

# La independencia de la interpretación con respecto a la traducción, y el cuestionamiento del Principio de Caridad de Davidson\*

MARÍA DE LA CONCEPCIÓN CAAMAÑO ALEGRE

Universidad de Valladolid

mariac@fyl.uva.es

La cuestión de la independencia de la traducción con respecto a la traducción se aborda en diferentes contextos filosóficos. Me centraré principalmente en aquel más próximo al ámbito de la Filosofía de la Ciencia, desde el cual Thomas S. Kuhn trata de responder a las críticas que Donald Davidson dirige contra su noción de inconmensurabilidad o intraducibilidad entre lenguajes. La respuesta kuhniana consiste, precisamente, en un intento de refutación de la idea de Davidson de que la traducibilidad es una condición necesaria para toda interpretación. Se tendrá en cuenta aquí la formulación del Principio de Caridad realizada por Davidson. Pero quizá resulte esclarecedor atender a la formulación original del mismo por parte de Wilson, para poder reparar en las peculiaridades de la formulación debida a aquél. Después de ello presentaré esquemáticamente el argumento de Davidson para defender que la interpretación depende de la aplicación del Principio de Caridad y, en consecuencia, de la traducción. Luego contrastaré, también muy brevemente, las nociones de traducción e interpretación manejadas por Davidson y por Kuhn. Finalmente, me referiré a aquellas críticas al Principio de Caridad de Davidson desarrolladas a partir de la tesis de que resulta posible interpretar sin traducir.

## 1. Formulación original y formulación davidsoniana del Principio de Caridad

Las siguientes citas condensan, respectivamente, la formulación original del Principio de Caridad por parte de N. L. Wilson, y la subsiguiente formulación del Principio de Caridad ofrecida por D. Davidson:

«We select as *designatum* [of a name] that individual which will make the largest possible number of [the speaker's] statements true». («Substance without Substrata», *Review of Metaphysics*, 12, 1959).

---

\* Este trabajo se realizó durante el disfrute de una Beca de Investigación de la Diputación de A Coruña en 2005. Quiero agradecer a Klaus Mulzer sus agudas aclaraciones y útiles comentarios sobre diversos puntos abordados en el artículo. También han resultado de gran ayuda las sugerencias hechas por Ana Rosa Pérez Ransanz, a quien deseo expresarle asimismo mi agradecimiento.

«If we cannot find a way to interpret the utterances and other behaviour of a creature as revealing a set of beliefs largely consistent and true by our own standards, we have no reason to count that creature as rational, as having beliefs, or as saying anything». («Radical Interpretation», *Dialectica*, 27, 3-4, 1973).

Como se explicará a continuación, la formulación de Davidson es mucho más extrema que la de Wilson. En la formulación inicial de Wilson, el principio se formula como una regla semántica para determinar los referentes de los nombres en la lengua del hablante. La versión posterior de Davidson difiere substancialmente de la Wilson, ya que el Principio de Caridad se presenta allí no sólo como una regla semántica sino también como una condición de posibilidad para reconocer a un agente lingüístico como tal, esto es, para reconocer cualquier conducta intencional. La divergencia entre las dos formulaciones es relevante porque de la formulación original no se siguen algunas consecuencias derivadas de la versión de Davidson, como la necesidad de acuerdo entre la mayoría de las creencias del intérprete y del agente interpretado, la dependencia de la interpretación con respecto a la traducción, o la necesidad de que la mayoría de nuestras creencias sean verdaderas. Si para reconocer a una criatura como un agente intencional o lingüístico ha de ser ya posible atribuirle una interpretación a su lenguaje, entonces el único modo de que el intérprete lleve a cabo tal atribución es que suponga que la evidencia que le sirve para interpretar su propio lenguaje coincide con la evidencia que sirve al agente intencional para interpretar el suyo. Por ello, el punto de vista de Davidson implica que el intérprete maximice no sólo la verdad de las emisiones de los hablantes sino también el acuerdo entre las creencias de los hablantes y las creencias del intérprete. Una última diferencia entre las dos versiones del Principio de Caridad que estamos viendo rebasa lo explicitado en las formulaciones del mismo, pero es de gran importancia y fácilmente detectable a partir de las explicaciones dadas por el propio Davidson. La diferencia es que la normatividad del Principio de Caridad de Davidson se justifica apelando a la estructura misma de los estados mentales, y, por tanto, este principio establece normas que los intérpretes, por su propia naturaleza intencional, están obligados a seguir, mientras que la normatividad del Principio de Caridad de Wilson se justifica apelando a la estructura de las relaciones semánticas que el intérprete está investigando, en dicho principio se establecen normas que el intérprete no está obligado a seguir, aunque seguirlas facilite la interpretación.

