



La determinación de la referencia de los términos para artefactos

Laura Ortega Cano

ADVERTIMENT. La consulta d'aquesta tesi queda condicionada a l'acceptació de les següents condicions d'ús: La difusió d'aquesta tesi per mitjà del servei TDX (www.tdx.cat) i a través del Dipòsit Digital de la UB (diposit.ub.edu) ha estat autoritzada pels titulars dels drets de propietat intel·lectual únicament per a usos privats emmarcats en activitats d'investigació i docència. No s'autoritza la seva reproducció amb finalitats de lucre ni la seva difusió i posada a disposició des d'un lloc aliè al servei TDX ni al Dipòsit Digital de la UB. No s'autoritza la presentació del seu contingut en una finestra o marc aliè a TDX o al Dipòsit Digital de la UB (framing). Aquesta reserva de drets afecta tant al resum de presentació de la tesi com als seus continguts. En la utilització o cita de parts de la tesi és obligat indicar el nom de la persona autora.

ADVERTENCIA. La consulta de esta tesis queda condicionada a la aceptación de las siguientes condiciones de uso: La difusión de esta tesis por medio del servicio TDR (www.tdx.cat) y a través del Repositorio Digital de la UB (diposit.ub.edu) ha sido autorizada por los titulares de los derechos de propiedad intelectual únicamente para usos privados enmarcados en actividades de investigación y docencia. No se autoriza su reproducción con finalidades de lucro ni su difusión y puesta a disposición desde un sitio ajeno al servicio TDR o al Repositorio Digital de la UB. No se autoriza la presentación de su contenido en una ventana o marco ajeno a TDR o al Repositorio Digital de la UB (framing). Esta reserva de derechos afecta tanto al resumen de presentación de la tesis como a sus contenidos. En la utilización o cita de partes de la tesis es obligado indicar el nombre de la persona autora.

WARNING. On having consulted this thesis you're accepting the following use conditions: Spreading this thesis by the TDX (www.tdx.cat) service and by the UB Digital Repository (diposit.ub.edu) has been authorized by the titular of the intellectual property rights only for private uses placed in investigation and teaching activities. Reproduction with lucrative aims is not authorized nor its spreading and availability from a site foreign to the TDX service or to the UB Digital Repository. Introducing its content in a window or frame foreign to the TDX service or to the UB Digital Repository is not authorized (framing). Those rights affect to the presentation summary of the thesis as well as to its contents. In the using or citation of parts of the thesis it's obliged to indicate the name of the author.

LA DETERMINACIÓN DE LA REFERENCIA DE LOS TÉRMINOS PARA ARTEFACTOS

Doctorando: Laura Ortega Cano

Directora de la tesis: Genoveva Martí

Tutora de la tesis: Genoveva Martí

Universitat de Barcelona, Facultat de Filosofia

Departament de Lògica, Història i Filosofia de la Ciència

Programa de doctorado: Ciència Cognitiva i Llenguatge (EEES)

ÍNDICE

Agradecimientos	v
Abstract (resumen en inglés).....	vii
Introducción	1
Capítulo 1. Los nombres propios y los términos de clase natural	9
1.1. El internismo clásico para los nombres propios	10
1.1.1. La propuesta internista de Frege.....	10
1.1.2 La propuesta descriptivista de Russell.....	13
1.2. El nuevo descriptivismo modificado para los nombres propios: la teoría cúmulo y el descriptivismo débil o de fijación de la referencia	17
1.2.1. La teoría cúmulo.....	17
1.2.2. El descriptivismo débil o de fijación de la referencia.....	21
1.3. Extensión del internismo a los términos generales: el caso de los términos de clase natural	22
1.3.1. El internismo de Locke para los términos de clase natural.....	22
1.3.2. El internismo de Mill para los términos generales.....	26
1.3.3. Las teorías descriptivistas modificadas aplicadas al caso de los términos de clase natural.....	29
1.4. La crítica de la nueva teoría de la referencia al internismo descriptivista clásico y a la teoría cúmulo para los nombres propios	31
1.4.1. La crítica kripkeana al descriptivismo tradicional.....	31

1.4.2. La crítica kripkeana a la teoría cúmulo y al descriptivismo de fijación de la referencia.....	34
1.5. La crítica externista al internismo descriptivista clásico y a la teoría cúmulo para los términos de clase natural.....	37
1.5.1 La crítica de Kripke.....	37
1.5.2. El argumento de Putnam de la Tierra Gemela.....	45
1.5.3. La propuesta positiva del externismo.....	49
Capítulo 2. Las clases artificiales y los términos de clase artificial.....	57
2.1. Diferencias y similitudes entre las clases naturales y las clases artificiales.....	57
2.1.1. Diferencias entre las clases naturales y las clases artificiales.....	57
2.1.2. Similitudes entre las clases naturales y las clases artificiales.....	60
2.2. La esencia de los artefactos.....	62
2.2.1. El principio de realizabilidad múltiple y la esencia nominal de los artefactos.....	62
2.2.2. Crítica a la idea de la esencia nominal de los artefactos.....	67
2.2.3. La defensa de las clases artificiales como clases no nominales.....	70
2.2.4. Las condiciones necesarias y suficientes de pertenencia a las clases artificiales.....	73
2.2.5. Alternativa a las condiciones necesarias y suficientes.....	79
2.3. La relación entre la metafísica y la semántica.....	83
2.3.1. La relación entre la metafísica de los artefactos y la semántica de los términos de clase artificial.....	83
2.3.2. El papel de la metafísica en la determinación de la referencia de los términos de clase natural.....	90
2.4. Una primera propuesta semántica para los términos de clase artificial.....	92
2.4.1. El argumento putnamiano de los lápices-organismos.....	92
2.4.2. Adecuación del modelo putnamiano a los términos para artefactos.....	97

Capítulo 3. Las réplicas descriptivistas al modelo putnamiano para los términos de clase artificial y objeciones a estas réplicas.....	101
3.1. La crítica del descriptivismo <i>puro</i> a la propuesta de Putnam.....	102
3.1.1. Consideraciones metafísicas.....	102
3.1.2. Un criterio para distinguir entre términos de clase natural y términos de clase artificial.....	105
3.2. Objeciones al descriptivismo <i>individualista</i> de Schwartz.....	106
3.2.1. Crítica a la concepción metafísica de Schwartz.....	106
3.2.2. Crítica a la analiticidad de los términos de clase artificial.....	110
3.2.3 Crítica al argumento de las generalizaciones estables.....	112
3.3. La teoría híbrida de la referencia de Thomasson.....	114
3.3.1. Descriptivismo <i>sui generis</i>	114
3.3.2. La propuesta metafísica.....	115
3.3.3. Derivación de conclusiones epistemológicas.....	119
3.3.4. Conclusiones semánticas: teoría híbrida de la referencia.....	122
3.4. Objeciones a la teoría thomassoniana.....	129
3.4.1 Crítica a la concepción metafísica y epistemológica de Thomasson.....	129
3.4.2. Crítica a la teoría híbrida de la referencia.....	133
Capítulo 4. Argumentos externistas de ignorancia y error contra el descriptivismo.....	140
4.1. ¿Un argumento externista <i>de Tierra Gemela</i> con un término de clase artificial?.....	141
4.1.1. El argumento de los lápices-organismos y las críticas descriptivistas.....	141
4.1.2. Ulteriores dificultades del argumento de los lápices-organismos.....	143
4.2. Dos argumentos externistas de ignorancia y error sobre la función de un artefacto.....	146

4.2.1. El argumento externista de Kornblith basado en la ignorancia.....	146
4.2.2. El argumento externista de Nelson basado en el error.....	148
4.3. Objeciones descriptivistas a los argumentos externistas de ignorancia y error y réplicas a esas objeciones.....	150
4.3.1. Objeciones de Schwartz.....	150
4.3.2. Objeciones de Thomasson.....	153
4.4. Análisis de los argumentos externistas de Kornblith y Nelson.....	155
4.4.1. Réplica a las críticas de Schwartz y Thomasson.....	155
4.4.2. El alcance de los argumentos externistas de Kornblith y Nelson en la disputa con el descriptivismo.....	163
4.5. La posibilidad de construir argumentos de ignorancia y error para toda comunidad lingüística.....	166
4.5.1. Argumentos universales de ignorancia y error.....	166
4.5.2. La posible objeción a los argumentos externistas <i>universales</i> de la falta de referencia.....	171
4.5.3. La idiosincrasia de los argumentos universales propuestos.....	175
Conclusiones.....	182
Conclusions (conclusiones en inglés).....	188
Bibliografía.....	193

Agradecimientos

En primer lugar, quiero agradecer a mi directora de tesis, la doctora Genoveva Martí, su dedicación, paciencia y apoyo. Su entusiasmo y sus conocimientos han sido de una gran ayuda para mí.

También quiero expresar mi agradecimiento al grupo Logos, grupo de investigación al que he estado vinculada estos años, por las actividades organizadas y por crear el entorno ideal para la investigación, sobre todo por sus brillantes miembros.

También agradezco al departamento de Lògica, Història i Filosofia de la Ciència de la facultad de Filosofia la oportunidad de haber podido trabajar este tiempo en un departamento que, merecidamente, tiene una excelente reputación.

Quiero agradecer igualmente la financiación recibida durante estos años y sin la que no habría sido posible este trabajo: por un lado, la financiación de la beca FI (2009FI_B00342) de la Generalitat de Catalunya, y, por otro, la financiación recibida de la beca FPU (AP2008-00909) del Ministerio de Educación, Cultura y Deporte de España.

Asimismo, debo mencionar los dos proyectos de investigación en el marco de los cuales he desarrollado mi tesis: el proyecto “The Determination of Reference” (FFI2008-04263), financiado por el Ministerio de Ciencia e Innovación de España, y el proyecto “Reference, self-reference and empirical data” (FFI2011-25626), financiado por el Ministerio de Economía y Competitividad de España.

Una mención especial merece el profesor Diego Marconi, profesor de acogida en mi estancia en la Universidad de Torino. Pese a su valiosísima ayuda, me temo que no he sabido sacar todo el provecho a los conocimientos que ha compartido conmigo.

Agradezco también a los miembros del tribunal, constituido por el doctor Massimiliano Carrara, el doctor Jesús Vega y el doctor Diego Marconi, su deferencia a la hora de aceptar evaluar esta tesis.

Y, por último, quiero dar las gracias a mi familia. Especialmente, a dos personas sin las cuales este proyecto y muchos otros no habrían sido posibles: mi hermana Ester y mi marido Pablo.

Abstract

This dissertation is about the determination of reference of terms for artifacts. In order to study the question of how reference of those terms is determined, I follow the Putnamian suggestion that terms for artifacts respond to an externalist semantics and explore this option.

The dissertation has four chapters. In the first chapter, **chapter 1**, I analyze the question of the determination of reference of proper names and natural kind terms. This analysis is a starting point for the central issue of this dissertation, the determination of reference of artifactual kind terms. In the historical debate about the semantics of proper names and natural kind terms, internalism has been the first theory in providing an explanation. The internalist idea is that reference of a proper name is determined by some cognitively available entity (a description or another entity) that competent speakers attach to the term in question. And the same for the case of natural kind terms: the determination of reference of terms like ‘water’ is given by some cognitively accessible entity to the speakers.

However, despite its initial plausibility, internalism is not free from problems. The new theory of reference presents its objections against different versions of internalism, and, in particular, against descriptivism, a form of internalism. The main criticism from externalism is that, if internalism maintains that the reference of our terms is determined by something cognitively accessible (a description, say), then it would not be possible that speakers can be ignorant or mistaken about what identifies the referents of our terms (the descriptions that singularizes the referents of our terms could not be mistaken or insufficient to pick up the referents). But this has counterexamples that show that we can successfully use, for instance, a proper name without a description associated to the term that singularizes the referent of the term. Or we can be mistaken about the property of the bearer of the name that is expressed by the description we attach to the proper name. So, descriptivism does not explain how reference of proper names is determined. And the same goes for natural kind terms. In contrast, externalism can accommodate the ignorance or error of speakers, defending that the reference both of proper names and natural kind terms is determined by something objective in the world that could turn out not to be cognitively accessible.

In **chapter 2**, I focus on the issue of artifactual kinds and artifactual kind terms. Artifactual kinds are conceived in contrast with natural kinds. Certainly, artifactual kinds present differences with respect to natural kinds, but also some similarities. Philosophers have proposed different theories concerning the essence of artifacts, but a significant part of the discussion has been about whether artifacts have a *real* or a *nominal* essence. I think that artifactual kinds are not nominal kinds, and that artifactual kind terms are not nominal kind terms, at least in the traditional sense of *nominal kind* and *nominal kind term*. That is, I think that terms like ‘pencil’ are not terms like ‘bachelor’. From my point of view, the difficulty (or impossibility) in providing necessary and sufficient conditions for belonging to an artifactual kind encourages an alternative conception for artifactual kinds, inspired by philosophers as Wittgenstein or Putnam, according to which an artifact x belongs to an artifactual kind X if and only if x is similar (in the *relevant* respects) to paradigmatic instances of X . This is a parallel picture to that of the proposed by Putnam for natural kinds.

But, in spite of the metaphysical similarities between artifactual kinds and natural kinds, there are some intuitive metaphysical differences between them. And it has been said that these metaphysical differences justify a different semantics for artifactual and natural kind terms, respectively. I think that most of the metaphysical considerations that can be made with respect both natural and artifactual kinds are independent from the semantic issues concerning the determination of reference. The fact is that different metaphysical conceptions about natural and artifactual kinds are compatible with different semantic proposals for the corresponding terms. So, I think that reasons for arguing for a different semantics for artifactual kind terms with respect to natural kind terms are not to be found in the metaphysical realm. Semantics for artefactual kind terms does not depend on metaphysics of artefactual kinds, in the same way that semantics for natural kind terms does not depend on metaphysics of natural kinds.

I think that the question about whether an externalist or an internalist semantics is the appropriate semantics for artifactual kind terms depends, as in the case of natural kind terms, on the possibility of construing arguments of ignorance or error concerning the *relevant* nature of a certain artifact. If this is possible, externalism for artifactual kind terms is, in principle, an admissible explanation for the determination of reference for artifactual kind terms. If such arguments cannot be construed, then internalism seems

more plausible. Putnam is one of the first philosophers in presenting a semantic proposal for terms for artifacts, and he gives an argument of error concerning what he thinks that is the *relevant* nature of a certain artifact in order to defend an externalist semantics for those terms.

In **chapter 3**, I expose the criticisms that descriptivists have made to the Putnamian externalist argument and their positive theories for terms for artifacts. Basically, descriptivists argue against externalism in the basis of metaphysical considerations regarding the nature of artifacts, and, in order to give a positive account of how reference of terms for artifacts is determined, they pretend to derive semantic conclusions from metaphysical considerations about artifacts.

In the literature, there are two basic descriptivist options for artifactual kind terms: Schwartz's *individualistic* descriptivism and Thomasson's *communitarian* or *social* descriptivism. In the former case, we have a *pure* descriptivist theory according to which the reference of an artifactual kind term is determined by a singularizing description of the corresponding term that every competent speaker, *qua* competent speaker, associates with the term. In the case of Thomasson's theory, we have not just a communitarian theory, in which deference to experts is allowed, but also a *hybrid* theory of reference of artifactual terms. According to this *hybrid* theory, in contrast with the *pure* descriptivist theory of Schwartz, the reference of our terms for artifacts is determined by two factors: a descriptivist factor (the conception of the creator of the artifact in question about its *relevant* nature) and an *external* or causal factor (properties of the artifact that can be ignored by the creator). Although she defends that her theory incorporates a descriptivist and an externalist element, I consider that her proposal is basically a descriptivist one because I think the descriptivist element has finally all the weight in the determination of reference.

I present the main objections that can be made to both Schwartz and Thomasson's proposals. I consider that Schwartz's theory does not account for the semantics of artifactual kind terms. In sum, I think that it is not the case that we associate with every artifactual kind term an analytic definition, known *a priori* and necessarily linked to the term. Moreover, appealing to linguistic intuitions in hypothetical situations in order to defend descriptivism, as Schwartz does, is not so conclusive as he states. In the case of

Thomasson's theory, I think that her hybrid theory of reference is not free from problems and I find especially hard to understand how it is supposed to work. On the other hand, the main motivation for such an hybrid theory, the so called '*qua* problem' can be solved appealing to a causal externalist theory. So, even if we accept the *qua* problem, we are not forced to adopt an hybrid theory of reference.

Finally, in the last chapter, **chapter 4**, I consider the viability of externalist arguments of ignorance or error regarding artifactual kind terms. As I said, Putnam presented such an argument, the arguments of pencils-organisms, but it has not been so convincing as the Twin Earth thought experiment with 'water'. The reason is that, according both externalist and internalist philosophers, the Putnamian argument is not an argument about the *relevant* nature of an artifact: its function. So, both parts in the debate accept that the externalist needs a different argument.

Such arguments of ignorance or error concerning the function of a certain artifact have been provided, from an externalist perspective, by Kornblith and Nelson. However, those arguments have been replied by Schwartz and Thomasson. One objection is that the artifactual kind terms involved in the externalist arguments are terms without referent, and, thus, the arguments do not show what they pretend. I think that this objection does not work because there is a significant difference between general terms without reference and terms for artifacts as those in the arguments.

