eds.), pp. 98-109.
- 1995 «In what Sense Is Language Public?», en Leonardi, P. y Santambrogio, M. (eds.) 1995, pp. 53-67.
- Føllesdal, D. y Quine, D. B. (eds.) 2008 *W. V. Quine. Confessions of a Confirmed Extensionalist and Other Essays*, Harvard University Press, Cambridge, Massachusetts, Londres.
- Foster, J. 1976 «Meaning and Truth Theory», en Evans, G. y McDowell, J. (eds.) 1976.
- Fricke, M. F. 2007 «Davidson y la Autoridad de la Primera Persona», *Diánoia*, vol. LII, n°. 58, mayo, pp. 49-76.
- Garrido, M., 1976 «El Principio de Indeterminación en la Semántica de Quine», en VVAA 1976.
- Georgalis, N. 2005 *The Primacy of the Subjective. Foundations for a Unified Theory of the Mind*, The MIT Press.

- Gibson Jr., R. F. 1982 *The Philosophy of W. V. Quine: An expository Essay*, University Presses of Florida.
- 1995 «Quine on the Naturalizing of Empistemology», en Leonardi, P. (ed.) 1995, pp. 89-103.
- (ed.) 2004 *The Cambridge Companion to Quine*, Cambridge University Press.
- Gibson Jr., R. F., y Barrett, R. (eds.) 1990 *Perspectives on Quine*, Oxford, Blackwell.
- Glock, H-J. 1996 «On Safari with Wittgenstein, Quine and Davidson», en Arrington R. L. y Glock H-J. (eds.) 1996, pp. 144-173.
- Gochet, P. (ed.) 1997 *Quine with His Replies*, *Revue Internationale de Philosophie* 51, nº 202, diciembre.
- Guttenplan, S. (ed.) 1975 *Mind and Language*, Clarendon Press, Oxford.
- Hacking, I. 1975 *Why Does Language Matter to Philosophy?*, Cambridge University Press.
- Hahn, L. E. (ed.) 1999 *The Philosophy of Donald Davidson*, The Library of Living Philosophers, vol. XXVII, Chicago, Il: Open Court.
- Hahn, L. E. y Schlipp, P. A. (eds.) 1986 *The Philosophy of W. V. Quine*, The Library of Living Philosophers, La Salle, III: Open Court.
- Hale, B. y Wright, C. (eds.) 1998 *A Companion to Philosophy of Language*, Blackwell, Oxford, Massachusetts.
- Hernández Iglesias, M. 1990 *La Semántica de Davidson*, Visor, Madrid.
- Higginbotham, J. 1995 «The Place of Natural Language», en Leonardi, P. y Santambrogio, M. (eds.) 1995, pp. 113-139.

- 1998 «On Knowing One's Own Language», en C. Wright (*et al.*) (ed.) 1998, pp. 429-441.
- Hilton, P. 2004 «Quine on Reference and Ontology», en Gibson, R. (ed.) 2004, pp. 115-150.
 - Hookway, C. 1998 *Quine*, Basil Blackwell, Oxford.
 - Jutronic, D. (ed.) 1997 *The Maribor Papers in Naturalized Semantics* (Acta de la conferencia sobre Semántica Naturalizada y su Metodología, junio de 1996), Maribor University.
 - Kirk, R. 1987 *Translation Determined*, Clarendon Press, Oxford.
- 2004 «Indeterminacy of Translation», en Gibson, R. (ed.) 2004, pp. 151-180.
- Koppelberg, D. 1995 «Skepticism about Semantic Facts», en Leonardi, P. y Santambrogio, M. (eds.) 1995, pp. 336-346.
 - Kotatko, P., Pagin, P., Segal, G. (eds.) 2001 *Interpreting Davidson*, CSLI Publications, Stanford.
 - Kripke, S. 1982 *Wittgenstein on Rules and Private Language*, Cambridge, Mass: Harvard University Press.
 - Kuhn, T. S. 1962 *The Structure of Scientific Revolutions*, University of Chicago Press, Chicago, 2ª ed.
 - Lee, O. H. (ed.) 1936 *Philosophical Essays for A. N. Whitehead*, Longmans, Nueva York.
 - Leonardi, P. y Santambrogio, M. (eds.) 1995 *On Quine*, Cambridge University Press, Cambridge.