## **2. Presentación esquemática del argumento de Davidson en favor de la dependencia de la interpretación con respecto a la traducción**

Veamos ahora el tipo de argumentación que Davidson desarrolla para defender que toda interpretación exige una traducción al requerir la aplicación del Principio de Caridad. Las cinco premisas se corresponden con cinco tesis de Davidson. La dependencia entre la segunda y cuarta premisa, afirmada en el marco de las tesis de Davidson, es lo que determina la dependencia de la interpretación con respecto a la traducción. El argumento se puede esquematizar del siguiente modo:

P1. La única evidencia disponible para interpretar es la observación de la conducta que el agente a interpretar manifiesta en circunstancias públicamente observables

P2. La única evidencia disponible para interpretar es insuficiente, la interpretación, por ello, requiere la atribución de creencias a los agentes a interpretar

P3. La creencia y el significado son interdependientes, y carecemos de acceso a las creencias de los agentes a interpretar

P4. La atribución de creencias requiere la aplicación del Principio de Caridad, que exige la maximización del acuerdo entre las creencias del intérprete y las creencias de los agentes a interpretar

P5. La maximización del acuerdo entre las creencias requiere que la mayoría de éstas coincidan en su contenido proposicional, y ello requiere que sean expresables en un mismo lenguaje

┆ La interpretación requiere la traducción

### 3. Discrepancias entre Quine, Davidson y Kuhn en su caracterización de las nociones de traducción e interpretación

Con el fin de comprender debidamente las críticas que Kuhn dirige contra el Principio de Caridad defendido por Davidson, es necesario mostrar las divergencias teóricas entre ambos autores en sus respectivas caracterizaciones de las nociones de traducción e interpretación. Sin embargo, conviene recordar, primero, aquellos rasgos estructurales básicos de la traducción y de la interpretación sobre los que hay acuerdo.

Se entiende que una traducción consiste en una correlación entre expresiones de dos lenguas distintas realizada en virtud de la equivalencia semántica entre las mismas, especificándose, en el caso de las expresiones ambiguas, los contextos en que cada una de ellas ha de correlacionarse con una u otra expresión de la otra lengua. La interpretación, en cambio, se considera como una relación entre expresiones de una lengua (o emisiones lingüísticas) y contenidos semánticos expresados o transmitidos por ellas. Tanto la traducción como la interpretación se expresarán en un metalenguaje en el cual se establecerá, por lo que concierne a la primera, una correlación entre expresiones semánticamente equivalentes, y, por lo que respecta a la segunda, una relación entre expresiones (o emisiones lingüísticas) y contenidos semánticos que ellas transmiten. Existen, igualmente, otros rasgos de la traducción y de la interpretación, en concreto rasgos de carácter epistemológico, conjuntamente reconocidos por Quine, Davidson y Kuhn. Se trata del tipo de conocimiento necesario para iniciar las labores de traducción e interpretación. Para lo primero se requiere que la persona que pretenda acometer una traducción sea bilingüe, esto es, que domine las dos lenguas cuyas expresiones semánticamente equivalentes pretenda correlacionar. Para lo segundo, sin embargo, no es necesario que la persona que pretenda llevar a cabo una interpretación sea bilingüe, sólo es preciso que sea capaz de llegar a entender el significado o a captar el contenido semántico de las expresiones de dicho idioma. La interpretación sería, por tanto, algo de lo que depende la traducción, sin seguirse, en cambio, de lo dicho hasta aquí, que la traducción sea a su vez algo de lo que depende la interpretación.