But there is another objection that can be made to the externalist arguments from descriptivism: the externalist arguments by Kornblith and Nelson have a limited, local, scope, in such a way that a descriptivist proposal such as that of Thomasson is not undermined in its main point. I think that this descriptivist reply is right, so externalism requires arguments of ignorance or error with a universal scope. And I think that such arguments are available. I present such arguments and, furthermore, I reply the possible objection of the lack of reference of the artifactual kind terms involved. My conclusion is that if those externalist *universal* arguments are coherent, I think that externalism has the needed arguments for its plausibility, but perhaps those arguments are not available for *every* artifact and *every* term for artifact. So, at last, those arguments could show that externalism is plausible at least for some artifactual kind terms, but perhaps not for all.

Introducción

La intuición que hay detrás de la distinción entre las clases artificiales y las clases naturales es que las primeras son creaciones humanas, en tanto que las segundas delimitan áreas naturales ajenas a la intervención humana. Así, ejemplos de clases naturales son el agua, el oro, o los gatos. De manera general, las sustancias químicas que se encuentran en la naturaleza y las especies biológicas constituyen clases naturales paradigmáticas, y quizá se podrían incluir otras clases o géneros, como algunos fenómenos físicos y los colores. Ejemplos paradigmáticos de clases artificiales son los diferentes tipos de artefactos: las mesas, los vestidos o los aviones. En general, los artefactos incluyen todos aquellos objetos que atribuimos a la creación humana, desde objetos muy sencillos, como un remo rudimentario, hasta creaciones muy sofisticadas, como el más sofisticado de los ordenadores.

Por tanto, una primera diferencia intuitiva entre las clases naturales y las clases artificiales es la intervención humana en la existencia y la evolución de las segundas, pero no de las primeras. Tenemos la idea que nosotros fabricamos y mejoramos los artefactos con el paso del tiempo; en cambio, la evolución que pueden sufrir algunas clases naturales, como las especies biológicas, es del todo ajena a nuestros intereses e intenciones (exceptuando algunas clases de especies animales y vegetales seleccionadas artificialmente por el hombre). La intervención humana no se limita al origen y a las sucesivas modificaciones de los artefactos, sino que también está la intención de otorgar una finalidad a las clases artificiales. Éstas suponen un agente inteligente con unas ciertas intenciones respecto a la cosa creada. En nuestro contacto cotidiano con los artefactos, tenemos la intención de usarlos con unas determinadas finalidades. Así, las intenciones humanas parecen jugar un papel importante tanto a la hora de crear un tipo de artefacto, como a la hora de modificarlo y usarlo con una cierta finalidad. Aparte de este factor intencional, también parece haber presente un factor funcional en las clases artificiales: la función se presenta, de entrada, como la característica distintiva de muchos tipos de artefactos. Es cierto que también hay clases consideradas *artificiales*, como todas las clases incluidas en la categoría de *obras de arte* (clases como *sinfonía*, *escultura*, etc.) que, a pesar de ser producidas intencionalmente por el hombre, no tienen a primera vista una función clara o fácilmente especificable. Dejaré de lado este tipo de

artefactos, y me centraré en los artefactos a los que típicamente atribuimos una función, como los lápices, las sillas o los aviones.

Hasta aquí, en una primera aproximación, podemos decir que lo distintivo de los artefactos es que son productos humanos intencionados a los que, en principio, atribuimos una función. Algunos autores (Schwartz 1978, Kornblith 2009) han enfocado la particularidad de las clases artificiales contrastándolas con las clases naturales: lo que caracteriza a una clase natural es la naturaleza *subyacente* de los miembros que la componen, mientras que lo característico de los artefactos, en tanto que artefactos, no es una naturaleza *subyacente* o microscópica. Parece que la distinción entre las diferentes clases naturales se basa en criterios que contemplan la composición molecular o genética, características *subyacentes* que tienen que ver con estructuras no accesibles cognitivamente si no es por investigación empírica; en cambio, la distinción entre las clases artificiales parece responder a un criterio que contempla otro tipo de características, como la función, la forma, etc., características que no necesitan de ningún estudio empírico ulterior porque ya las conocemos. Al menos, de entrada, parece que la división de las clases artificiales no está afectada por nada que haya que descubrir.

Los aspectos diferenciadores apuntados hasta aquí entre las clases naturales y las clases artificiales parecen justificar la habitual distinción y contraste entre las clases naturales y las clases artificiales. Pero hay que notar que esta primera aproximación a la distinción entre las clases artificiales y las naturales, basada en la actuación humana y/o en la no existencia de una naturaleza subyacente de los artefactos, no es demasiado fina porque puede dejar en un estatus indefinido a algunas clases como, por ejemplo, los analgésicos y otros compuestos químicos sintetizados por el hombre a partir de sustancias naturales, o las razas de diferentes animales domésticos o especies vegetales *creadas* por el ser humano mediante procesos de selección artificial. Así que puede haber artefactos, como ciertos medicamentos, cuya naturaleza *subyacente* juegue un papel importante a la hora de determinar que sean el tipo de artefacto que son. Y puede haber clases naturales que no existirían sin la intervención humana y a las que otorgamos una finalidad. Seguramente los mismos términos ‘clase natural’ y ‘clase artificial’ son vagos y pueden darse casos de indefinición respecto a algunas clases. Sin embargo, en la presente tesis, voy a ocuparme de artefactos o clases artificiales que, de

entrada, consideramos artefactos paradigmáticos, como los lápices o las sillas. Éstos son casos ejemplificados por objetos materiales concretos de creación humana intencional y a los que atribuimos una cierta función. Dejaré de lado, como he dicho, todos aquellos objetos considerados obras de arte porque, aunque también son objetos materiales concretos fruto de una creación humana intencional, no se les reconoce una función de manera unánime. Igualmente, no tendré en cuenta el caso de otros *artefactos* considerados propiamente como creaciones humanas intencionales, tales como las instituciones sociales (por ejemplo, el matrimonio) y los objetos o valores abstractos (por ejemplo, el dinero) porque, en este caso, a pesar de que podamos atribuirles una función, no son objetos materiales. Asumo, por tanto, que hay alguna diferencia entre las clases naturales y las clases artificiales, y que hay ejemplos paradigmáticos de ambas, aunque el límite entre ellas no sea preciso y formen un continuo con diferentes gradaciones antes que dos mundos totalmente separados e inconexos.

Acepto también la distinción entre términos (típicos) de clase natural, como ‘agua’, ‘oro’, ‘tigre’, y términos (típicos) de clase artificial, como ‘reloj’, ‘silla’, ‘bolígrafo’. Tanto los términos de clase natural como los términos de clase artificial son considerados términos generales. Los términos generales son aquéllos que, en oposición a los términos singulares como los nombres propios, se aplican a muchos individuos. Así, según una concepción muy extendida que asumo –y sin tener en cuenta la posición que considera que los nombres propios son predicados que se aplican a todos los individuos que tienen el mismo nombre- , los términos singulares como ‘Sócrates’ designan a un único individuo -al menos dentro de un contexto conversacional concreto- mientras que los términos generales se aplican a una serie de individuos o ejemplares que no quedan singularizados por el término. El uso que hacemos de un nombre propio es un uso designativo de la entidad portadora del nombre; en cambio, los términos generales admiten tanto un uso designativo como un uso predicativo. Podemos usar un término general como ‘agua’ para designar la sustancia agua o la propiedad de ser agua, como cuando decimos ‘El agua es común en nuestro planeta’, y podemos usar el mismo término para aplicarlo a una muestra particular de agua, como cuando decimos ‘Esto es agua’. Estos dos usos de los términos generales también son posibles para los términos de clase artificial. Así, podemos decir -al menos en castellano- ‘El ordenador ha ahorrado mucho tiempo al hombre’, donde ‘ordenador’ tiene un uso designativo, y

también podemos decir ‘Eso es un ordenador’, donde ‘ordenador’ tiene un uso predicativo.

La presente tesis se centra en un aspecto semántico de los términos para artefactos: la determinación de la referencia para tales términos. Esta cuestión ha sido bastante desarrollada para los términos de clase natural y se trata de ver cómo se desarrolla en el caso de los términos de clase artificial. Entenderé que la referencia es una relación que se da entre un término y una parte del mundo a la que el término en cuestión refiere, el referente. De manera muy general, se considera que el referente de un término juega un cierto papel semántico a la hora de aclarar cuál es el significado de los términos; podemos ver el referente de un término como una parte del significado del término o como un contribuyente a ciertos aspectos semánticos de nuestras palabras. Sin embargo, al menos en el caso de los términos generales, no es algo obvio cuál es esa parte del mundo, el referente, a la que un término general refiere. Podría considerarse que la parte del mundo a la que refiere el término es, por ejemplo, la extensión del término: el conjunto de todos los objetos a los que el término se aplica con propiedad. O podría considerarse que lo referido por un término general es algo abstracto como, por ejemplo, en el caso de los términos de clase natural, una especie o una sustancia. O también podría considerarse que el referente de un término general es una propiedad que presentan las muestras de la clase. Así, por ejemplo, el referente de ‘agua’ sería la propiedad de *tener la composición molecular H₂O* que presentan las muestras de agua. En este trabajo, no me comprometo con ninguna de estas posiciones acerca del referente de los términos generales; creo que lo que voy a defender en la tesis es compatible con diferentes concepciones sobre cuál es, en última instancia, el referente de un término general. Ahora bien, por comodidad expositiva y porque las propiedades que tienen las muestras de las clases parecen estar implicadas, en alguna medida, en cualquier explicación de cuál es el referente de los términos de clase general, a veces hablaré de las propiedades de las cosas como si fueran los referentes de los términos, pero sin querer comprometerme con esta posición y con la esperanza que lo que diga sea compatible con otras concepciones sobre el referente de los términos generales.

La cuestión central de esta tesis, más que cuál es exactamente el referente de los términos, es qué hace que un término general (y, en especial, un término para artefacto) se aplique a lo que se aplica, qué determina el dominio de aplicación del término. Me

centraré en *cómo se determina la referencia* de un término, cómo se establece y mantiene la relación entre un término y aquello a lo que se aplica. Esa manera en que se determina la referencia de un término puede implicar o no cierto estado cognitivo del hablante, según de qué teoría se trate. Hasta ahora, a grandes rasgos, el asunto de cómo se determina la referencia de los términos generales (y también de algunos términos singulares) ha girado básicamente en torno a dos grandes teorías: el internismo y el externismo semánticos.

El internismo semántico respecto a la determinación de la referencia es el punto de vista que defiende que la referencia sólo puede venir determinada por alguna entidad *interna* cognitivamente accesible a los hablantes competentes (bien a cada hablante competente, bien a la comunidad de hablantes en su conjunto) que selecciona únicamente la extensión del término. Dicha entidad cognitivamente accesible puede ser una descripción que especifique las características definitorias de la clase (en el caso de los términos generales) o puede ser otro tipo de entidad cognitivamente accesible sin material descriptivo (algo como una imagen u otro tipo de entidad). El descriptivismo es la postura internista que defiende que la entidad disponible cognitivamente a los hablantes y que determina la referencia del término es una descripción. Por tanto, consideraré que toda teoría descriptivista es también internista, pero no necesariamente a la inversa.

El externismo semántico respecto a la determinación de la referencia puede considerarse como la teoría que niega el internismo (y, por tanto, también el descriptivismo semántico). Esto es, como el punto de vista que sostiene que la referencia de los términos no está necesariamente determinada por algo cognitivamente accesible (sea una descripción o sea otro tipo de entidad). La formulación positiva de la teoría externista es que la referencia de los términos viene determinada, sin mediación de alguna entidad cognitiva, por características *externas*, por características del mundo, que podrían no ser cognitivamente accesibles a cada hablante competente o a la comunidad de hablantes. La diferencia básica, entonces, entre el internismo y el externismo semánticos, respecto a la determinación de la referencia, es la cuestión de la accesibilidad a lo que determina la extensión (o ámbito de aplicación) de los términos.

En este punto es preciso hacer varias aclaraciones. En primer lugar, el internismo semántico no está obligado a negar que los referentes de los términos generales puedan ser, en principio, características *externas* al hablante como son las propiedades que tienen los miembros de una clase. Lo crucial para el internismo es que los hablantes tengan *algún* acceso cognitivo a esas propiedades (bien porque se trata de una serie de características superficiales directamente observables, bien porque el hablante puede acceder cognitivamente a esas características por introspección, bien porque los hablantes disponen de algún mecanismo de reconocimiento de esas características, etc.) mediante algún mecanismo cognitivo (la observación, la introspección, el reconocimiento, etc.). El mecanismo cognitivo más célebre que se ha postulado desde el internismo, mecanismo que permite llegar al referente del término, es el mecanismo de la asociación con el término de una descripción de la propiedad o propiedades singularizadoras del referente; en este caso, se trata de un internismo descriptivista o descriptivismo respecto a la determinación de la referencia del término. En segundo lugar, conviene aclarar que el externismo no se compromete con que aquello que determina la referencia de los términos generales sea siempre cognitivamente inaccesible, sino con que, al menos en ocasiones, *podría* resultar inaccesible y, por tanto, no es verdad que la referencia de los términos generales se determine *necesariamente* de la manera que afirma el internismo.

Por último, quiero hacer otras dos puntualizaciones de tipo general: al hablar de *internismo* y *externismo* semánticos no tengo ninguna intención de referirme a ningún tipo de internismo o externismo sobre el contenido de nuestros conceptos o pensamientos. La cuestión del internismo/externismo psicológicos queda fuera de los propósitos presentes. No presupondré nada sobre cómo se determina el contenido de los pensamientos en estas dos propuestas semánticas y me limitaré a la cuestión puramente semántica de la determinación de la referencia de los términos. Por otro lado, al abordar la cuestión de la determinación de la referencia de los términos de clase artificial, pretendo ser neutral sobre si los términos para artefactos son designadores de clases o, más bien, tienen un uso preferente como predicados que se aplican a los ejemplares de una clase. Aunque, por comodidad, hablaré del *referente* de los términos para artefactos o de *designación del referente*, todo lo que expongo en la tesis pretende ser también compatible con la concepción – expuesta antes- de los términos de clase artificial como predicados que aplicamos a ciertos objetos.

El estudio presente sobre la determinación de la referencia de los términos para artefactos consta de cuatro capítulos. En el primero de ellos, me detengo en la cuestión de la determinación de la referencia para los términos más estudiados hasta ahora: los nombres propios y los términos de clase natural. Para ello, sigo a grandes rasgos el desarrollo histórico de la cuestión, empezando por las teorías de corte internista para los nombres propios y los términos de clase natural, y siguiendo con la crítica externista de la nueva teoría de la referencia y la propuesta positiva de esta teoría para los términos mencionados. Este estudio preliminar sobre la determinación de la referencia de los nombres propios y los términos de clase natural, y las conclusiones citadas, pretende servir como punto de partida y de comparación para el caso de los términos de clase artificial.

En el segundo capítulo, me centro ya en las clases artificiales y los términos de clase artificial. Primero, considero la cuestión de la naturaleza de las clases artificiales, estableciendo una comparación entre las clases naturales y las clases artificiales y presentando diferentes concepciones metafísicas sobre los artefactos. Después, tras haber separado la cuestión metafísica de la naturaleza de los artefactos de la cuestión propiamente semántica de la determinación de la referencia de los correspondientes términos, abordo la cuestión semántica de los términos de clase artificial. Para ello, presento la primera propuesta histórica de cierta importancia para los términos de clase artificial: la propuesta externista putnamiana.

En el tercer capítulo, expongo las principales críticas, que se han hecho desde diferentes versiones de descriptivismo, contra la teoría de Putnam para los términos para artefactos. Estas críticas son, básicamente, las críticas de Schwartz y de Thomasson, y van acompañadas de propuestas alternativas a la putnamiana acerca de la determinación de la referencia de los términos de clase artificial. Analizo estas propuestas descriptivistas y considero diferentes réplicas que se pueden hacer a sus críticas.

En el cuarto y último capítulo, considero la viabilidad de defender una postura externista para los términos para artefactos. Como el argumento presentado por Putnam no es tan convincente como su argumento externista para los términos de clase natural, presento los dos argumentos externistas más importantes que se han ofrecido hasta la fecha desde la sugerencia de Putnam. Sin embargo, estos argumentos no están libres de

críticas por parte del descriptivismo y, además, argumento que presentan un alcance limitado en lo que pretenden mostrar. Expongo entonces, algunos argumentos que considero que son los que tienen más opciones de mostrar que el externismo semántico puede ser plausible para los términos de clase artificial. Estos argumentos, si son coherentes, ciertamente tienen un alcance más amplio que los argumentos anteriormente presentados, pero, sin embargo, presentan una idiosincrasia particular que hace que quizá no sean extensibles a *cualquier* artefacto y *cualquier* término de clase artificial.

Capítulo 1. Los nombres propios y los términos de clase natural

En este primer capítulo, trato la cuestión de la semántica de los nombres propios y de los términos de clase natural. Este análisis sirve como punto de partida para el estudio, en capítulos siguientes, de la cuestión semántica específica de la determinación de la referencia de los términos de clase artificial, el tema central de la tesis. El estudio previo del funcionamiento de los nombres propios y de los términos generales me parece justificado porque la semántica de los términos para artefactos se concibe en gran parte a partir de, y por analogía con, los términos de clase natural. De hecho, los filósofos que han pensado en la semántica apropiada para los términos de clase artificial, lo han hecho considerando si su propuesta semántica favorita para los términos de clase natural es extensible o no a los términos para artefactos. Por otro lado, la consideración previa de la semántica de los nombres propios está justificada porque, así como la semántica de los términos de clase artificial se piensa en comparación con la semántica de los términos de clase natural, estos últimos se contraponen a menudo, como ejemplos paradigmáticos de términos generales, a los nombres propios, ejemplos paradigmáticos de términos singulares.