- LePore, E., (ed.) 1986 *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*, Oxford, Basil Blackwell.
- LePore, E. y Smith, B. C. (eds.) 2006, *The Oxford Handbook of Philosophy of Language*, Clarendon Press, Oxford.
- LePore, E. y McLaughlin, B. (eds.) 1985 *Actions and Events: Perspectives on the Philosophy of Donald Davidson*, Nueva York y Oxford: Blackwell.
- López Losada, C. 2002 «La Última Noción de Significado Estimular de Quine», *Theoria*, nº. 43, Vol.17/1, pp. 113-135.
- Lowinger, A. 1941 *The Methodology of Pierre Duhem*, Columbia University Press, New York.
- Ludwig, K. A. 1994 «First Person Knowledge and Authority», en Preyer, G. *et al.* (eds.) 1994, pp. 367-398.
- Malachowski, A. 1983 *Reading Rorty*, Basil Blackwell, Oxford.
- McGee, V. 2005 «Inscrutability and its Discontents», *Nous* 39, 3, pp. 397-425.
- Moya Espí, C. 1992 *Introducción a la Filosofía de Davidson*, en Davidson 1992b, pp. 9-50.
- Mühlhölzer, F. 1995 «Science Without Reference?», *Erkenntnis* 42, pp. 203-222.
- Müller, O. 1997 *Perspektiven der Analytischen Philosophie*, (ed. Meggle, G.), tomo 17, Walter de Gruyter, Berlin-Nueva York.
- Nelson, R. J. 1978 «The Competence-Performance Distinction in Mental Philosophy», *Synthese* 39, pp. 337-381.
- Orenstein, A. 2002 *W. V. Quine*, Acumen Publishing Limited, Chesham.

- Orenstein, A. y Kotatko, P. (eds.) 2000 *Knowledge, Language and Logic. Questions for Quine*, Boston Studies in the Philosophy of Science, vol. 210, Kluwer Academic Publishers.
- Preyer G., Siebelt, F. y Ulfig, A. (eds.) 1994 *Language, Mind and Epistemology. On Donald Davidson's Philosophy*, Kluwer Academic Publishers.
- Puhl, K. 1994 «Davidson on Intentional Content and Self-Knowledge», en Preyer, G. *et al.* (eds.) 1994, pp. 339-352.
- Putnam, H. 1975 «The Refutation of Conventionalism», en *Mind, Language and Reality*, Philosophical Papers, Vol. 2, Cambridge: Cambridge University Press, pp. 153-191.
- 1981 *Reason, Truth and History*, Cambridge University Press, Cambridge.
- 1998 *Representetion and Reality*, The MIT Press, Massachusetts.
- Quine, W. V. O. 1936 «Truth by Convention», en Lee, O. H. (ed.) 1936. Citado de Quine 1976, pp. 77-106.
- 1951 «Two Dogmas of Empiricism», *Philosophical Review*, enero, 60, 1, pp. 20-43. Citado de Quine 1963, pp. 20-46.
- 1960 *Word and Object*, The M.I.T. Press, Cambridge, Massachusetts.
- 1963 *From a Logical Point of View*, W. V. Quine, Harper and Row, New York.
- 1969a *Ontological Relativity and Other Essays*, Columbia University Press, New York y Londres.
- 1969b «Reply to Chomsky», en Davidson, D. and Hintikka, J. (eds.) 1969, pp. 302-311.
- 1969c «Reply to Davidson», en Davidson, D. y Hintikka, J. (eds.) 1969, pp. 333-335.

- 1970a «On the Reasons for Indeterminacy of Translation», *Journal of Philosophy* 26, pp. 178-183.
- 1970b «Methodological Reflections on Current Linguistic Theory», *Synthese* 21, pp. 386-398. Citado de Føllesdal, D. y Quine, D. B. (eds.), 2008, pp. 215-227.
- 1974 *The Roots of Reference*, La Salle, Illinois: Open Court, 1974. Esta obra se cita de la traducción española: *Las raíces de la referencia*, Alianza Editorial, 1988.
- 1975a «The Nature of Natural Language», en Guttenplan, S. (ed.) 1975, pp. 67-81.
- 1975b «Mind and Verbal Dispositions», en Guttenplan, S. (ed.) 1975, pp. 83-95. Citado de Føllesdal, D. y Quine, D. B. (eds.) 2008, pp. 244-256
- 1976 *The Ways of Paradox and Other Essays*, Cambridge, MA: Harvard University Press.
- 1977 «Facts of the Matter», en Shahan, R. W. y Merrill, K. R. (eds.) 1977. Citado de Føllesdal, D. y Quine, D. B. (eds.) 2008, pp. 271-286.
- 1981 *Theories and Things*, Harvard University Press, Cambridge.
- 1985 «Comment on Føllesdal », en LePore, E. y McLaughlin, B. (ed.) 1985, pp.
- 1986 «Reply to Robert Nozick», en Hahn, L. E. (ed.) 1986, pp. 364-7.
- 1987 «Indeterminacy of Translation Again», *The Journal of Philosophy*, vol. LXXXIV, n° 1, enero.
- 1990 «Reply to Føllesdal », en Gibson, R., y Barrett, R (eds.), pp. 110-111.
- 1993 «In Praise of Observation Sentences», *Journal of Philosophy* 90, pp. 107-116.
- 1995a *From Stimulus to Science*, Harvard University Press, Cambridge, Mass.
- 1995b «Naturalism; Or, Living within One's Means», *Dialectica* 49, pp. 251-261.
- 1995c «Reactions», en Leonardi, P. y Santambrogio, M. (eds.) 1995, pp. 347-361.