A juicio de Kuhn, la confusión ordinaria entre estas dos nociones a la que él alude se debería, precisamente, al hecho de que toda traducción contiene un componente interpretativo, y de que siendo habitualmente local la intraducibilidad entre dos lenguas distintas, las personas familiarizadas con ambas pueden no percatarse de que en algunos casos están interpretando y no traduciendo.

*El enfoque extensional de Davidson*, al igual que el de Quine, en el que se inspira, responde a un intento de prescindir del recurso a entidades abstractas (como proposiciones o universales) en la caracterización del significado. Prescindiendo de tal recurso, sendos autores dan cuenta de los

contenidos semánticos y de la equivalencia semántica, involucrados en las prácticas de interpretación y traducción, apelando principalmente a hechos de la conducta en circunstancias públicamente observables. En el caso de Quine, esto conduce a una indeterminación del significado y de la traducción, y en el de Davidson, a una apelación al Principio de Caridad, que permitiría determinar el significado y la traducción.

Davidson (a diferencia de Quine) mantiene que para interpretar es necesario poseer alguna teoría de la interpretación. La interpretación requiere el reconocimiento común del entorno y la confirmación de una teoría de la verdad (o las T-sentences) para dicho lenguaje. Ello se realiza a partir de la evidencia que proporciona la conducta observable del hablante indígena manteniendo ciertas emisiones como verdaderas en un cierto momento bajo circunstancias públicamente observables. Discrepando de Quine, Davidson considera que el tipo de estímulo constitutivo del significado no es el proximal (interno al sujeto) sino el distal (externo al mismo). De modo que, a su juicio, el conocimiento que adquiere un sujeto lingüístico no está sólo causado por los estímulos distales (como sostiene Quine) sino también en parte especificado por tales estímulos.<sup>1</sup> Sin embargo, la evidencia disponible no es concluyente, ya que remite necesariamente a una combinación de aspectos relativos al significado y de aspectos relativos a las creencias del hablante. El significado y la creencia guardan una relación de interdependencia manifiesta en el hecho de que no hay forma de descubrir el significado con independencia de la creencia, ni forma de descubrir la creencia con independencia del significado. El acceso a las condiciones de verdad de un enunciado se alcanza mediante el acceso a sus condiciones de asertabilidad (o condiciones para ser considerado como verdadero) por parte de los hablantes del lenguaje desconocido. La conexión entre asertar y creer genera otra entre atribuir significado a emisiones y atribuir creencias (y deseos) a los emisores de tales emisiones. De acuerdo con Davidson, esto constituye un problema, ya que no tenemos acceso a las creencias de los hablantes de otra lengua y la interpretación sólo es posible si las actitudes proposicionales que les atribuimos están disponibles públicamente. Para tratar de salvar esta dificultad metodológica, Davidson apela, por razones pues metodológicas, al Principio de Caridad, según el cual, ha de maximizarse el acuerdo entre las creencias del intérprete y las creencias del hablante cuyas emisiones se interpretan. La teoría de la verdad necesaria para interpretar presupone pues la elaboración de una teoría de las creencias.

Téngase en cuenta que la interpretación se caracteriza a partir de una noción primitiva o no definida de verdad, considerándose ésta translingüísticamente determinable, al igual que los estímulos distales, pues de lo contrario se carecería de toda evidencia para la interpretación. Así entendida, la interpretación requiere presuponer que se comparten los estándares de verdad, pues, de lo contrario, la atribución de creencias sólo podría realizarse aplicando estándares de verdad patentemente inadecuados y considerando irracionales a los agentes cuyas emisiones tratan de interpretarse. Por otra parte, al Principio de Caridad subyace la idea de que determinada conducta sólo puede reconocerse como lingüística y como racional en la medida en que se ajuste a nuestras propias creencias. Adviértase que del recurso davidsoniano a una noción de verdad translingüística y al Principio de Caridad se siguen, respectivamente, dos llamativas consecuencias: que la mayoría de nuestras creencias son verdaderas y que nuestras actitudes proposicionales son constitutivamente racionales.