Presento la semántica tanto de los nombres propios como de los términos de clase natural como una discusión entre el internismo y el externismo, en la que, siguiendo a grandes rasgos el curso histórico del debate, primero el internismo presenta su propuesta y, después, ésta es criticada por el externismo o nueva teoría de la referencia. Empiezo por hacer un breve repaso, en la primera sección, **1.1**, de las propuestas semánticas clásicas de Frege y de Russell para los nombres propios. A continuación, en la sección **1.2**, presento otras propuestas semánticas más recientes para los nombres propios en la línea de lo defendido por Frege y por Russell, propuestas que intentan superar los problemas del modelo clásico. En la tercera sección, **1.3**, expongo la extensión de ese modelo, al caso de los términos generales, centrándome en los términos de clase natural. En la sección **1.4**, repaso las críticas que se han hecho, desde la nueva teoría de la referencia, a ese modelo clásico, en sus diferentes versiones, para los nombres propios. Y, por último, en la quinta sección, **1.5**, presento, por un lado, las críticas de Kripke y Putnam al modelo de Frege y Russell aplicado a los términos de clase natural, y, por otro, la propuesta positiva que hace la nueva teoría de la referencia para estos términos generales.

1.1. El internismo clásico para los nombres propios

1.1.1. La propuesta internista de Frege

A finales del siglo XIX, Frege desarrolló una teoría del significado para los términos y expresiones de nuestro lenguaje. Me centraré en su propuesta semántica para los términos singulares y, especialmente, para los nombres propios. El análisis de la función semántica de los nombres propios ha preocupado a los filósofos interesados en cuestiones semánticas, pues los nombres propios constituyen un caso paradigmático respecto a la relación entre las palabras y el mundo: intuitivamente, los nombres parecen representar la cosa que nombran.

La teoría fregeana introdujo dos componentes semánticos presentes en los términos singulares: el referente y el sentido¹. De manera informal, el referente de un término singular es la entidad que denota. Así, por ejemplo, considerando casos de nombres propios, el referente del nombre propio ‘Héspero’ es el planeta Venus, y el referente del nombre propio ‘Fósforo’ es también el planeta Venus. Estos dos términos singulares comparten, pues, el referente. Pero Frege se dio cuenta de que entre estos dos términos hay alguna diferencia desde un punto de vista semántico, porque entre los enunciados ‘Héspero es Héspero’ y ‘Héspero es Fósforo’ hay una diferencia en valor cognoscitivo: el primer enunciado no es informativo, y el segundo sí. Un usuario competente del lenguaje que asintiera a la verdad del primer enunciado podría negar la verdad del segundo. Por tanto, ‘Héspero’ y ‘Fósforo’ deben de contribuir de maneras diferentes al valor cognoscitivo de los enunciados en que aparecen. Esta diferente contribución se explica por la aportación a la oración, por parte de los dos términos singulares, de diferentes sentidos, a pesar de que ambos términos, ‘Héspero’ y ‘Fósforo’, comparten el mismo referente. Podemos decir de manera informal que el sentido es la manera o modo concreto en que el término presenta su referente, y este modo de presentación es diferente en los casos de ‘Héspero’ y de ‘Fósforo’. El modo de presentación que los hablantes asocian con cada término singular es algo así como un conjunto de características que singularizan el referente del término². Así, a cada término singular

¹ Estos dos componentes semánticos también están presentes en otras expresiones lingüísticas, pero sólo consideraré la teoría fregeana en lo relativo a los términos singulares.

² O, al menos, ésta es una interpretación estándar de los sentidos fregeanos.

los hablantes asocian una serie de características que pueden ser diferentes incluso para dos términos correferenciales como ‘Héspero’ y ‘Fósforo’.

En el caso de la comparación entre dos descripciones definidas correferenciales como ‘el maestro de Aristóteles’ y ‘el autor de *La República*’ es fácil para Frege explicar la diferencia de sentido o de modo de presentación asociado entre los dos términos.

Ambos presentan el mismo referente, Platón, de maneras diferentes mediante ciertas informaciones individualizadoras claramente diferentes que las propias descripciones definidas especifican en cada caso. ‘El maestro de Aristóteles’ presenta a Platón como maestro de Aristóteles, y ‘el autor de la *República*’ presenta a Platón como autor de la *República*. Ahora bien, ¿cuál es el sentido o la manera de presentar a Venus mediante ‘Héspero’ y ‘Fósforo’, respectivamente? Aquí el problema es que los nombres propios no especifican de manera explícita ningún tipo de información que individualice a su referente. Pero, como se ha visto, por el problema de la posible diferencia en valor cognoscitivo de ciertos enunciados que sólo difieren en los nombres propios correferenciales que contienen, es necesario, para Frege, que éstos tengan diferentes sentidos. Una respuesta, inspirada en Frege, es que los hablantes asocian, por el mero hecho de ser hablantes competentes del lenguaje, ciertas características a cada nombre propio. Así, los hablantes asocian a ‘Héspero’ ciertas características individualizadoras de Venus (*ser cuerpo luminoso visible algunos días del año por poniente tras el atardecer*), mientras que asocian a ‘Fósforo’ otras características diferentes individualizadoras de Venus (*ser cuerpo luminoso visible algunos días del año por levante al amanecer*).

Otro elemento importante de la teoría fregeana es que el sentido determina el referente, pero no a la inversa: un mismo sentido (como pasa con ‘Héspero’ y ‘el cuerpo luminoso visible algunos días del año por poniente tras el atardecer’) no puede conducir a dos referentes distintos, pero dos sentidos distintos sí pueden conducir a un mismo referente (como se ha visto con ‘Héspero’ y ‘Fósforo’). Para llegar hasta el referente del término singular, hay que conocer previamente el sentido del término. Así, aunque el significado de los términos singulares tiene los dos componentes mencionados, el sentido es la entidad mediadora que permite determinar y llegar hasta el referente (si lo hay). Por otro lado, los sentidos o modos de presentación de las expresiones lingüísticas son, además,

los constituyentes de las proposiciones, y su contribución a las proposiciones en que aparecen consiste en determinar sus condiciones de verdad.

Se ha discutido mucho qué son y cómo hay que entender, en la teoría fregeana, los sentidos de los nombres propios. Durante mucho tiempo, la interpretación estándar de la teoría de Frege ha sido que el sentido de todo nombre propio equivale al sentido de alguna descripción definida que es conocida por los hablantes y que éstos asocian al nombre propio. Tal interpretación es la que consideraré cuando presente las críticas de la nueva teoría de la referencia al descriptivismo clásico, pues es la interpretación que adoptan los introductores de la mencionada teoría. Esta interpretación es plausible porque si, por ser hablantes competentes, asociamos con cada nombre propio un conjunto de características que individualizan al portador del nombre de una determinada manera, parece natural pensar que conocer esas características es conocer alguna descripción que las especifica. Entonces, conocer el sentido de un nombre parecería que es conocer esa descripción que individualiza al portador del nombre. Pero Frege no dice explícitamente en sus escritos que el significado de un nombre propio sea sinónimo al significado de una descripción definida. Además, hay otras interpretaciones de su teoría, como la de Evans o la de McDowell, que defienden que no es una teoría *descriptivista*, basándose, entre otras cosas, en consideraciones de Frege sobre los nombres propios sin referente³.

Dejaré de lado la cuestión exegética acerca del descriptivismo en Frege. Pero, desde la terminología que adopto, lo que importa es considerar la teoría de Frege como una teoría semántica *internista* para los nombres propios. Considero internista a toda teoría semántica para la que aquello que determina la referencia de los términos (sea o no conocido mediante una descripción) es algo cognitivamente accesible para los hablantes competentes o, al menos, para algunos hablantes de la comunidad lingüística. Es decir, toda teoría para la que no es posible que se determine la referencia de un término sin que *ningún* hablante tenga acceso cognitivo a las características singularizadoras del referente del término. En este sentido, la teoría de Frege es internista, pues presupone

³ Según la exégesis de estos autores, un nombre propio sin referente, como ‘Vulcano’ o ‘Superman’, no puede tener sentido para Frege, aunque haya una descripción definida asociada al término. Por tanto, concluyen, el sentido no puede equivaler a una descripción definida.

que el acceso cognitivo a los sentidos (el componente semántico que, a su vez, determina el referente) se produce mediante la identificación de las características distintivas del portador del nombre, se produzca o no esta identificación mediante alguna descripción definida. No se trata de que el significado de un nombre propio sea algo *interno* en el sentido de algo *subjetivo* o *mental*, pues, para Frege, los sentidos son intersubjetivos y no psicológicos. Es la accesibilidad de cada hablante a las propiedades singularizadoras del objeto referido la que marca el carácter internista del significado de los nombres propios en la teoría de Frege, aunque el reconocimiento de esas propiedades no implique asociar una descripción individualizadora al nombre propio. Así, una teoría puede ser internista (para los nombres propios) sin que ello implique que sea descriptivista. Por ejemplo, una teoría internista para los nombres propios podría defender que los hablantes (o algunos hablantes) asocian a cada nombre propio algo así como una imagen mental que singulariza al referente del nombre. Tal entidad sería cognitivamente accesible a los usuarios (o a algún usuario) del nombre por el mero hecho de ser hablantes competentes sin que la capacidad para identificar esas características individualizadoras mediante una imagen implique necesariamente que el hablante sea capaz de explicitar una descripción que las enumera o cómo las reconoce.

1.1.2 La propuesta descriptivista de Russell

Russell no comparte la idea fregeana de que todos los términos de nuestro lenguaje tienen dos componentes, el sentido y el referente. Para Russell, los auténticos términos singulares sólo tienen referente: su significado se agota en la cosa referida por el término. Esta concepción monista del significado, respecto a los términos singulares, es la que dirige toda su teoría semántica. A partir de esta concepción de los términos singulares, concluye, por una serie de consideraciones, que las descripciones definidas (términos de la forma *el F*) no pueden ser términos singulares; esto es, su función semántica en los enunciados en que aparecen no es –tal como debería ser si fuesen auténticos términos singulares- aportar un objeto a la proposición. Las consideraciones que le llevan a afirmar esto son varios *puzzles* o problemas graves que se generan si consideramos a las descripciones definidas como términos singulares. Brevemente, estos problemas tienen que ver con la violación del principio de sustituibilidad para términos correferenciales en contextos indirectos cuando hay descripciones definidas implicadas y con problemas derivados del hecho que algunas descripciones definidas no tienen referente. Esto le lleva a proponer un análisis lógico de las descripciones

definidas, análisis ampliamente compartido por los filósofos, en el que se muestra que las descripciones definidas funcionan como expresiones cuantificacionales, y no como términos singulares.

En 1905, Russell piensa que el análisis que propone de las descripciones definidas es ampliable a los nombres propios ordinarios (nombres como ‘Venus’, ‘París’ o ‘Jorge IV’). Éstos tampoco son auténticos términos singulares, pese a las apariencias. El motivo de Russell para rechazar que los nombres propios sean términos singulares es que los problemas o puzzles que se generaban a partir de considerar a las descripciones definidas como términos singulares se generan de la misma manera si tratamos a los nombres propios como términos singulares. Así, en contextos indirectos, como en ‘Juan cree que Héspero es visible al atardecer’, al sustituir el nombre propio ‘Héspero’ por ‘Fósforo’ (término correferencial), el valor de verdad del enunciado puede cambiar. Se viola, entonces, en estos contextos, el principio de sustituibilidad para términos correferenciales. Si los nombres propios fueran términos singulares y, por tanto, su única función semántica fuera la de aportar un objeto a la proposición, dos enunciados cualesquiera que sólo difirieran entre sí en la aparición de diferentes nombres propios correferenciales, el principio de sustituibilidad no se violaría en ningún caso. Otro problema de considerar a los nombres propios como términos singulares consiste en que no se puede explicar que enunciados de la forma *p o no p* que contienen nombres propios sin referente, como ‘Vulcano es un planeta o Vulcano no es un planeta’, sean enunciados necesariamente verdaderos. Para explicar las condiciones de verdad de estos enunciados debemos apelar al objeto que aportan a la proposición los (supuestos) términos singulares que aparecen. Pero si el término no aporta ningún objeto, las condiciones de verdad quedan inexplicadas. Por último, otro problema relacionado con los nombres propios sin referente y derivado de tratar a los nombres propios como términos singulares es que tampoco podemos explicar que los enunciados negativos de existencia con nombres propios sin referente, como ‘Vulcano no existe’, sean significativos y tengan, por tanto, condiciones de verdad. De nuevo, si aparece un nombre propio sin referente y lo consideramos un término singular, no puede aportar un objeto a la proposición. Y no podemos dar cuenta de las condiciones de verdad que parece tener la proposición.

Así pues, como en el caso de las descripciones definidas, la función semántica de los nombres propios no puede ser aportar un objeto a las proposiciones en que aparecen. Entonces, ¿cuál es su función semántica? Para Russell, los nombres propios comunes son descripciones definidas encubiertas, meras abreviaturas de descripciones. Así, por ejemplo, ‘Scott’ es una mera abreviación de ‘el autor de *Waverly*’. Entonces, la función semántica de los nombres propios es la misma que la de las descripciones definidas: una función análoga a la de las expresiones cuantificacionales, que en ningún caso implica aportar un objeto a las proposiciones en que aparecen. Con este tratamiento de los nombres propios, los problemas o puzzles mencionados no se generan⁴.

Así, para Russell, los nombres propios comunes de nuestro lenguaje no son términos singulares, no refieren a un objeto, sino que se analizan como descripciones definidas (en términos de cuantificadores y predicados). Los auténticos términos singulares, aquéllos cuyo significado se agota en el objeto que denotan, son otros términos diferentes a los nombres propios ordinarios.⁵ Pero lo que me interesa aquí es que, para éstos últimos, Russell propone una semántica explícitamente descriptivista, pues los nombres propios no son más que descripciones definidas encubiertas. El significado de un nombre propio equivale al de una descripción definida. Y el significado de esa descripción es, además, accesible a los hablantes, pues, en un análisis lógico ulterior las descripciones definidas se descomponen en términos que significan entidades que conocemos *directamente*: nuestras propias sensaciones y algunos universales.

Si formulamos, por ejemplo, un juicio sobre Julio César, es evidente que Julio César mismo no está presente en nuestro espíritu, puesto que no tenemos un conocimiento directo de él. Tenemos en el espíritu alguna

⁴ No expondré aquí cómo se solucionan los puzzles.

⁵ Son, para Russell, los nombres de datos de los sentidos (*sense data*). A partir de consideraciones epistemológicas de tipo cartesiano, Russell afirma que lo único que conocemos *directa e infaliblemente* son los propios datos de los sentidos y algunos *universales*. Por otro lado, tenemos que *conocer* el significado de nuestras palabras. Por tanto, nuestras palabras no pueden significar sino datos de los sentidos y universales en última instancia. Y los nombres de los datos de los sentidos propios son los auténticos términos singulares, los *nombres lógicamente propios*.

*referencia*⁶ de Julio César: “El hombre que fue asesinado en los Idus de marzo”, “el fundador del Imperio romano”, o acaso simplemente, “el hombre cuyo nombre era *Julio César*”. (En esta última definición, Julio César es un ruido del cual tenemos un conocimiento directo.) Así, nuestro juicio no significa exactamente lo que parece significar, sino que significa algo que envuelve, en lugar de Julio César, alguna definición de él, compuesta exclusivamente de particulares y universales, de los cuales tenemos un conocimiento directo. (Russell 1986, 56-7)

Para Russell, los significados son entidades que tenemos que conocer; si no, nuestras palabras serían meros sonidos y nada más. Russell comparte con Frege este principio internista de transparencia de los significados: los significados de nuestros términos (en este caso, de los nombres propios) nos resultan transparentes, sabemos de lo que hablamos cuando usamos nuestros términos. El internismo semántico se fundamenta en la plausibilidad del principio de transparencia del significado de nuestros términos. Conocemos lo que significan los nombres propios que usamos, pues tenemos un acceso cognitivo a las propiedades individualizadoras del portador del nombre en virtud de ser hablantes competentes. En el caso de Frege, el sentido de un nombre propio es algo cognitivamente accesible (mediante una descripción o mediante otro mecanismo) a los hablantes y que les permite llegar al referente. En el caso de Russell, no sólo tenemos un acceso cognitivo a las características que individualizan al portador del nombre, sino que, además, el acceso a la cosa nombrada sólo puede ser mediante una descripción, pues conocemos las cosas del mundo de manera *indirecta*, únicamente por descripción.