1996 «Progress on Two Fronts», *The Journal of Philosophy*, vol. XCIII, n° 4, Abril, pp. 159-163.

1997 «Reply to Orenstein», en Gochet, P. (ed.) 1997, pp. 573-574.

1999 *Pursuit of Truth*, Harvard University Press, Cambridge, Massachusetts y Londres.

2000 «Quine's Responses», en Orenstein, A. y Kotatko, P. (eds.) 2000, pp.407-430.

- Ramberg, B. T. 1989 *Donald Davidson's Philosophy of Language. An Introduction.*, Basil Blackwell, Oxford.

2001 «What Davidson Said to the Skeptic», en Kotatko, P. *et al.* (ed.) 2001, pp. 213-235.

- Rawling, P., 2001 «Davidson's Measurement Theoretic Reduction of the Mind», en P. Kotatko (*et al.*) (ed.) 2001, pp.237-255.

- Searle, J. 1969 *Speech Acts. An Essay in the Philosophy of Language*, Cambridge University Press, Londres.

1983 *Intentionality: an Essay in the Philosophy of Mind*, New York: Cambridge University Press.

1987 «Indeterminacy, Empiricism, and the First Person», *The Journal of Philosophy*, vol. LXXXIV, n° 3, pp. 123-146.

1997 «An Interview With John Searle», University of California, Berkeley. Publicada únicamente en versión digital: www.stanford.edu/group/dualist/vol4/pdfs/searle.pdf.
Págs. 1-7.

- Segal, G. 2006 «Truth and Meaning», en LePore, E. y Smith, B. C. (eds.) 2006, pp. 189-212.

-
- Shahan, R. W. y Merrill, K. R. (eds.) 1977 *American Philosophy from Edwards to Quine*, University of Oklahoma Press.
 - Smith, Barry C. 1998 «On Knowing One's Own Language», en Wright, C., Smith, B. C., Macdonald, C. (eds.) 1998, pp. 391-429.
 - 2006 «Davidson, Interpretation and First-Person Constraints on Meaning», *International Journal of Philosophy Studies*, Vol. 14, 3, pp. 385-406.
 - Soames, S. 1999 «The Indeterminacy of Translation and the Inscrutability of Reference», *Canadian Journal of Philosophy*, vol. 29, nº. 3, pp. 321-370.
 - Soldati, G. 2002 «Knowledge of Meaning in the First Person», *Topoi* 21, pp. 21-24.
 - Stoecker, R. 1993 *Reflecting Davidson. Donald Davidson Responding to an International Forum of Philosophers*, Walter de Gruyter, Berlín y Nueva York.
 - Stroud, B. 1995 «Quine on Exile and Acquiescence», en Leonardi, P. y Santambrogio, M. (eds.) 1995, pp. 37-52.
 - Stüber, K. 1996 «Indeterminacy and the First Person Perspective», Acta del congreso: *Verdad: Lógica, representación y mundo*, Santiago de Compostela, pp. 333-341.
 - 1997 «Holism and Radical Interpretation: the Limitations of a Formal Theory of Language», *Analyomen* 2, ed. Meggle, G., DeGruyter, Berlin y Nueva York, pp. 290-298.
 - Taylor, B. M. 1987 *Michael Dummett*, Dordrecht.
 - Thöle, B. 1993 «The Explanation of First Person Authority», en Stoecker, R. (ed.) 1993, pp. 213-247.
 - Thompson, M. 1986 «Quine's theory of knowledge», en Hahn, L. E. y Schlipp, P. A. (eds.) 1986, pp. 537-563.

- Unger, P. 1984 *Philosophical Relativity*, Blackwell, Oxford.
- Viger, C., Stainton, R., M. Ezcurdia, (eds.) 2004 *New Essays in Philosophy of Language and Mind. Supplementary issue 30 of the Canadian Journal of Philosophy*.
- VVAA 1976 *Aspectos de la filosofía de Quine*. Actas del V simposio de lógica y filosofía de la ciencia (Cullera, 28 y 29 de junio de 1974), Departamento de Lógica y Filosofía de la Ciencia, Universidad de Valencia, 1976.
- Wiggins D. 1998 «Meaning and Truth Conditions: From Frege's Grand Design to Davidson's», en Hale, B. y Wright, C. (eds.) 1998, pp. 3-28.
- Wittgenstein, L. 1953 *Philosophische Untersuchungen*, en: *Werkausgabe*, Suhrkamp, vol. I, Frankfurt a. M., 1984.
- Wright, C. 1998 «The indeterminacy of translation», en Hale, B. y Wright, C. (eds.) 1998, pp. 397-426.
- Wright, C., Smith, Barry C., Macdonald, C. (eds.) 1998 *Knowing Our Own Minds*, Clarendon Press, Oxford.