---

1 Cfr. Rawling, P., «Radical Interpretation», *Donald Davidson*, ed. Kirk Ludwig, Cambridge University Press, Cambridge, 2003, 85-112, 98-99.

Ha de quedar claro, por último, que, para Davidson, la equivalencia semántica entre dos expresiones consiste en que ambas compartan sus mismas condiciones de verdad, basándose la interpretación en el establecimiento de equivalencias semánticas.

*Kuhn asume un punto de vista intensional*, conforme al cual la extensión de las expresiones de un lenguaje se determina necesariamente recurriendo a entidades intensionales como los conceptos, que son presentados como entidades abstractas con las que se categorizan los distintos tipos de entidades que configuran el mundo. Para Kuhn los conceptos se aprenden y se aplican de manera holista, en interrelación con aquellos otros necesarios en la identificación de los referentes comúnmente aceptados de los términos que transmiten dichos conceptos.<sup>2</sup> Los conceptos poseen pues, para Kuhn, un carácter unitario, es decir, no resulta posible mantener la identidad del concepto si se alteran drásticamente los vínculos semánticos del término que transmite dicho concepto. De acuerdo con Kuhn, la equivalencia semántica entre dos expresiones se da siempre que tengan la misma extensión y que transmitan la misma categoría, definida dentro de estructuras léxicas homólogas. La interpretación y la traducción serían determinables porque no sólo disponemos o podemos llegar a disponer de los mismos estímulos y conducta observable, sino también de los mismos patrones de reconocimiento de similitudes y de tipos de entidades, en otras palabras, de las mismas categorías para estructurar el mundo.

#### 4. Principales críticas al Principio de Caridad

Paso pues a comentar las principales críticas al Principio de Caridad. Se han planteado tres tipos de objeciones a la utilización del Principio de Caridad por parte de Davidson,<sup>3</sup> todos ellos notablemente convergentes con las tres líneas de crítica que Kuhn desarrolla contra su noción de interpretación. Estos tres focos de controversia se enuncian a continuación:

- 1º el carácter *a priori* del Principio de Caridad, carente de una justificación empírica;<sup>4</sup>
- 2º la alta probabilidad de que el acuerdo necesario para interpretar sea mucho menor del que Davidson considera necesario;<sup>5</sup>
- 3º la existencia de diferentes patrones de conducta racionales susceptibles de evolución a lo largo de la historia, e incluso de cierto margen de irracionalidad obedeciendo, también, ciertos patrones susceptibles de evolución.<sup>6</sup>

Los autores que siguen la primera línea de crítica (Kuhn, McGinn, Hacking) indican que, en general, los intérpretes no hacen coincidir la mayoría de sus creencias con otros sujetos integrados en un entorno muy remoto. Es, por tanto, empíricamente más adecuado relativizar la interpretación a la evidencia disponible para los sujetos a interpretar, de manera que los intérpretes atribuyan a los sujetos creencias que los intérpretes tendrían dada la evidencia del sujeto a interpretar. Este modo de interpretar no presupone la maximización del acuerdo entre creencias sino que las creencias respondan, en cada caso, a la evidencia disponible por un sujeto, y que, dada la misma evidencia, dos sujetos tengan las misma creencias.

2 Cfr. «Commensurability, Comparability, Communicability», *cit.*, p. 680.

3 Una esclarecedora revisión crítica del Principio de Caridad se encuentra en Lewis, D., «Radical Interpretation» (1974), *Philosophical Papers*, vol. I, Oxford University Press, Oxford, 1983.