Aparte de que, para algunos intérpretes de la obra de Frege, éste no es un descriptivista como Russell, una diferencia importante entre estos dos filósofos es que, mientras para Frege los sentidos son intersubjetivos, compartidos, para Russell los significados son, en última instancia, completamente subjetivos. Esto, sin duda, es un punto débil de la teoría russelliana, ya que no puede dar cuenta del carácter comunicativo que presenta, al menos a primera vista, el lenguaje. Una consecuencia indeseable de su teoría es que no podemos entender realmente los significados de los términos de los otros hablantes; cada hablante sólo conoce –eso sí, de manera infalible– el significado de sus propias

⁶ Aquí, hay que entender ‘*referencia*’ como *descripción* o *información descriptiva*.

palabras pues éstas siempre remiten, en un último análisis a sensaciones propias, lo único que conocemos de manera *directa* (junto con algunos universales intersubjetivos). Pero las teorías de Frege y de Russell también presentan puntos en común. Ambas son teorías semánticas internistas. Como tales, presentan la ventaja, como se ha visto, de respetar la plausibilidad inicial que parece tener el principio de la transparencia para los significados. Otra ventaja de las teorías fregeana y russelliana es que son capaces de explicar, bien apelando a la diferencia de sentidos o bien a la diferencia de descripciones definidas encubiertas implicadas, la diferencia en valor cognoscitivo de los enunciados informativos de identidad respecto a los enunciados de identidad que no son informativos. Además, la teoría russelliana permite atribuir significado a nombres propios que no tienen referente, ya que simplemente el nombre encubriría alguna descripción definida cuyo significado sí puede analizarse hasta llegar a sus últimos componentes semánticos. Es controvertido si la teoría fregeana presenta también esta última ventaja, y dejaré al margen la cuestión.

1.2. El nuevo descriptivismo modificado para los nombres propios: la teoría cúmulo y el descriptivismo débil o de fijación de la referencia

1.2.1. La teoría cúmulo

Aunque he calificado la propuesta de Frege para los nombres propios de *internista*, dejando de lado si es o no también descriptivista, las primeras interpretaciones de la teoría fregeana la consideraron una teoría descriptivista, a la par con la teoría de Russell para los nombres propios. Esta interpretación, previa al estudio exegético de los textos de Frege, tiene cierta plausibilidad, y, bajo esta interpretación, se expusieron los inconvenientes que presentaba la teoría descriptivista *Frege-Russell*. Así, Searle en 1967 introduce una teoría descriptivista modificada para los nombres propios con el fin de salvar los inconvenientes del descriptivismo clásico que atribuye a Frege y a Russell. Una de las desventajas que Searle ve en el descriptivismo clásico es que la descripción asociada con un nombre propio y que proporciona el significado del término puede variar entre los hablantes. Aunque Frege afirma que el sentido de un nombre propio es intersubjetivo, compartido por los hablantes, Searle cree que es posible que, al menos a veces, diferentes personas asocien descripciones diferentes con el nombre. Por ejemplo, un hablante *A* puede asociar con ‘Aristóteles’ la descripción ‘el discípulo más brillante

de Platón' y otro hablante, *B*, puede asociar con el mismo nombre propio la descripción 'el autor de la *Metafísica*'. En este caso, el problema es, para Searle, que, aunque el nombre 'Aristóteles' dicho por *A* y dicho por *B* refiere al mismo objeto, tiene un significado cuando lo usa *A* y otro significado cuando lo usa *B* (en virtud de las diferentes descripciones definidas asociadas por *A* y *B*). Entonces, ¿cómo puede 'Aristóteles' significar dos cosas diferentes si ninguno de los dos hablantes está equivocado sobre Aristóteles y ambos tienen éxito al referir? Una teoría descriptivista necesita, entonces, determinar cuál es la descripción que determina la referencia y da el significado del término, ya que personas diferentes asocian diferentes descripciones con el mismo término.

Otro problema que detecta Searle en lo que considera la teoría descriptivista clásica tiene que ver con ciertas consecuencias modales que se derivan de la idea que el contenido descriptivo expresado por un nombre propio equivale al significado del mismo nombre propio. Si un nombre propio y su descripción asociada funcionan como sinónimos, entonces la atribución de la propiedad expresada por la descripción al referente del término tendría que ser necesariamente verdadera. Si, por ejemplo, 'Platón' significa lo mismo que significa 'el maestro de Aristóteles', entonces 'Platón es el maestro de Aristóteles' tendría que ser un enunciado necesariamente verdadero, y no lo es. Es sólo contingentemente verdadero de Platón que es, o fue, el maestro de Aristóteles. Platón podría haber existido sin haberse dedicado a la filosofía o sin haber conocido a Aristóteles. Y lo mismo puede decirse de la mayoría de descripciones que asociamos a un nombre propio: si se pueden predicar con verdad del portador del nombre, esta predicación es sólo contingentemente verdadera. Por tanto, Searle concluye que el significado de un nombre propio no puede ser el significado de una descripción definida asociada.

En respuesta a estos problemas del descriptivismo tradicional, Searle introduce su teoría descriptivista modificada: una teoría cúmulo (*cluster*) del significado para los nombres propios. Searle propone que sea un cúmulo de descripciones, y no una única descripción, el que dé el significado y determine el referente del término. Así, el significado de un nombre propio viene determinado por un conjunto difuso de diferentes descripciones. Por ejemplo, el significado de 'Platón' es equivalente al significado de "el maestro de Aristóteles" o 'el discípulo de Sócrates' o 'el autor de *La*

República’ o ‘el autor del *Fedón*’ o...’ En general, si N es un nombre propio cualquiera y Dn cualquier descripción definida asociada con N, el significado de N = el significado de (D1 o D2 o D3 o...). Esta disyunción de descripciones no debe entenderse como una disyunción lógica estricta. En una disyunción lógica, bastaría que uno de los disyuntos fuera verdadero para atribuir significado al nombre propio. La idea de Searle es, más bien, que *algunas* (no especificadas) descripciones de la disyunción pueden ser falsas del referente, pero no todas. Muchas de ellas, una *mayoría cualificada* o una *mayoría ponderada* de ellas, tiene que ser verdadera del referente del nombre.

Esta versión modificada de la teoría descriptivista mantiene las ventajas de la propuesta clásica, y pretende superar los inconvenientes. Para resolver el problema de decidir cuál entre diversas descripciones es la que da el significado de un nombre propio, la solución de Searle consiste en decir que un nombre suficiente pero inespecificado de las descripciones asociadas por los hablantes son verdaderas del portador del nombre. Su idea es como si preguntásemos a todos los usuarios de ‘Platón’ que digan lo que consideran hechos establecidos sobre Platón, y tomar estas respuestas como un conjunto de descripciones identificadoras de Platón. Entonces, la condición bajo la cual podríamos decir de un objeto que es Platón es que un número suficiente pero inespecífico de estas descripciones son verdaderas de Platón. Así, ‘Platón’ refiere a Platón no en virtud del hecho que hay una única descripción identificadora que proporciona el sentido del nombre, sino más bien porque el objeto Platón satisface la mayoría o una gran mayoría o un número suficiente de descripciones identificadoras del cúmulo como para contar como el único referente del nombre. No hay problema en que diferentes hablantes asocien diferentes descripciones identificadoras con el mismo nombre propio, el significado del nombre propio viene dado por el cúmulo de descripciones, no por esta o aquella descripción concreta.

Respecto al problema que la descripción asociada al nombre tiene que ser necesariamente verdadera del portador del nombre, Searle dice que aunque ninguna de las descripciones definidas comúnmente asociadas al nombre propio es un sinónimo de éste, la disyunción de las descripciones sí es sinónima del nombre y resulta necesariamente verdadera del referente del nombre. Cualquier descripción aislada podría ser falsa o contingentemente verdadera si se predica del referente, pero hay *algunas* descripciones, en un número no especificado, del cúmulo asociado que sí se

predican de manera necesariamente verdadera del portador del nombre. En términos de mundos posibles esto significa que al menos una descripción identificadora de la disyunción será verdadera en cada mundo posible, aunque no necesariamente la misma. Lo que no puede ser es que todas las descripciones asociadas sean falsas del referente. Por ejemplo, mientras cualquier descripción asociada a 'Aristóteles' podría ser falsa (o haber sido falsa) predicada de Aristóteles, no tiene sentido suponer, según Searle, que todo lo que cualquiera ha creído de Aristóteles es de hecho falso del Aristóteles que existió.

La teoría cúmulo de Searle pretende corregir algunos defectos de lo que él considera el descriptivismo clásico, pero desde una postura internista descriptivista: lo que determina el referente y el significado de los nombres propios es una serie de descripciones accesibles cognitivamente a los hablantes. La novedad que introduce Searle es que lo cognitivamente accesible que determina la referencia de los términos, el cúmulo de descripciones, no es accesible a *cada* hablante competente, sino a toda la comunidad lingüística en su conjunto. Hay que modificar, pues, el principio de transparencia porque es cierto que no todo hablante conoce necesariamente el significado de los términos que usa, ya que un hablante puede asociar una o diversas descripciones al nombre que no sean las que determinan la referencia del término. El principio de transparencia, entonces, sólo es válido aplicado socialmente: es el conjunto de los hablantes, mediante el cúmulo de descripciones que éstos asocian al término, el que conoce el significado de los términos. La teoría de Searle pierde así parte del atractivo que tenía la teoría descriptivista clásica; aquello a lo que un hablante competente concreto tiene acceso cognitivo ya no determina necesariamente la referencia de los términos que utiliza. Además, no está claro que este internismo modificado pueda explicar, como sí lo hacía el internismo clásico, la diferencia en valor cognoscitivo *para un hablante particular* entre enunciados con nombres propios correferenciales como 'Héspero es Héspero' y 'Héspero es Fósforo'. El cúmulo social descriptivo asignado a 'Héspero' y a 'Fósforo' puede diferir, respectivamente, de la información que asocie con cada uno de estos dos términos un hablante particular y, entonces, el cúmulo no explica la posible diferencia en valor cognitivo para el individuo.

1.2.2. El descriptivismo débil o de fijación de la referencia

Otra forma de descriptivismo que pretende superar los inconvenientes del descriptivismo tradicional que asimila el significado de un nombre propio al significado de una descripción definida es el descriptivismo de *fijación de la referencia* o descriptivismo débil⁷. Esta versión no pretende decir que un nombre propio es necesariamente la abreviatura de una descripción definida; la descripción no tiene por qué funcionar como un mero sinónimo del nombre, pues, como había notado Searle, esto conlleva consecuencias indeseables (entre ellas, consecuencias modales indeseables). La idea, entonces, de esta forma de descriptivismo es simplemente que los nombres propios no comparten significado con alguna descripción definida, pero hay alguna descripción que tiene que *fijar* el referente del nombre y conducir hasta él, aunque no proporcione el significado del nombre en cuestión. Sin esta conexión entre el nombre y el referente, quedaría indeterminado a qué refieren los usos de los nombres propios.

De nuevo, la idea descriptivista es que los hablantes, al usar un nombre propio, asocian una descripción que es satisfecha únicamente por el portador del nombre y, en virtud de esta asociación semántica, se fija el referente del nombre. Pero, según esta versión, el nombre y la descripción asociada no funcionan como sinónimos. Así, suponiendo que la descripción asociada a 'Platón' es 'el maestro de Aristóteles', 'Platón fue el maestro de Aristóteles' no tiene porqué ser (y no lo es) una oración necesariamente verdadera: 'Platón' y 'el maestro de Aristóteles' no significan lo mismo. Sin embargo, el único modo en que se determina el referente de 'Platón' es mediante una descripción como la sugerida.

⁷ Esta versión *débil* del descriptivismo fue modificada tras la crítica de Kripke y una de sus versiones es conocida como 'descriptivismo causal'. El descriptivismo causal, como se verá más adelante, sigue siendo una teoría sólo de fijación de la referencia, pero descartando que la descripción asociada al nombre y que fija su referencia sea una descripción que especifica un hecho o característica notable (y contingente) del referente.

1.3. Extensión del internismo a los términos generales: el caso de los términos de clase natural

1.3.1. El internismo de Locke para los términos de clase natural

Las propuestas internistas para los nombres propios, términos singulares, pueden extenderse a los términos generales, los que se aplican a más de un individuo. La idea internista es que la extensión de los términos generales queda determinada por una serie de características cognitivamente accesibles a los hablantes (o a la comunidad lingüística en su conjunto) que son compartidas por los miembros de la clase designada por el término. Como en el caso de los nombres propios, no todas las propuestas internistas para los términos generales son de tipo descriptivista; esto es, el acceso cognitivo a las características individualizadoras de la clase no se produce necesariamente mediante material descriptivo. Éste es el caso en la teoría de Locke.

Este filósofo, ejemplo paradigmático del internismo semántico, es uno de los poquísimos autores de la tradición anterior a Frege que habla explícitamente de los términos de clase natural. Locke, a finales del siglo XVII, afirma, por un lado, que tenemos conocimiento de lo que significan nuestras palabras y, por otro lado, que *sólo* conocemos (con certeza) nuestras ideas. Por tanto, las palabras significan ideas. Ahora bien, como las ideas, en condiciones normales, son causadas en la mente por la presencia de entidades extramentales, resulta que las palabras también significan estas entidades extramentales, pero *sólo secundariamente*. Sólo de manera indirecta conocemos los objetos extramentales representados por nuestras ideas. A partir de estas consideraciones semánticas generales, Locke propone, para las clases naturales, la distinción entre esencia nominal y esencia real. Una interpretación plausible de esta distinción en Locke es considerar la *esencia nominal* que corresponde a una clase natural como el conjunto de propiedades perceptibles comunes a todos los individuos de la clase. En cambio, la *esencia real* sería una constitución o estructura interna, de la que dependen causalmente las cualidades que integran la esencia nominal. Así, por ejemplo, la esencia nominal del oro es un conjunto de propiedades perceptibles del oro (su brillo, solidez, color amarillo) que causan en cada uno de nosotros las ideas de brillo, de solidez y de amarillez. Su esencia real es la estructura interna del oro, de la que depende causalmente la esencia nominal; es decir, es aquello del oro que hace que sea brillante, sólido y amarillo.

Bien, ¿cuál es, entonces, el significado de ‘oro’, atendiendo a esta distinción entre esencia real y esencia nominal del oro? Según Locke, el significado primario de ‘oro’ se descompone en una serie de ideas simples que asociamos con el término: la idea de solidez, de brillo, de un cierto color. Todas estas ideas simples conforman la idea compleja de oro, que constituye el significado primario del término. Pero Locke admite que los términos de clase natural también significan secundariamente las propiedades extramentales que causan esas ideas. En este punto, es preciso decir que no está claro en la obra lockeana si esas propiedades extramentales constituyen la esencia nominal y, por tanto, los términos de clase natural significan secundariamente esencias nominales, o bien si la esencia nominal, concebida como la idea compleja constituida por ideas simples, es la significación primaria de los términos de clase natural, y las propiedades objetivas de las cosas (que causarían la esencia nominal) constituirían la significación secundaria de los términos para clases naturales. Estas dos posibilidades al interpretar la esencia nominal en Locke se debe a la propia ambigüedad presente en su obra pues Locke habla a veces de la esencia nominal de una clase natural como si fuera el conjunto de las propiedades objetivas y superficiales que presentan los miembros de la clase, y, otras veces, considera la esencia nominal como la idea compleja que asociamos con su correspondiente término. Sin intentar dirimir esta cuestión, lo que me interesa destacar es que, para Locke, los términos de clase natural significan primariamente ideas y, constituyan éstas o no una clase nominal, las ideas en Locke tienen el carácter de entidades cognitivamente accesibles necesariamente: son algo así como impresiones o *qualia* que, en condiciones normales, las cosas del mundo producen en nosotros.

Como he dicho, para Locke, los términos de clase natural tienen también una significación secundaria: las entidades extramentales que son la causa de nuestras ideas. En condiciones normales de percepción, la idea de brillo, que es un componente de la idea compleja que constituye el significado primario de ‘oro’, ha sido causada por el propio brillo que presenta el oro. Y lo mismo para el resto de ideas hasta configurar la idea compleja de oro que ha sido causada por una serie de propiedades perceptibles del oro. Esta idea compleja constituye el significado primario de ‘oro’, pero en virtud de la relación causal entre las ideas y las propiedades de las cosas, los términos de clase natural significan de manera derivada esas propiedades extramentales. Es controvertido si esas propiedades objetivas de las cosas constituyen o no lo que Locke llama ‘esencia nominal’, pero lo que parece claro es que en ningún caso el significado de ‘oro’ es la

esencia real del oro. Aunque Locke reconoce que tendemos a creer que sí, quizás porque la esencia real es, a su vez, la (supuesta) causa de las propiedades superficiales de las cosas que nos causan ideas.

[...] they [Men] have *supposed a real Essence belonging to every Species*, from which these Properties all flow, and would have their name of the Species stand for that. But they not having any *Idea* of that real Essence in Substances, and their words signifying nothing but the *Ideas* they have, that which is done by this Attempt, is only to put the name or sound, in the place and stead of the thing having that real Essence, without knowing what the real Essence is; and this is that which Men do, when they speak of Species of Things, as supposing them made by Nature, and distinguished by real Essences. (Locke 1975, 469-70)

Locke excluye la esencia real como significado del término porque, partiendo de una concepción cartesiana del conocimiento, no podemos tener certeza que haya una esencia real subyacente que cause las propiedades que sí percibimos. Es sólo una suposición que existe, pero, como no tenemos ideas causadas por la esencia real (pues no es perceptible), no podemos conocerla. Se podría objetar a Locke que la misma falta de certeza se da respecto a las propiedades extramentales que sí percibimos y que Locke admite que son la significación secundaria de los términos de género natural.

Ciertamente, sólo conocemos de manera infalible nuestras propias ideas. Una posible respuesta coherente con la teoría de Locke sería decir que el conocimiento (falible) que tenemos del mundo es derivado del conocimiento (absolutamente cierto) de nuestras ideas, pero que, en condiciones normales, no parece *demasiado* arriesgado suponer que nuestras ideas han sido causadas por objetos externos percibidos; en cambio la inferencia que va de las propiedades perceptibles de las cosas a su estructura interna no es tan *segura*⁸.

De todos modos, sea éste o no un gran problema para la teoría lockeana, lo que me interesa de ésta es que el significado primario de nuestros términos de clase natural son nuestras ideas (acerca de las respectivas clases), lo único que conocemos con certeza. Naturalmente, esta exigencia de certeza para el conocimiento también implica que no

⁸ Ésta, al menos, es la solución interpretativa que propone García-Carpintero (1996) y que sigo aquí.

podemos conocer las ideas de los otros. Las ideas de cada uno son privadas, y cada individuo tiene cierta libertad sobre ellas: la libertad de tener las ideas que quiera y la libertad de nombrarlas como quiera. Es sólo una suposición que nuestras palabras representan lo mismo que las palabras de otro; no podemos estar del todo seguros de que es así. Y a pesar de que las ideas de los diferentes individuos son muy similares y, generalmente, son causadas por propiedades del mundo *externo*, no hay ninguna garantía de que esto sea así.

Locke no es propiamente un filósofo descriptivista, pues Locke no habla de descripciones accesibles que los hablantes asocian con los términos de clase natural. Los significados primarios de las palabras, las ideas, son directamente accesibles e infalibles, algo como los *qualia*⁹, pero, en principio, no tienen un carácter descriptivo. Y las esencias nominales, sean ideas complejas o sean propiedades extramentales, tampoco tienen un carácter descriptivo. Otra cosa es que a menudo podamos describir una esencia nominal, pues se trata de describir propiedades sensibles. Sin embargo, Locke es un ejemplo paradigmático de filósofo internista, al menos para los términos de clase natural, pues el significado de tales términos viene dado por algo cognitivamente disponible, accesible al hablante competente, sea o no especificable mediante contenido descriptivo: las propias ideas y, en segunda instancia, las propiedades perceptibles de los miembros de la clase natural (cuando sean causa de tales ideas).

Cabe remarcar que el hecho que las propiedades asociadas con las clases naturales sean conjuntos de características superficiales o perceptibles supone un problema para la teoría de Locke. Como internista, necesita que las propiedades de las cosas que determinan la extensión de los términos de clase natural sean accesibles cognitivamente y la esencia real, dice Locke, no es algo cognitivamente accesible, no puede causarnos ideas. Pero si la teoría de Locke fuera la teoría correcta para los términos de clase natural, resulta que no podemos decir que, a veces, en el pasado, las personas se equivocaban cuando, en condiciones normales¹⁰, aplicaban ‘oro’ a algo con toda la

⁹ Al menos, las ideas *simples*, las más básicas, pueden asimilarse a *sensaciones* o *experiencias*, algo como los *qualia*.

¹⁰ Es decir, cuando sus ideas eran causadas por propiedades existentes de las cosas.

aparición del oro, algo que tenía la esencia nominal del oro, pero que, en realidad, no era oro, sino otra sustancia perceptiblemente indistinguible como la pirita de hierro. Asumiendo los puntos de vista de Locke, tampoco podemos corregir a alguien que decide usar los términos según sus ideas de las cosas. En resumen, una consecuencia contraintuitiva de su teoría es que, como basa la correcta aplicación de los términos en la competencia lingüística de cada hablante, no hay un criterio *objetivo* para corregir la aplicación de otro hablante competente. En último término, cada hablante competente tiene un criterio propio subjetivo de aplicación del término que viene determinado por las ideas que posee.

Men must suppose the same Word to signify different Things in different Men: since they cannot doubt, but different Men may have discovered several Qualities in Substances of the same Denomination, which others know nothing of. (Locke 1975, 469)

Las ideas propias que determinan el significado de nuestras palabras son del todo internas y *sólo* accesibles al poseedor de tales ideas. Cada hablante, entonces, habla su propio idiolecto.

1.3.2. El internismo de Mill para los términos generales

Mill es un autor que, en 1843, también trata explícitamente la semántica de los términos de clase natural, así como de los nombres propios. En general, hace una distinción entre dos componentes semánticos de los términos: la denotación y la connotación, que pueden considerarse análogos al referente y el sentido fregeanos, respectivamente. Para Mill, los nombres propios son términos excepcionales de nuestro lenguaje y sólo presentan denotación, no connotan nada, y, al carecer de uno de los componentes semánticos básicos, estrictamente hablando no tienen significado. Simplemente *están por* la cosa nombrada, pero carecen de un significado pleno. Son como marcas que mentalmente conectamos con el objeto (o la idea del objeto) correspondiente. Este punto de vista acerca de los nombres propios es llamado ‘millianismo’ o, a veces, ‘teoría ‘Pluto’ - Pluto’¹¹.

¹¹ Frege reacciona contra esto y presenta su teoría dual del significado también para los nombres propios, expuesta en la sección 1.1.1.

Respecto a los términos de clase natural, y como el resto de términos generales, al tratarse de términos diferentes a los nombres propios, Mill dice que tienen un significado pleno: presentan denotación y connotación.

All concrete general names are connotative. The word *man*, for example, denotes Peter, Paul, John, and an indefinite number of other individuals, of whom, taken as a class, it is the name. But it is applied to them, because they possess, and to signify that they possess, certain attributes. These seem to be, corporeity, animal life, rationality, and a certain external form, which for distinction we call the human. (Mill 1973, 31)

La denotación de un término general se corresponde con su extensión, y la connotación, con lo que podríamos considerar su intensión. Cada término general, como los términos para clases naturales, se aplica a una serie de individuos o muestras en virtud de que poseen ciertas características comunes. Estas características comunes son las connotadas por el propio término de clase natural.

The word *man*, therefore, signifies all these attributes, and all subjects which possess these attributes. (Mill 1973, 32)

Para Mill, los términos de clase natural significan *directamente* los particulares denotados e *indirectamente* las propiedades o atributos que son connotadas por los propios términos. La serie de atributos connotados por el término son conocidos por los hablantes competentes por el mero hecho de serlo, pues la idea de Mill es que el propio término ya connota: a diferencia de un nombre propio, ya proporciona alguna información (que automáticamente asociamos) sobre los atributos característicos de los miembros de la clase. Pero Mill reconoce que, a menudo, las connotaciones asociadas a los términos generales pueden ser ambiguas o imprecisas; es el caso de términos generales como ‘gente’. De la misma manera, pensar en las connotaciones como definiciones de los términos que especifican los atributos característicos de una clase no siempre funciona, ya que, según Mill, no siempre disponemos de todas las especificaciones concretas que caracterizan una clase natural.

Thus, it is clear that the word *man*, besides animal life and rationality, connotes also a certain external form; but it would be impossible to say

precisely what form; that is, to decide how great a deviation from the form ordinarily found in the beings whom we are accustomed to call men, would suffice in a newly-discovered race to make us to refuse them the name of man. (Mill 1973, 37)

Y lo mismo para otras características connotadas por los términos generales (entre ellos, los términos de clase natural). Las connotaciones son, muchas veces, algo vagas como para intentar explicitarlas en definiciones precisas.

Por tanto, aunque habitualmente se ha interpretado la teoría de Mill como una teoría descriptivista para los términos de clase natural, Mill ve dificultades en equiparar su significado con el de alguna descripción o definición. De todos modos, para Mill, la parte determinante del significado de los términos connotativos es su connotación:

[...] whenever the names given to objects convey any information, that is, whenever they have properly any meaning, the meaning resides not in what they *denote*, but in what they *connote*. (Mill 1973, 34)

Así que, dejando de lado si su propuesta es estrictamente descriptivista, Mill ofrece una teoría internista, ya que la connotación presente en los términos generales, aunque más o menos vaga, es cognitivamente accesible a los hablantes que emplean competentemente los términos.

Para entender cómo es posible una teoría internista que no sea descriptivista, es útil quizás considerar algún ejemplo de un término que sepamos aplicar correctamente sin saber dar ninguna descripción de las características que sabemos identificar. Así, por ejemplo, consideremos el término general 'rojo'. Desde una postura internista no descriptivista, se podría decir que los hablantes tienen acceso cognitivo a las características distintivas de las cosas rojas y aplican correctamente el término aun cuando no puedan especificar mediante una descripción en qué consiste ser rojo. Lo importante es que la rojez es algo cognitivamente accesible a los hablantes. Así, posturas internistas como la de Mill, tienen sentido, pues la idea internista no descriptivista es que tener acceso cognitivo a ciertas características, poder reconocerlas, no es lo mismo que ser capaz de describirlas.

1.3.3. Las teorías descriptivistas modificadas aplicadas al caso de los términos de clase natural

La teoría cúmulo descriptivista, inicialmente concebida para los nombres propios, puede ampliarse sin problemas a los términos generales, en particular, a los términos de género natural, los términos generales más estudiados. De nuevo, la motivación para un descriptivismo *social* como el que propone la teoría cúmulo es que un descriptivismo clásico en el que cada hablante asocie a cada término de clase natural una descripción que seleccione la extensión del término presenta problemas. Ciertamente, un descriptivismo clásico para los términos de clase natural puede presentar uno de los problemas aludidos por Searle: diferentes personas pueden asociar diferentes descripciones a un término de clase natural y, entonces, queda indeterminada cuál es, entre todas las descripciones asociadas, la que determina la referencia del término. Pero no está tan claro, de entrada, que el defecto que ve Searle al descriptivismo clásico de los nombres propios, consistente en que el contenido descriptivo asociado al término tendría que predicarse necesariamente del referente del término, sea un defecto obvio en el caso del descriptivismo clásico aplicado a los términos de clase natural. Simplemente, no es tan intuitivo que, si la descripción asociada con ‘oro’ es ‘metal brillante, amarillo y duro’, ‘el oro es un metal brillante amarillo y duro’ no sea una verdad necesaria como es claramente intuitivo que ‘Aristóteles es el autor de la *Metafísica*’ no es una afirmación necesariamente verdadera. El brillo, la amarillez y la dureza del oro podrían parecer, de entrada, propiedades no contingentes del oro en contraste con la propiedad de haber sido el autor de la *Metafísica*, propiedad claramente contingente de Aristóteles¹². Más adelante, en la sección 1.5.1, expongo las consideraciones de Kripke acerca de esta cuestión.

Pero un descriptivismo cúmulo para los términos de clase natural permite que algunos hablantes competentes no dispongan de una descripción que selecciona el referente o el ámbito de aplicación del término general. En una teoría como la de Searle, aplicada a los términos de clase natural, bastaría que hubiese un cúmulo social de descripciones, la

¹² Una de las razones por las que parecería que el oro es necesariamente un metal brillante, amarillo y duro es que consideramos que éstas son propiedades ligadas causalmente a la estructura interna del oro. Esto contrasta, de entrada, con el tipo de propiedades que el descriptivismo clásico típicamente atribuye al portador de un nombre propio.

suma o disyunción de todas las descripciones asociadas por los hablantes, para que se determinase la extensión del término mediante *alguna* o *algunas* de esas descripciones. Naturalmente, en el cúmulo de descripciones están también las descripciones que asocian los expertos en cada género natural al correspondiente término. De nuevo, como en el caso del cúmulo de descripciones asociado a un nombre propio, la idea es que no puede suceder que todas las descripciones del cúmulo sean falsas o no seleccionen la extensión del término de clase natural. Un número *ponderado* de descripciones, presumiblemente las de los expertos, tienen que ser verdaderas predicadas de los miembros de la clase, y seleccionarán el ámbito de aplicación del término. Los usos de los hablantes, también de los no expertos, consiguen referir porque remiten al cúmulo social de descripciones asociadas al término general.

Igualmente, la versión descriptivista de fijación de la referencia puede ser adaptada al caso de los términos de clase natural para superar los inconvenientes del descriptivismo clásico que considera como sinónimos el significado del término y el de la descripción asociada a él. Para esta propuesta, el significado de ‘oro’ no es sinónimo del significado de ‘metal brillante, amarillo y duro’ (u otra descripción similar); la descripción asociada no proporciona necesariamente el significado del término. Pero hay alguna descripción asociada al término que fija la referencia del término en cada uso de éste. Esta descripción que fija la referencia puede ser una descripción que recoja las características superficiales que presentan las muestras de la clase. Pero, finalmente, tras las críticas de Kripke al descriptivismo (que discutiré en las secciones **1.4** y **1.5**), que también afectan a esta versión, algunos partidarios de un descriptivismo de fijación de la referencia acaban proponiendo un descriptivismo causal según el cual la descripción que fija la referencia de un término de clase natural como ‘oro’ es algo así como ‘la sustancia referida por aquéllos de quienes tomo la referencia’, y no una descripción de las características perceptibles del oro.

1.4. La crítica de la nueva teoría de la referencia al internismo descriptivista clásico y a la teoría cúmulo para los nombres propios

1.4.1. La crítica kripkeana al descriptivismo tradicional

Kripke hace una revisión y una crítica del descriptivismo aplicado a los nombres propios que engloba las diferentes modalidades del descriptivismo: el descriptivismo clásico en la versión *Frege-Russell*¹³, la versión cúmulo de los nombres propios y el descriptivismo débil o de fijación de la referencia. Kripke caracteriza primero una teoría descriptivista aplicada a los nombres propios en cuatro tesis. Dos de ellas son de tipo semántico, otra es de tipo epistémico y la última, de tipo modal. Estas tesis características del descriptivismo son:

- (i) Los hablantes creen que una o algunas de las propiedades de un cúmulo F seleccionan un único individuo como la única entidad que las posee.
- (ii) Para todo objeto x , un nombre propio refiere a x si y sólo si x es el único objeto que satisface F (satisfacer F es poseer la mayoría o una mayoría cualificada de las propiedades de F).
- (iii) Si 'N' es un nombre propio, entonces el enunciado 'Si N existe, entonces N satisface F' es conocido *a priori* por los hablantes.
- (iv) El enunciado 'Si N existe, entonces N satisface F' expresa una verdad necesaria (donde 'N' es un nombre propio del lenguaje).

La objeción de Kripke a estas tesis es que son falsas. Contra la tesis (i), Kripke dice que habitualmente los hablantes usamos nombres propios sin creer que disponemos de descripciones (una o diversas) que individualicen su referente. Por ejemplo, dice Kripke, podemos usar el nombre 'Feynman' para referir a Feynman aunque sólo sepamos de su referente que es un físico. Es decir, podemos referir con éxito al usar un nombre propio aunque ignoremos la propiedad o propiedades que individualizan al referente del nombre. Éste es el llamado 'argumento de la ignorancia' que propone

¹³ De nuevo, la interpretación que Kripke hace de la obra de Frege es en clave descriptivista, asumiendo que, para Frege, el significado de un nombre propio es sinónimo al significado de una descripción definida. Por otro lado, no está del todo claro que el descriptivismo en versión cúmulo que Kripke caracteriza se corresponda exactamente con el defendido por Searle.

Kripke contra el descriptivismo. Contra la tesis (ii), Kripke dice que no sólo el bicondicional que establece la tesis es falso, sino que los dos sentidos del bicondicional son falsos. Por un lado, la tesis dice que, para todo objeto x , si un nombre propio refiere a x , entonces x es el único objeto que satisface la mayoría de propiedades de un cúmulo F asociado al nombre. Para rebatir esta afirmación, Kripke pide que imaginemos que, en vez de Gödel, hubiese sido Schmidt el genuino pero desconocido descubridor de la incompletud de la aritmética y que Gödel se hubiera limitado a plagiarlo haciendo pasar con éxito los trabajos de Schmidt como propios. En este caso, ‘Gödel’ refiere a Gödel, pero no es verdad que Gödel sea el objeto que satisface *ser el descubridor de la incompletud de la aritmética* (Kripke supone, plausiblemente, que esta descripción es prácticamente la única que asociamos a ‘Gödel’). De hecho, este argumento de Kripke parece desacreditar sobretudo el descriptivismo clásico, y quizás no tan claramente la teoría cúmulo. Si fuera verdad que en nuestra comunidad lingüística hay muchas más descripciones asociadas a ‘Gödel’, entonces sería más difícil construir un contraejemplo a (ii)).

Por otro lado, en el otro sentido del bicondicional, (ii) afirma que si x es el único objeto que satisface la mayoría de descripciones del cúmulo, entonces el nombre propio asociado a las descripciones refiere a x . Y esto también es falso, dice Kripke, porque en la situación imaginada, Schmidt es el único objeto que satisface *ser el descubridor de la incompletud de la aritmética*, pero ‘Gödel’ (el nombre asociado a la descripción) no refiere a Schmidt en la situación ficticia, sino a Gödel. (De nuevo, Kripke sólo contempla una descripción: ‘el descubridor de la incompletud de la aritmética’). Este contraejemplo de Kripke a la tesis (ii) es el argumento del error: podemos referir al usar un nombre propio aun cuando la descripción singularizadora que asociamos al nombre sea equivocada. Hay que notar que, para Searle, si todas las descripciones asociadas a ‘Gödel’ en nuestra comunidad lingüística expresasen propiedades que de hecho no tuvo Gödel, entonces ‘Gödel’ no referiría, sería un nombre sin sentido. Y esto no parece ser así.