4 B. Ramberg desarrolla este punto en *Donald Davidson's Philosophy of Language*, capítulos 6, 7 y 9.

5 Este problema es puesto de relieve por S. Stich en *The Fragmentation of Reason*, capítulo 1.

6 Sobre esta cuestión véase McGinn, C. «Charity, Interpretation and Belief», *Journal of Philosophy* 74, 1977.

Otras evidencias empíricas contrarias al Principio de Caridad serían la posibilidad del aprendizaje de una primera lengua y la necesidad que surge frecuentemente de emplear glosas para dar una interpretación de expresiones para las que no se ha logrado hallar una traducción.

El estudio empírico de la interpretación revela que el acuerdo necesario para interpretar es considerablemente menor de lo que Davidson afirma. La idea ya mencionada de relativizar la interpretación a la evidencia disponible en cada caso se recoge en el denominado Principio de Humanidad, formulado por Richard Grandy,<sup>7</sup> donde se establece que los agentes a interpretar razonan desde su perspectiva como el intérprete razona desde la suya. Este principio, a diferencia del Principio de Caridad, es sensible a las asimetrías en cuanto a la información de que disponen los distintos agentes lingüísticos. Otro principio similar al de Humanidad, el Principio de Plausibilidad Psicológica de la Interpretación, reivindicado por Quine, apela igualmente a la necesidad de tener en cuenta el entorno cultural de los agentes a interpretar, lo cual a menudo supone tener que atribuirles creencias falsas. Estos principios alternativos sugieren que el acuerdo necesario para interpretar puede limitarse a una coincidencia suficiente en nuestras habilidades perceptivas y cognitivas, lo que nos capacitaría para aprender cualquier lenguaje humano.

Finalmente, en la medida en que la racionalidad depende del conocimiento disponible, si el conocimiento disponible varía de una comunidad a otra y a lo largo de la historia, entonces la racionalidad también variará de una comunidad a otra y a lo largo de la historia. Por otra parte, la irracionalidad constituye una vertiente importante de la conducta humana, ya que supone un modo de economizar la conducta racional, el esfuerzo que ella requiere. Además de cumplir otras funciones más lúdicas o más psicológicas, como la de satisfacer la necesidad de marcar la propia individualidad. También la conducta irracional se verá condicionada por el entorno cultural y por el momento histórico.

## 5. Argumentos en favor de la independencia de la interpretación con respecto a la traducción

A partir de las tres líneas de crítica contra el Principio de Caridad esbozadas más arriba pueden desarrollarse por lo menos dos argumentos en favor de la independencia de la interpretación con respecto a la traducción. Tal convergencia, además de evidenciar la variedad de enfoques desde los que es posible determinar dicha independencia, refuerza la evidencia en favor de la tesis correspondiente. Por ello, los argumentos que se bosquejarán a continuación constituyen un contundente foco de crítica contra el Principio de Caridad de Davidson.

Uno de los argumentos a los que me referiré es más directo que el otro, en el que sentido de que la conclusión se sigue directamente del mero reconocimiento de un determinado hecho, mientras que, en el otro argumento, la misma conclusión se sigue de tal reconocimiento junto con una determinada asunción en cuanto a su relevancia. De acuerdo con el primer argumento, es una cuestión de hecho que la interpretación se da incluso cuando no hay ninguna traducción disponible. Entre los casos de este tipo, cabe mencionar el aprendizaje de una lengua materna, así como el adquisición de un lenguaje o fragmento de lenguaje estructuralmente incompatible con aquel usado por el intérprete. Dado que las capacidades cognitivas humanas se muestran suficientes para posibilitar la interpretación de cualquier lenguaje humano incluso careciéndose de traducción alguna para dicho lenguaje, la interpretación está garantizada por tales capacidades y es, por tanto, independiente de la traducción.