Contra la tesis (iii), Kripke dice que, incluso en casos en que se dan (i) y (ii), nuestro conocimiento que el referente satisface las correspondientes descripciones es *a posteriori*. Ciertamente, sabemos *a posteriori* que Aristóteles es al autor de *La Metafísica* (y lo mismo respecto a todas o casi todas las descripciones que se puedan

asociar a ‘Aristóteles’). Y, por último, contra la tesis (iv), Kripke argumenta que, aunque se den las condiciones (i), (ii) y (iii), la última condición, (iv), podría ser falsa. Para ver esto, consideremos el caso en que estipulamos que llamaremos ‘Julius’ a quien quiera que sea que inventó la cremallera. En virtud de esta estipulación, se puede defender que se cumplen (i), (ii) y (iii). Pero, sin embargo, (iv) no se cumple porque ‘Julius’ es un designador rígido (designa a la misma entidad en todo mundo posible en que dicha entidad existe) y, en cambio, ‘el inventor de la cremallera’ no es un designador rígido (no designa al mismo objeto en todo mundo posible). Y, así, ‘Si Julius existe, Julius inventó la cremallera’ no es una verdad necesaria pues hay mundos posibles en que la oración es falsa porque, por ejemplo, otra persona antes que Julius inventó la cremallera.

Respecto a las diferentes objeciones que Kripke hace al descriptivismo, quiero destacar que, desde mi punto de vista, las críticas más importantes son las críticas a las tesis (i) y (ii). La razón es que las tesis (iii) y (iv) no son suscritas por toda forma de descriptivismo; en particular, no son suscritas por un descriptivismo *débil* de fijación de la referencia según el cual el término y su descripción asociada no tienen significados sinónimos, sino que la descripción asociada *sólo* sirve para fijar el referente del término sin que esté ligada necesariamente y *a priori* al nombre propio¹⁴. Entonces, los contraejemplos de Kripke a estas tesis no se aplican al citado tipo de descriptivismo. En cambio, los argumentos kripkeanos contra (i) y (ii), llamados a veces ‘argumentos semánticos’, pueden aplicarse, en principio, a todas las versiones de descriptivismo¹⁵. En definitiva, estos argumentos de tipo semántico, los argumentos de ignorancia y error, son argumentos con un alcance más amplio contra el descriptivismo.

Por otro lado, dada la plausibilidad inicial -basada sobre todo en el principio de transparencia- que presenta el internismo semántico, creo que la carga de prueba recae del lado del externismo para intentar mostrar que el internismo no es la teoría adecuada para explicar cómo funcionan nuestros términos. Y tal como he caracterizado el

¹⁴ O, al menos, el descriptivismo de fijación de la referencia no se compromete con (iv). Seguramente tampoco con (iii), aunque esto último quizá podría considerarse dudoso.

¹⁵ Excepto a un tipo de descriptivismo, el descriptivismo causal, que, como veremos en la sección 1.4.1, puede refutarse con argumentos independientes.

externismo y el internismo semánticos¹⁶, la viabilidad del externismo depende crucialmente, desde mi punto de vista, de la posibilidad de construir argumentos de ignorancia y error contra el descriptivismo como los que propone Kripke al argumentar contra (i) y (ii). Pues si tales argumentos son posibles y el externista puede mostrar que es posible que los hablantes asocien con un término material descriptivo equivocado o insuficiente para seleccionar su referente, entonces se mostraría que los hablantes no tienen acceso cognitivo a una descripción que determine la referencia del término, aunque intuitivamente sus usos del término en cuestión consigan referir. Creo que, antes de proponer una teoría positiva acerca de cómo y a qué consiguen referir los términos de nuestro lenguaje, el externismo tiene que mostrar que es viable porque *podría* haber casos en los que el descriptivismo no se sostuviera. Y considero que esta argumentación pasa por la construcción de argumentos como los que propone Kripke con ‘Feynman’ y ‘Gödel’ para mostrar que el descriptivismo no da cuenta de cómo consiguen referir los nombres propios. Igualmente, por los mismos motivos, considero que, en el caso de los términos generales, que será tratado más adelante, los argumentos decisivos para el externista en su disputa con el descriptivista son también los argumentos de ignorancia y error respecto a las características relevantes de las clases en cuestión. De ahora en adelante, me centraré en esta parte de la crítica kripkeana, y dejaré de lado su crítica contra las tesis descriptivistas (iii) y (iv) por las consideraciones expuestas.

1.4.2. La crítica kripkeana a la teoría cúmulo y al descriptivismo de fijación de la referencia

Los contraejemplos a las tesis (i) y (ii) ofrecidos por Kripke se aplican claramente a la teoría descriptivista clásica de los nombres propios, y Kripke piensa que también a una teoría descriptivista en versión cúmulo. Quizás, para aplicarlos a una teoría cúmulo de los nombres propios, los contraejemplos tendrían que ser un poco más elaborados. Para Kripke, uno de los puntos débiles de una teoría cúmulo como la de Searle es que asume que nunca se da el caso que todas las descripciones asociadas con un nombre propio resultan ser predicaciones falsas del referente. Según Searle es necesario que, por

¹⁶ Según mi caracterización, el internismo sobre la determinación de la referencia es la tesis que defiende que necesariamente tenemos un acceso cognitivo a la entidad que determina la referencia de nuestros términos (el descriptivismo es el caso particular que defiende que la entidad que determina la referencia y a la que accedemos cognitivamente es una descripción). El externismo es la negación de tal tesis.

ejemplo, Aristóteles tenga un número suficiente de las propiedades comúnmente atribuidas a él. Es una verdad necesaria sobre Aristóteles que al menos una de las descripciones en el cúmulo asociado a él en este mundo, sea también verdad de Aristóteles en todo mundo posible en que existe. Así, necesariamente Aristóteles tiene al menos una de las propiedades habitualmente atribuidas a él en este mundo en cada mundo posible (aunque la propiedad que tiene puede ser diferente en los diferentes mundos posibles). Pero esto es lo que Kripke rechaza: Aristóteles podría haber existido sin poseer ninguna de las propiedades que son habitualmente asociadas a él en este mundo. O poniendo otro ejemplo, parece posible descubrir, dice Kripke, que todas las descripciones del cúmulo que asociamos con ‘Colón’ resultasen ser falsas aplicadas a Colón, y el nombre parece que continuaría refiriendo.

Quizás se podría pensar que el hecho que *todas* las descripciones del cúmulo puedan resultar erróneas o falsas del referente no es tan plausible como quisiera Kripke. De todas formas, no hace falta que todas sean falsas para que la teoría cúmulo no dé cuenta de la relación de referencia entre el nombre con sus descripciones asociadas y el referente. En los argumentos expuestos de Kripke contra las tesis (i) y (ii), argumentos basados en la ignorancia y el error de los hablantes, ya se encuentra la clave para intentar refutar una teoría cúmulo como la de Searle aplicada a los nombres propios. Aunque no todas las descripciones del cúmulo social sean (o puedan ser) erróneas, basta que algunas, quizás en un número *suficiente*, lo sean para conducir a un referente *equivocado*. De manera similar al caso de ‘Gödel’, si un número suficiente de descripciones del cúmulo asociado al nombre propio son erróneas, el cúmulo podría establecer la relación de referencia con otro referente. Igualmente, también parece posible que haya nombres propios para los que el cúmulo social no dispone de ninguna descripción singularizadora. Podría haber un caso como el de ‘Feynman’ ampliable a toda la comunidad. Entonces, el cúmulo sería insuficiente para determinar el referente, dándose un caso de ignorancia por parte de la comunidad lingüística.

Kripke también incluye en su crítica al descriptivismo débil o de fijación de la referencia. Obviamente, si, a veces, la descripción que los hablantes asocian con un nombre propio es insuficiente para seleccionar el referente del nombre o selecciona un referente *equivocado*, entonces tampoco se trata de una descripción que meramente fije el referente (aunque no dé el significado del nombre). En general, los nombres propios

comunes no tienen sus referentes semánticamente fijados por una descripción¹⁷. La conclusión de todo esto es que los nombres propios no funcionan como las descripciones definidas y no pueden ser equivalentes en significado al significado de una o diversas descripciones. Tampoco es verdad que haya siempre una descripción disponible asociada con cada nombre propio que, al menos, fije su referente.

El problema de fondo del descriptivismo, en cualquiera de sus versiones, es que parece que podemos usar un término y referir con él a pesar de ignorar o estar equivocados sobre lo que individualiza al referente. Una posible salida para Searle que no tiene el internismo clásico es decir que *cualquier* hablante puede ser ignorante o estar equivocado sobre las descripciones que seleccionan el referente. Pero, en el caso de una teoría cúmulo el problema se presenta cuando todos o la mayoría en la comunidad lingüística podemos ser ignorantes o estar equivocados sobre cómo individualizar el referente. Y esto es algo que parece plausible. En general, la crítica de Kripke al descriptivismo aplicado a los nombres propios ha sido aceptada por muchos autores, que han descartado una semántica de este tipo para los nombres propios, tanto en su versión tradicional como en la versión modificada de Searle.

A raíz de la crítica de Kripke, que también se extiende al descriptivismo en su versión *débil*, surgió una adaptación del descriptivismo de fijación de la referencia. Esta modificación de la teoría, que dio origen al llamado ‘descriptivismo causal’, propone que la descripción que fija la referencia del nombre propio, antes que ser una descripción que especifica algún hecho notable o característico del referente, es una descripción del tipo *la cosa referida por aquéllos de los que el presente uso del término toma prestada la referencia*. Lewis, tal como recoge Soames (2005), propone que el material descriptivo que los hablantes asocian con un nombre como, por ejemplo, ‘Londres’, podría ser ‘el lugar del que he oído hablar bajo el nombre de ‘Londres’ o ‘la fuente causal del ejemplar Londres’. Así, mediante una descripción que apela al proceso por el que el nombre adquirió su significado¹⁸, el uso del término ‘Londres’ por

¹⁷ Hay algunas excepciones, como, por ejemplo, ‘Jack el destripador’.

¹⁸ El descriptivismo causal extrae una descripción de la explicación histórico-causal que propone Kripke de cómo se transmite la referencia de los nombres. En la sección 1.5.3 expongo dicha propuesta de Kripke.

parte de los hablantes consigue alcanzar su referente, Londres. De este modo, la teoría descriptivista causal se asegura que siempre habrá una descripción disponible que singularice el referente del nombre (sin que esa descripción sea sinónima al nombre) y salva los problemas de la ignorancia y el error que afectan a las otras versiones descriptivistas.

Sin embargo, esta modificación de la teoría descriptivista no está libre de problemas. Soames hace una crítica a esta versión modificada notando que una descripción asociada al nombre que describe el proceso causal que se origina en el referente no puede ser el mecanismo que determina el referente, pues tal explicación ya asume que hay otro mecanismo de transmisión de la referencia. La propia descripción asociada al término explicita que la referencia se toma prestada de los usos previos de otros hablantes mediante una cadena causal. Sin este proceso previo, independiente, de transmisión de la referencia, la descripción no determina cómo se fija la referencia, ni consigue designar nada. Más bien, describe el proceso por el que un uso de un nombre propio consigue referir. Es cierto que tal descripción sí selecciona el referente del nombre, pero una vez que la referencia ya ha sido determinada por el proceso histórico causal. Si tal proceso histórico no existiera, es obvio que la descripción no conseguiría llegar al referente. Por otro lado, tal como dice Soames, es inverosímil que los hablantes al usar un nombre propio asocien de alguna manera una descripción que hace explícito cómo el uso de ese nombre consigue referir. Naturalmente, los hablantes no tienen por qué saber nada acerca de cómo se determina la referencia de los términos que usan; su uso de los términos es independiente de sus conocimientos semánticos.

1.5. La crítica externista al internismo descriptivista clásico y a la teoría cúmulo para los términos de clase natural

1.5.1 La crítica de Kripke

Kripke y otros autores de la nueva teoría de la referencia ponen de relieve que los hablantes pueden usar un nombre propio y referir con él a pesar de que la descripción o descripciones asociadas al término sean equivocadas o del todo insuficientes para seleccionar al individuo portador del nombre, desacreditando la teoría descriptivista, en su versión clásica o modificada. Es decir, el internismo semántico para los nombres

propios es susceptible de sufrir los problemas de la ignorancia y del error. Pero los argumentos ofrecidos por Kripke también son aplicables al descriptivismo para los términos generales. En particular, los mismos problemas de la ignorancia y el error pueden darse también en la explicación descriptivista de los términos de clase natural. Un hablante competente puede usar un término de clase aunque desconozca una descripción de la clase en cuestión que sólo seleccione la extensión del término. Y también puede pasar que un hablante asocie con el término una descripción que no es satisfecha por las muestras que caen en la extensión del término. Pero, aunque un hablante competente pueda ser ignorante o estar equivocado sobre lo que define una clase, puede referir con éxito cuando usa un término general. Kripke propone el ejemplo de la ilusión colectiva respecto al oro: supongamos que descubrimos que, debido a un efecto óptico causado por las condiciones atmosféricas, el oro nos parece de color amarillo cuando realmente, una vez eliminadas esas condiciones atmosféricas, es azul. En tal situación, dice Kripke, diríamos que, aunque parecía que el oro era amarillo, de hecho ha resultado ser de color azul. Esto muestra que al usar ‘oro’ queremos referirnos a una cierta clase de sustancia, no a algo que tiene que cumplir ciertas propiedades que le atribuimos (pues podríamos equivocarnos en la atribución de esas propiedades). También podría pasar que hubiese una sustancia con todos los rasgos identificadores que comúnmente atribuimos al oro, pero que fuese otra sustancia, pirita de hierro. Diríamos, entonces, que tal sustancia parece oro pero no es oro. Así, el significado de ‘oro’ no puede venir dado por una descripción que recoge una serie de características atribuidas al oro, pues esa descripción podría no seleccionar la extensión de ‘oro’ (bien por error en nuestra descripción, bien por insuficiencia de la descripción para seleccionar solamente el ámbito de aplicación del término).

Kripke proporciona algún otro ejemplo de argumento de ignorancia y error. Supongamos que un hablante competente del lenguaje asocia con ‘tigre’ una descripción como ‘felino grande de cuatro patas, carnívoro, de color amarillo tostado con rayas negras transversales’¹⁹. De nuevo, podríamos descubrir, por ejemplo, que los

¹⁹ En la definición propuesta, aclara Kripke, el término ‘felino’ debe entenderse como *con apariencia de gato grande*, no como designando una familia biológica determinada, pues el hecho de que los tigres pertenezcan a la familia de los felinos es algo que hemos tenido que descubrir, no algo que asociemos con ‘tigre’ por el mero procedimiento de observar las características superficiales de los tigres.

tigres no tienen cuatro patas, sino tres. Igual que es concebible un tigre con tres patas, es concebible que *ningún* tigre tenga de hecho cuatro patas; podríamos haber sufrido una ilusión óptica que nos hiciera ver animales cuadrúpedos donde había unos animales con tres patas. No hay contradicción, dice Kripke, en suponer que los tigres podrían tener tres patas. En tal situación, diríamos que hemos descubierto que los tigres tienen realmente tres patas. Además, satisfacer una descripción como la anterior no es necesario ni suficiente para que algo sea un tigre. No es necesario que algo tenga las características que atribuimos a los tigres para ser un tigre: es concebible que, debido a alguna alucinación colectiva los tigres no sean amarillos con rayas negras o, como se ha dicho, no tengan cuatro patas. Por otro lado, para que algo sea un tigre no es suficiente que sea similar superficialmente a las entidades que llamamos ‘tigres’: es concebible algo con todo el aspecto de un tigre pero que realmente fuera un robot. Del mismo modo, algo podría tener toda la apariencia del oro y ser pirita de hierro. En definitiva, una descripción asociada con ‘tigre’ como la apuntada anteriormente no puede dar el significado de ‘tigre’.

Estos problemas de la ignorancia y el error también pueden afectar a una teoría cúmulo adaptada al caso de los términos de clase natural. Quizá se podría pensar que algo que es un tigre podría no cumplir *alguna* característica que generalmente atribuimos a los tigres, como el ser cuadrúpedo o ser de color amarillo con rayas negras, pero no podría incumplir *todas* las características que atribuimos a los tigres. Kripke piensa que no es así:

[...] so we might also find out tigers had *none* of the properties by which we originally identified them. Perhaps *none* are quadrupedal, none tawny yellow, none carnivorous, and so on; all these properties turn out to be based on optical illusions or other errors, as in the case of gold. So the term ‘tiger’, like the term ‘gold’, does *not* mark out a ‘cluster concept’ in which most, but perhaps not all, of the properties used to identify the kina must be satisfied. On the contrary, possession of most of these properties need not be a necessary condition for membership in the kind, nor need it be a sufficient condition. (Kripke 1988, 121)

Así, el cúmulo de descripciones asociado a un término de clase natural podría resultar

falso porque *todas o la mayoría* de las descripciones contenidas en él podrían ser falsas. Tal vez se podría pensar que esta posibilidad es plausible en el caso de ‘tigre’ porque, a fin de cuentas, quizá las características superficiales que asociamos a los tigres, como las que menciona Kripke, no son, consideradas por separado, esenciales para que algo sea un tigre. Claramente, un tigre puede tener tres patas o presentar un problema de coloración en la piel o, incluso, no ser fundamentalmente carnívoro, y no por eso dejaría de ser un tigre. Y si ninguna de las características que asociamos a los tigres es, al menos considerada aisladamente, esencial para que algo sea un tigre, quizá parece admisible una situación en que *ninguna* de esas características está presente en *ningún* tigre. En cambio, se podría pensar que en los casos de las clases de sustancias químicas, como el oro, muchas o, al menos, algunas de las características que asociamos son propiedades necesarias de los miembros de la clase en tanto que tales. Así, por ejemplo, si con ‘oro’ asociamos una descripción del tipo ‘metal amarillo, duro y brillante’, nos parece que estas características (algunas, o, quizás, todas) son necesarias del oro en la medida en que dependen causalmente de la estructura interna del oro. Algo que no tuviera estas características, no podría ser oro. Nos podría parecer inverosímil que el oro no fuera duro o brillante o metálico, por ser características necesarias para que algo sea oro. Entonces, el hecho que *todas o la mayoría* de las descripciones del cúmulo asociado con ‘oro’ resultasen ser falsas puede parecer, de entrada, poco plausible.

Kripke es consciente de la dificultad de plantear situaciones contrafácticas respecto a una clase natural cuando al respectivo término asociamos una descripción que incluye características necesarias para que los miembros pertenezcan a la clase. Naturalmente, el problema no afecta sólo a las clases químicas, pues, a veces, también podemos asociar a las especies biológicas propiedades necesarias tales como ser un animal, ser un ave, ser mamífero, etc. Kripke, en su caracterización estándar de los tigres, elude mencionar este tipo de características pero es consciente de que podríamos incluir características como *animal* o *felino* (en un sentido biológico). Y si la descripción asociada con un término de clase natural expresa propiedades necesarias para que un objeto caiga en la extensión del término, sean propiedades observables (como *brillante*) o no lo sean (como *metal*), podría parecer ilegítimo plantear una situación contrafáctica en la que los miembros de esa clase no tienen esas propiedades necesarias. La clave, en estos casos de asociación de propiedades que nos parecen necesarias, está, creo, en ver la diferencia que hay entre *la posibilidad de descubrir* que el oro, o cualquier otro

género natural, no tiene ninguna de las características que le atribuimos y *la posibilidad que suceda* que el oro no tenga ninguna de las características que le atribuimos (suponiendo que algunas de esas características son necesarias del oro). Para entender mejor esta diferencia quizá es útil considerar el experimento mental de los gatos-robots propuesto por Putnam (1975). Aquí se trata de un término de especie biológica con el que, según Putnam, asociamos una descripción que hace referencia a la propiedad de ser animales. Podemos suponer, dice Putnam, que descubrimos que todos los individuos que considerábamos gatos son, en realidad, robots enviados desde Marte, y no unos determinados animales, tal como habíamos supuesto. En tal situación, dice Putnam, descubriríamos que los gatos son robots, no que no hay gatos. Si el argumento es plausible, muestra que todos podríamos estar equivocados sobre la naturaleza de los gatos y asociar con 'gato' una descripción que no es verdadera de los gatos y, sin embargo, con 'gato' habríamos estado refiriendo a esas entidades que ahora sabemos que son robots. Entonces, la descripción que asociamos con el término 'gato, y que incluiría alguna referencia a la propiedad de ser animal, no puede determinar la referencia de 'gato'.

Contra este argumento de Putnam parece que se podría objetar que no podríamos descubrir que los gatos son robots sencillamente porque ser animales es una propiedad esencial de los gatos. Es decir, en la situación descrita por Putnam, nunca ha habido gatos porque éstos, en todo mundo posible, son animales. Sin embargo, Putnam, como Kripke, está de acuerdo con que los gatos, puesto que son animales, son necesariamente animales; un individuo que no fuera un animal no podría ser un gato. Pero la clave, en este experimento mental, es que no es cognoscible *a priori* que los gatos son animales (en sentido biológico). Antes de conocer la estructura interna de los gatos, por todo lo que sabíamos, los gatos podrían haber resultado ser robots. Es esta posibilidad epistémica la que se describe en el experimento mental de los gatos robots. Igualmente, respecto al oro, no es cognoscible *a priori* que el oro es un metal. Respecto a las propiedades necesarias pero cognoscibles *a posteriori* siempre es posible el error o la ignorancia.

Sin embargo, se podría decir que, a veces, asociamos con un género natural, propiedades necesarias pero que, para algunos, parecen cognoscibles *a priori* (o, al menos, no claramente cognoscibles *a posteriori*) como la dureza, brillo y amarillez del

oro. La idea es que, en tanto que hablantes competentes del lenguaje, conocemos estas características del oro sin necesidad de investigación empírica, pues son características observables. Quizá, para algunos, si un hablante no sabe que el oro es amarillo, duro y brillante, no es un usuario competente del término. Pero, incluso dejando de lado si esto es el caso y si tenemos un acceso a ese tipo de características *a priori* o *a posteriori*, la asociación de la característica con la clase, requiera o no investigación empírica, siempre puede ser equivocada. Es decir, aún concediendo que asociamos *a priori* con ‘oro’ una descripción que hace referencia a su amarillez y brillo, esa asociación puede estar equivocada. Como vimos en el anterior argumento kripkeano, podríamos descubrir que hemos sufrido una ilusión colectiva que nos hacía ver el oro amarillo y brillante, cuando, realmente, era azul y mate. De nuevo, afirmar esta posibilidad no equivale a afirmar la posibilidad metafísica que el oro (en la medida en que su color y brillo dependen de su estructura interna) podría haber sido de otro color y haber presentado otro brillo. El tipo de posibilidad contemplada es epistémica: por todo lo que sabíamos (y sabemos) en un momento dado, podríamos haber sufrido una ilusión óptica colectiva que nos hiciera ver unas cualidades en el oro que no posee.

Hay que notar, además, que aun cuando asociemos *a priori* al oro ciertas características necesarias del oro, lo que en ningún caso sabemos *a priori* es que esas características son necesarias del oro. Esto es algo que hay que descubrir. Y, por tanto, por todo lo que sabíamos, antes de conocer que la dureza es una característica necesaria del oro (si es que lo es), el oro podría haber resultado no ser duro. Por otro lado, también hay que destacar que podríamos asociar ciertas características con el oro que son esenciales para que algo sea oro (pues dependen de la estructura del oro) e, incluso, si se quiere, podríamos conocerlas *a priori*, en tanto que hablantes competentes, pero que podrían ser del todo insuficientes para seleccionar únicamente la extensión del término, pues alguna otra sustancia podría presentar esas mismas características (por ejemplo, la pirita de hierro). En la sección 1.5.2 presento el argumento de Putnam de la Tierra Gemela que explota esta idea respecto al agua y puede considerarse como un argumento de ignorancia, pues las características asociadas al agua, aunque verdaderas y quizá necesarias predicadas del agua, podrían ser insuficientes para seleccionar únicamente el referente de ‘agua’.

En suma, la posibilidad del error respecto a las propiedades necesarias que, a veces,

podemos asociar con nuestros términos de clase natural siempre está presente. Cuando se trata de propiedades necesarias, la posibilidad en juego es de tipo epistémico, ya se trate de propiedades que son observables o de propiedades que requieren investigación empírica. En el primer caso, existe la posibilidad de la ilusión colectiva; en el segundo, la posibilidad de la falibilidad de nuestros descubrimientos. Otra cosa es que *sepamos*, como de hecho sabemos, que los gatos son animales y que el oro es un metal, y que estas propiedades son, en ambos casos, necesarias para pertenecer a las respectivas clases, pero nuestra situación epistémica podría ser diferente.

Sin embargo, estos argumentos externistas del error aplicados a ‘gato’ o a ‘oro’ parecen diferentes de los argumentos del error aplicados a nombres propios. En este último caso, tenemos características contingentes asociadas al portador del nombre propio, y, por tanto, la posibilidad implicada en los contraejemplos externistas parece de un tipo más fuerte que la epistémica. Así, en el argumento kripkeano del error respecto a Gödel, se asocia con ‘Gödel’ una descripción, ‘el descubridor de la incompletud de la aritmética’, que expresa una característica claramente contingente de aquél del que se predica. Entonces, el experimento mental parece basarse en la mera posibilidad metafísica consistente en que Gödel podría no haber descubierto la incompletud de la aritmética. En cambio, en el caso de los gatos-robots, la plausibilidad del experimento se basa en la falibilidad de nuestro conocimiento y no en una mera posibilidad metafísica pues, como se ha dicho, si los gatos son animales, entonces no podrían ser robots. La asimetría que se produce entre los nombres propios y los términos de clase natural en este tipo de experimentos mentales basados en el error se deriva del hecho que asociamos características individualizadoras de tipo diferente con los portadores de los nombres propios y con los miembros de una clase natural, respectivamente. En el caso de los nombres propios, típicamente asociamos descripciones que explicitan algún hecho notable y singularizador del portador del nombre, pero se trata muy habitualmente de características claramente contingentes (como *ser el descubridor de la incompletud de la aritmética*, *ser el discípulo más famoso de Platón*, etc.). En cambio, en el caso de los términos de clase natural, aunque típicamente asociamos características superficiales comunes a los miembros de una clase, solemos asociar (algunas) características que son esenciales a los miembros de la clase en tanto que tales ya que son características causalmente dependientes de aquello que hace que los miembros sean de la clase que son.

Pero esta asimetría que, de hecho, se da, entre las características que asociamos a los miembros de una clase natural y a las entidades con nombre propio no es algo que tiene que darse necesariamente. Algunos nombres propios pueden ir asociados con descripciones que hagan referencia a alguna característica esencial del portador del nombre: así, por ejemplo, es plausible pensar que prácticamente todas las personas que asocian alguna descripción con ‘Lamprocles’ asocian la descripción ‘el hijo primogénito de Sócrates’. Si acordamos que ser hijo de los individuos que fueron sus padres es una propiedad esencial de Lamprocles (pues si no, sería otro individuo), entonces es fácil generar un argumento del error análogo al que se puede generar con un término de clase natural. Podríamos descubrir (mediante pruebas de ADN realizadas sobre los restos que habríamos encontrado de Sócrates y Lamprocles) que, realmente, Lamprocles no fue hijo de Sócrates. En tal situación, ‘Lamprocles’, presumiblemente referiría a Lamprocles y la determinación de la referencia del nombre propio no la proporcionaría la descripción asociada al nombre. Tal situación es análoga a la anterior situación expuesta de los gatos robots: en ambos casos, tenemos alguna característica esencial asociada al portador del nombre y a los miembros de la clase, respectivamente, sobre la que podríamos haber estado equivocados, dado todo lo que sabíamos en algún momento.

Por otro lado, aunque, como un asunto de hecho, típicamente asociamos a los nombres propios descripciones que expresan características contingentes de sus portadores, y el caso de ‘Lamprocles’ o alguno similar puede parecer excepcional, hay que notar que podríamos estar en una situación en la que fuera habitual asociar características necesarias al portador de un nombre propio. Supongamos que estamos en una situación tal que en nuestro planeta hay individuos procedentes de diferentes planetas. En una situación así, es plausible imaginar que con ‘Aristóteles’ asociamos la descripción ‘el autor humano de *La Metafísica*’, para distinguirlo, quizás, de autores de procedencia extraterrestre que también han escrito obras parecidas. Pero, de nuevo, en esta situación, podríamos habernos equivocado sobre la condición humana de Aristóteles (pese a que *si*, en tal situación, Aristóteles fuera humano, sería necesariamente humano).

En resumen, aunque los argumentos kripkeanos de la ignorancia y el error aplicados a los términos de clase natural pueden parecer, a primera vista, más problemáticos o menos convincentes que aplicados a los nombres propios (por el tipo de características

que habitualmente asociamos a los portadores de los nombres propios y a los miembros de una clase natural, respectivamente), en ambos casos, la posibilidad importante implicada es, en el fondo, del mismo tipo. Creo que los argumentos de ignorancia y error kripkeanos no sólo son extensibles a los términos de género natural sino que son tan plausibles como los argumentos de ignorancia y error en el caso de los nombres propios. Del mismo modo que podríamos equivocarnos sobre ciertas características contingentes que atribuimos a un individuo, también podríamos equivocarnos en la atribución de ciertas características necesarias a un individuo o a una serie de individuos de una clase. El cúmulo de descripciones que asociamos tanto a un nombre propio como a un término de clase natural es siempre, en principio, revisable.

Esta característica de la revisabilidad de nuestras descripciones asociadas a los términos de clase natural es importante tenerla en cuenta porque los defensores de una teoría cúmulo podrían apelar a las descripciones de los expertos como descripciones que sí proporcionan el significado de los términos como 'oro'. Decir que la descripción de los expertos 'el elemento con número atómico 79' sí proporciona el significado de 'oro' también es problemático. Primero, porque tal descripción es revisable, 'el oro es el elemento con número atómico 79' es una afirmación necesariamente verdadera *si es verdadera*. Las descripciones científicas proporcionan un *criterio de aplicación* de los términos de clase natural, pero no dan su significado ni determinan su referente. Y, segundo, porque, antes del desarrollo de la química, el uso normalizado de los términos químicos de clase natural tenía un referente en una situación en la que nadie disponía de los criterios de correcta aplicación de los términos. Así que, la determinación de la referencia de tales términos tiene lugar de manera independiente a los criterios científicos y al descubrimiento de verdades acerca de las clases naturales.

1.5.2. El argumento de Putnam de la Tierra Gemela

Seguramente, el principal y más potente argumento contra el internismo semántico y su principio de transparencia respecto a los términos de clase natural lo proporciona Putnam (1973). Según él, las dos tesis principales del descriptivismo tradicional son las siguientes:

(i) El significado de un término determina su extensión.

(ii) Los significados se encuentran en la cabeza: nos resultan transparentes (principio de transparencia).

Putnam sostiene que las dos tesis conjuntamente no se pueden mantener para los términos de clase natural. Y como (i) parece una tesis verdadera, su conclusión es que (ii) es falsa, cosa que muestra con su experimento mental de la Tierra Gemela. Éste consiste en considerar una Tierra Gemela, un duplicado exacto de la Tierra incluso en los términos del lenguaje, excepto que en la Tierra Gemela hay un líquido de apariencia indistinguible del agua pero con una estructura molecular diferente: XYZ, digamos, en vez de H₂O. En una época anterior al descubrimiento de los elementos químicos, un individuo de la Tierra, Óscar, y su duplicado exacto en la Tierra Gemela, Óscar Gemelo²⁰, utilizan de manera del todo análoga el término ‘agua’ y asocian a él exactamente los mismos pensamientos, ambos tienen en la mente la misma descripción asociada a ‘agua’. Pero ‘agua’ dicho por Óscar y ‘agua’ dicho por Óscar Gemelo tienen claramente significados diferentes, dice Putnam, ya que refieren a sustancias diferentes. ‘Agua’ dicho por Óscar refiere a la sustancia compuesta por H₂O, y ‘agua’ dicho por Óscar Gemelo refiere a la sustancia compuesta por XYZ. La obvia diferencia en extensión de ambos términos implica una diferencia de significado, ya que el significado de un término determina su extensión. Los significados no pueden encontrarse en la cabeza, concluye Putnam, sino que dependen de factores externos a la mente de los individuos.

El argumento intenta mostrar que tanto el internismo como el descriptivismo para los términos de clase natural son falsos. Los significados no son algo que se encuentre en la cabeza y a lo que tengamos un acceso cognitivo (sea mediante material descriptivo o no). Cuando usamos ‘agua’ nos referimos a lo que sea H₂O. El significado de los términos de clase natural no viene dado por una descripción que especifique las características perceptibles del agua o por alguna entidad a la que tengamos acceso cognitivo y nos faculte para aplicar correctamente el término.

²⁰ Como han notado algunos, Óscar Gemelo no puede ser un duplicado *exacto* de Óscar, pues el 70% de su cuerpo está formado por XYZ en vez de H₂O. Salvo por este detalle, sin relevancia para el argumento, podemos considerar a Óscar Gemelo como un duplicado de Óscar.

Aunque hay muchas y variadas réplicas al experimento mental de Putnam, me detendré en exponer las dos líneas de respuesta a Putnam que considero más relevantes. Desde posturas contextualistas, se podría objetar al experimento mental de Putnam que el término ‘agua’ puede tener distintos referentes dependiendo del contexto. En el contexto de la Tierra, ‘agua’ refiere a la sustancia compuesta por H₂O, y en la Tierra Gemela, ‘agua’ refiere a la sustancia compuesta por XYZ. Así, si Óscar viaja a la Tierra Gemela, sus preferencias de ‘agua’ referirán a XYZ antes que a H₂O, ya que el contexto de la Tierra Gemela importa para determinar el significado del término. Esta objeción contextualista podría buscar apoyo en otros casos similares donde un cambio de contexto impone un cambio en el referente de un término. Así, por ejemplo, la descripción ‘el presidente’ proferida por un individuo en España hace referencia al presidente del Estado español, Rajoy. Ahora bien, la misma descripción proferida por el mismo individuo en los EUA hace referencia al presidente de los EUA, Obama. Parece que el contexto de la preferencia es el que determina, en parte, el significado del término. Ahora bien, esta objeción contextualista puede ser replicada de varias maneras. Se podría decir, por ejemplo, como Davidson (1992), que si Óscar viaja a la Tierra Gemela, sus usos de ‘agua’ allí no referirán a la sustancia compuesta por XYZ, sino que continuarán refiriendo a la sustancia compuesta por H₂O porque lo que importa para determinar la referencia de un término es el origen de su uso, en qué circunstancias ambientales se empezó a usar el término. Y los usos originales de ‘agua’ por parte de Óscar refieren a H₂O y están conectados históricamente con la sustancia compuesta por H₂O, no con la sustancia compuesta por XYZ. Otra línea de respuesta a la objeción contextualista hace recaer el acento en la *indexicalidad* antes que en el origen del uso: ‘agua’ dicho por Óscar en la Tierra Gemela refiere a lo que tenga la misma naturaleza que *esta materia* (la que sea relevante según el lenguaje que adopte Óscar). Si Óscar usa el lenguaje de la Tierra Gemela, entonces Óscar, con el término ‘agua’ de ese lenguaje, refiere a XYZ, pero cuando usa ‘agua’ en *su* lenguaje, el de la Tierra, refiere a H₂O.