7 Grandy, R., «Reference, Meaning, and Belief», *Journal of Philosophy*, 70, 1973, pp. 442-443.

Según lo establecido en el segundo argumento, la racionalidad tiene un carácter plural derivado del hecho de que depende de un conocimiento y unos objetivos variables. Teniendo en cuenta que la conducta lingüística observable es la principal fuente de información necesaria para interpretar, y que esta conducta sólo es comprensible asumiendo algún esquema racional subyacente a la misma, puede inferirse que habrá distintos esquemas racionales subyaciendo a conductas lingüísticas radicalmente distintas. Tomando igualmente en consideración que, de hecho, los seres humanos son capaces de identificar diferentes patrones de conducta racional, podemos concluir que la comprensión de la conducta lingüística está garantizada incluso cuando no existe un esquema racional compartido. Puesto que interpretar requiere la aplicación de un esquema racional, y traducir requiere la aplicación un esquema racional compartido, la interpretación es posible incluso cuando la traducción no lo es. De ello se sigue que la interpretación es independiente de la traducción.

## Conclusión

Las consideraciones precedentes apoyan las siguientes tesis:

- El Principio de Caridad de Davidson no es empíricamente adecuado, como lo demuestran los diversos tipos de evidencia empírica en su contra: la posibilidad de aprender una primera lengua, la existencia de lenguajes interpretables mutuamente intraducibles, la variabilidad de los esquemas racionales operando en diferentes comunidades y períodos históricos.
- La posesión de capacidades cognitivas compartidas, tales como las habilidades perceptivas e inferenciales, es lo que garantiza a los seres humanos la posibilidad de aprender e interpretar cualquier lenguaje humano, independientemente de que exista o no una traducción para él.
- La racionalidad que rige la conducta humana, y por consiguiente determinante de la fuente primaria de información sobre cómo ha de interpretarse un lenguaje, varía a lo largo de la historia y de una cultura a otra.

El Principio de Caridad propuesto por Davidson comporta, por tanto, consecuencias difícilmente admisibles, ya que parece estar en conflicto con evidencia empírica básica relativa a la capacidad de interpretación en los humanos.

## Referencias bibliográficas

- Davidson, D., «Radical Interpretation», *Dialectica*, vol. 27, n° 3-4, 1973, pp. 313-328.
- «On the Very Idea of a Conceptual Scheme», *Inquires into Truth and Interpretation*, Oxford U. P., Oxford, 1984, pp. 183-198.
- Grandy, R., «Reference, Meaning, and Belief», *Journal of Philosophy*, 70, 1973, pp. 442-443.
- Hacking, I., *Why does Language Matter to Philosophy*, Cambridge University Press, Cambridge, 1975.
- Kitcher, P., «Theories, Theorists and Theoretical Change», *The Philosophical Review*, 87, 1978, pp. 519-547.
- Kuhn, T. S., «Commensurability, Comparability, Communicability», *Philosophy of Science Association*, 1983, pp. 669-688.
- Lewis, D., «Radical Interpretation» (1974), *Philosophical Papers*, vol. I, Oxford University Press, Oxford, 1983, pp. 108-118.
- McGinn, C. «Charity, Interpretation and Belief», *Journal of Philosophy* 74, 1977, pp. 521-535.
- Pérez Ransanz, A. R., *Kuhn y el cambio científico*, Fondo de Cultura Económica, México, 1999.



- Putnam, H., *Reason, Truth and History*, Cambridge University Press, Cambridge, 1982.
- Quine, W. V. O., *Word and Object*, MIT Press, Cambridge, Mass., 1960.
- Ramberg, B., *Donald Davidson's Philosophy of Language*, Basil Blackwell, Oxford, 1989.
- Rawling, P., «Radical Interpretation», *Donald Davidson*, ed. Kirk Ludwig, Cambridge University Press, Cambridge, 2003, pp. 85-112.
- Simon, H., «Computational Theories of Cognition», William O'Donohue and Richard F. Kitchener (eds.), *The Philosophy of Psychology*, SAGE, London, 1996.
- Stich, S., *The Fragmentation of Reason*, the MIT Press, Cambridge, 1990.
- Wilson, N. L., «Substances without Substrata», *Review of Metaphysics*, 12, 1959, pp. 521-539.