Otra vía diferente para intentar objetar el experimento mental de Putnam sería decir que su experimento muestra que podría existir agua que no fuera H₂O, sino XYZ y, por tanto, ‘agua’ podría referir, o refiere (si hay una tierra gemela), indistintamente tanto a la sustancia compuesta por H₂O como a la sustancia compuesta por XYZ. La extensión del término abarcaría tanto las muestras de H₂O como las muestras de XYZ. De hecho, Searle (1983) propone que, si descubriéramos una Tierra Gemela como la del

experimento de Putnam, podríamos elegir libremente ampliar el referente de ‘agua’ para que incluyera tanto H₂O como XYZ y, además, esto sería, según él, lo más sensato. La réplica a esta objeción modal es que no es posible, de acuerdo con Putnam, que exista agua que no esté compuesta por H₂O. Si el agua tiene esa estructura interna, entonces la tiene necesariamente; es decir, todo aquello que no está compuesto por H₂O no puede ser agua. De nuevo, es conveniente aclarar que la Tierra Gemela no es un mundo posible en la medida en que las propiedades observables que presenta el agua estén causadas por su estructura interna y sólo tal composición interna pueda generar esas propiedades superficiales. La Tierra Gemela es una posibilidad epistémica: teniendo en cuenta todo lo que sabíamos sobre el agua antes del descubrimiento de la química, la Tierra Gemela podría haber existido.

En definitiva, el experimento mental de la Tierra Gemela desacredita el internismo semántico para los términos de clase natural, pues parece, desde mi punto de vista, un argumento coherente y convincente sin una réplica obvia. Si Putnam tiene razón y los significados no se encuentran *en la cabeza*, entonces no tenemos un acceso cognitivo a los significados. En el experimento mental, Óscar y Óscar Gemelo comparten absolutamente todos los estados mentales y asocian exactamente lo mismo con el término ‘agua’. Si se trata de una descripción asociada a ‘agua’, asocian la misma descripción; si se trata de tener acceso cognitivo a una serie de características individualizadoras del agua, también poseen el mismo acceso cognitivo. Por tanto, su argumento refuta tanto al descriptivismo como al internismo semánticos, en cualquiera de sus versiones, pues sea lo que sea que conocen, conscientemente o inconscientemente, en relación al agua, conocen lo mismo. Esta refutación exige renunciar al principio de transparencia internista: los significados de nuestros términos pueden no resultarnos transparentes por no ser cognitivamente accesibles (ni individual ni colectivamente).

Para Putnam, la extensión de los términos de clase natural no está determinada por descripciones asociadas a los términos por los hablantes, sino por algo *externo* que puede ser no accesible directamente a la mente, por características que constituyen la extensión de los términos. Y es posible que estas características tengan que ser descubiertas. Así, por ejemplo, la extensión del término ‘agua’ está determinada por la característica de estar constituida por moléculas de H₂O. Y esta característica, que

consideramos *subyacente*, tiene que ser descubierta. Dentro de este marco explicativo está claro por qué se pueden dar los problemas de la ignorancia y del error: las características del mundo que determinan la referencia de los términos de clase natural nos pueden ser desconocidas o podemos equivocarnos sobre ellas.

1.5.3. La propuesta positiva del externismo

Tras la crítica al descriptivismo y al internismo en sus diferentes versiones, los autores de la nueva teoría de la referencia expusieron sus teorías positivas acerca de cómo se determina la referencia de los nombres propios y de los términos generales (en particular, de los términos de género natural). La explicación externista propuesta, inicialmente pensada para los nombres propios, es ampliable a los términos de clase natural. En efecto, Kripke destaca las similitudes que se dan entre los nombres propios y algunos términos generales, en especial, los términos para clases naturales²¹, y considera que su explicación de cómo se determina la referencia de los nombres propios es válida para el caso de los términos de clase natural.

La teoría positiva de Kripke acerca de cómo se determina la referencia de los nombres propios y los términos de clase natural tiene dos fases: la introducción del término y la transmisión de éste. En un primer momento, la introducción del término en el lenguaje se produce mediante un bautizo inicial. En el caso de un nombre propio, en el momento del bautizo de un individuo, las intenciones de los hablantes presentes que participan en el bautismo juegan un papel: los hablantes intentan que con ese nombre se haga referencia a aquella entidad. Kripke admite que, en este momento inicial de introducción del nombre, cuando se fija la referencia, es perfectamente plausible que el hablante o los hablantes presentes en la ceremonia tengan en la mente alguna descripción de la entidad bautizada que acompañe la introducción del nombre. Pero esta descripción o descripciones no dan el significado del término; pueden servir para *fijar* la referencia, pero no determinan la referencia. La determinación de la referencia del nombre se produce mediante la presencia del individuo bautizado y la intención de los

²¹ Una de las similitudes, aparte del hecho que tanto en el uso de los nombres propios como en el uso de los términos de clase natural pueden generarse los problemas de la ignorancia y el error, es que ambos tipos de términos son, de acuerdo con Kripke, designadores rígidos (designan la misma entidad en todo mundo posible en que la entidad existe).

hablantes de querer referir a ese individuo con ese nombre.

Una explicación análoga sirve para la introducción de los términos de clase natural. Así, en la introducción de un término como ‘oro’, es la propia sustancia oro (cuyas propiedades podríamos desconocer) junto con las intenciones de los hablantes de querer referir a esa sustancia (mediante un ejemplar presente) lo que determina la referencia del término. Hay que notar que no se trata meramente de que las propiedades que hacen que algo sea oro son las que determinan la referencia del término, pues un internista podría estar de acuerdo con esto siempre que se trate de propiedades cognitivamente accesibles. La idea externista es que esas propiedades objetivas del oro y que determinan la referencia de un término de clase natural *podrían no ser* cognitivamente accesibles a los hablantes. Habitualmente, en el caso de las clases naturales, tenemos la intuición que hay alguna propiedad *más allá de las apariencias* que presentan los ejemplares que hace que éstos pertenezcan a una cierta clase natural. Esta intuición se muestra en el ejemplo que ofrece Kripke sobre el posible descubrimiento de que el oro fuera realmente azul, y no amarillo. En tal circunstancia, como vimos, la intuición es que diríamos que el oro realmente es azul, no que no hay oro puesto que éste *debe ser* amarillo. Este descubrimiento no cambiaría el significado del término ‘oro’, pues con ‘oro’ intentamos referir a lo que es verdaderamente oro, y no a lo que se ajusta a una descripción de las características perceptibles del oro. Cuando usamos un término de clase natural, queremos referirnos a lo que es *realmente* la sustancia correspondiente, a aquello que Locke llamaba ‘la esencia real’. Así, cuando usamos ‘oro’ o ‘gato’ nos referimos a una sustancia y a una especie, respectivamente. Pero podría haber clases naturales cuya naturaleza no fuera subyacente, de manera que sus miembros sólo compartieran propiedades observables²². Entonces, el internista podría decir que el acceso cognitivo a esas propiedades está garantizado. Pero, de nuevo, aún en el caso de propiedades meramente superficiales, la posibilidad epistémica del error respecto a esas propiedades siempre está presente (debido a una posible ilusión colectiva).

Además, el hecho mismo que las características singularizadoras de una clase natural sean solamente características superficiales es un hecho al que no tenemos un acceso

²² Éste parece ser el caso de algunas enfermedades como la artritis o la fibromialgia.

cognitivo infalible, sino un hecho que debemos descubrir. Y, por tanto, podríamos equivocarnos sobre si los miembros de una determinada clase natural comparten o no una naturaleza subyacente. Al introducir, entonces, un término de clase natural, la idea externista es que pretendemos referir a aquella propiedad o propiedades objetivas, *externas*, de los miembros de la clase que hacen que sean de la clase que son, y son propiedades tales que, sean subyacentes (como normalmente pensamos) o sean superficiales, podríamos equivocarnos sobre ellas. Así, el acceso cognitivo a tales propiedades no está garantizado. Desde este punto de vista, la extensión del término queda determinada por la propia naturaleza de las clases, sea la que sea, no por una idea o una descripción que comprenda las características que presumimos conocer de los miembros de la clase.

En una segunda fase, Kripke explica cómo los términos conservan su referencia en el lenguaje a medida que pasa el tiempo²³. Una vez que el término ha sido introducido en el lenguaje, los hablantes presentes en la ceremonia de bautismo transmiten el término a otros hablantes, y estos hablantes tienen la intención de usar el término de la misma manera que lo usan o usaban aquellas personas de las que lo recibieron; es decir, intentan referir al individuo o a la sustancia que fue bautizada, independientemente de si poseen o no una descripción que singularice a aquel individuo o a aquella sustancia natural. Así, mediante una cadena histórica entre los hablantes, o entre los usos del término que hacen los hablantes, los nombres propios y los términos de clase natural se van transmitiendo de unos hablantes a otros conservando su referente.

La teoría constructiva de Kripke implica que podemos usar tanto un nombre propio como un término de clase natural sin tener acceso cognitivo a una descripción que singularice el referente o el ámbito de aplicación del término. Los problemas de la ignorancia y el error son posibles cuando usamos nombres propios o términos de clase natural y, sin embargo, intuitivamente, conseguimos referir cuando los usamos aun siendo ignorantes o estando equivocados sobre aquello que singulariza al referente. Un nombre propio simplemente refiere al portador del nombre, conozcamos o no alguna

²³ Hay que notar que algunas veces se producen cambios en la referencia original de un término con el paso del tiempo; es el caso de 'Madagascar' sugerido por Evans. La posibilidad de estos casos hace que la teoría histórico-causal de Kripke sea una teoría que proporciona condiciones necesarias para que un término refiera, pero no suficientes. Kripke es consciente de este hecho.

característica que lo singularice; un término de clase natural refiere a una sustancia o a una especie, según el caso, conozcamos o no alguna característica distintiva de la clase referida.

Putnam también presenta una teoría sobre cómo se determina la referencia de los términos de clase natural. La determinación de la referencia se produce mediante dos elementos *externos* a los hablantes: un elemento indéxico y un componente social. El componente indéxico se muestra en la introducción del término; pues ésta se realiza idealmente de manera ostensiva señalando una muestra de la clase. Delante de una sustancia natural, los introductores del término de clase intentan referir a lo que sea que tiene la sustancia que hace que sea el tipo de sustancia que es. Así, delante del agua, los hablantes que introducen el término de clase ‘agua’ intentan referir a aquella propiedad o propiedades que tiene el agua que hacen que sea precisamente agua. En adelante, la intención de los hablantes es usar ‘agua’ para referirse a todo lo que tiene las mismas propiedades *relevantes* que tenía la sustancia paradigmática y que hacían que fuera agua. Así, la referencia de un término de clase natural se determina por una relación de similitud con un paradigma. La ostensión inicial puede ir acompañada de una descripción que previsiblemente expresará características superficiales de la clase o de la muestra de la clase, pero, de nuevo, esta descripción no es la que determina la referencia del término porque estas descripciones pueden no especificar las propiedades *importantes* que hacen que la muestra pertenezca a la clase natural, propiedades que, seguramente, habrá que descubrir. Así, la extensión de los términos de clase natural está determinada, al menos en parte, por la sustancia nombrada. La referencia de ‘agua’ está determinada por la propia sustancia agua (o por propiedades de la sustancia agua) sin la mediación de ninguna descripción o serie de descripciones que expresen características superficiales del agua. A partir de aquí, lo que cuenta como de la misma clase que el paradigma es un asunto de investigación empírica.

En este punto vale la pena apreciar que, respecto a lo que determina la referencia de los términos de clase natural, donde Kripke habla de propiedades *esenciales*²⁴, Putnam

²⁴ Parece, aunque es controvertido, que, al hablar de propiedades esenciales a un objeto natural, Kripke quiere decir tanto que esas propiedades son esenciales al objeto *en tanto que perteneciente a esa clase natural* como al objeto *en tanto que objeto particular*. Es decir, parece que defendería un esencialismo de

habla propiedades *importantes* o *relevantes*²⁵. Estas propiedades importantes son las propiedades que hacen que, por ejemplo, una muestra particular sea agua; es una propiedad esencial a la muestra pero en tanto que es agua (no necesariamente en tanto que entidad particular). Así explica Putnam cómo se determina la extensión de ‘agua’:

[...] an entity x , in an arbitrary possible world, is *water* if and only if it bears the relation *same L* [*same liquid as*] (construed as a cross-world relation) to the stuff we call "water" in the actual world. (Putnam 1973, 708)

La relación *mismo líquido que* hay que entenderla como una relación a través de los mundos posibles por la que si un líquido en el mundo W_1 tiene las mismas propiedades físicas importantes que un líquido en W_2 , entonces el primer líquido mantiene la relación *mismo L* con el último líquido. Cuáles sean las propiedades físicas importantes es algo que requiere investigación empírica por parte de la ciencia, con la posibilidad inherente de la revisabilidad de nuestros descubrimientos.

Putnam habla de un segundo componente, de tipo social, en la determinación de la referencia de los términos para clases naturales. Este componente social explica cómo es posible la referencia una vez introducidos los términos de clase natural en el lenguaje. En nuestra comunidad lingüística se da una cooperación social a la hora de determinar la referencia de los términos. Esta cooperación social puede tomar diferentes formas. Una de ellas es la división del trabajo lingüístico. Así, por ejemplo, algunas personas expertas conocen las características distintivas que hay entre los olmos y las hayas, y otras no. Como los expertos en botánica conocen las características distintivas de estas clases naturales, el resto de la comunidad lingüística, ignorante, puede referir a ellas mediante la deferencia a los expertos. Hay que notar que la deferencia a los expertos puede remitir a una descripción singularizadora de la clase en cuestión que los

clase y también un esencialismo individual. Sin embargo, no quiero comprometerme con esta interpretación de la postura kripkeana.

²⁵ Estas propiedades son *importantes* o *relevantes* para nuestros intereses clasificatorios de las clases naturales, pero también para hacer que algo pertenezca a la clase natural a la que pertenece. Es decir, hay diferentes posibilidades clasificatorias de las clases naturales, pero siempre atendiendo a propiedades objetivas y susceptibles de ser descubiertas que tienen los miembros de dichas clases.

expertos conocen, pero esta descripción es conocida después de llevar a cabo investigación empírica. Después de ulterior investigación empírica, los expertos podrían descubrir que la descripción es equivocada. Así, la descripción, aunque verdadera y singularizadora de la clase, no está conectada ni *a priori* ni, quizás, necesariamente con el término de clase natural. Lo que quiere destacar Putnam apelando a la división de la tarea lingüística es que la extensión de muchos términos de clase natural está determinada socialmente, no individualmente. Con ‘haya’ conseguimos referir a las hayas porque en nuestra comunidad lingüística hay expertos que, basándose en el conocimiento empírico de las características de las hayas, saben cuáles son las características *relevantes* para ser un haya. Pero el conocimiento descriptivo que poseen los expertos no determina por sí solo la referencia de los términos de clase natural; puede fijar la referencia de un término proporcionando un criterio para distinguir la clase, haciendo así que los usos de los hablantes sean parasíticos de los usos expertos.

Los dos elementos citados, el elemento indécico y el elemento social presentes en la referencia de los términos generales son factores que remarcan el externismo de Putnam, ya que actúan en la determinación de la referencia como factores *externos* al individuo (o a la mente del individuo) que usa los términos y consigue referir con ellos. Sin embargo, el factor social putnamiano en la determinación de la referencia de los términos de clase natural puede llevar a pensar que Putnam está defendiendo algún tipo de descriptivismo *social* para los términos de clase natural, cercano al defendido por una teoría cúmulo descriptivista. Kripke alerta del peligro de considerar que la teoría putnamiana, al introducir un elemento social en la determinación de la referencia de los términos de clase natural, es solamente una versión social de la vieja teoría descriptivista. Ésta parece ser la interpretación que hace Dummett, de acuerdo con Kripke (1986), de la teoría de Putnam, y Kripke argumenta contra esta interpretación. Según Kripke, en los casos típicos de términos de clase natural, como ‘oro’ o ‘agua’, los expertos no determinan de ninguna manera la extensión de los términos. Su papel se reduce a encontrar y distinguir las cosas que caen bajo la extensión del término de clase natural, pero no poseen ningún privilegio semántico especial a la hora de determinar la extensión de los términos. La extensión del término está fijada previamente al criterio de los expertos para distinguir aquello que cae bajo su extensión, y es independiente de la existencia de expertos o no en el asunto. Además, los expertos podrían equivocarse completamente sobre la extensión de un término de clase natural. Por tanto, dice Kripke,

no hay que interpretar lo que dice Putnam acerca de la división del trabajo lingüístico como diciendo que la presencia de expertos es lingüísticamente crucial para determinar la extensión de un término. Es cierto que esa interpretación parece sugerida por lo que dice Putnam, pero no puede ser la interpretación correcta porque sería incompatible con lo que defiende Putnam.

Kripke hace notar que sí puede haber casos en que algunos expertos tengan ciertos poderes lingüísticos especiales, como en el caso de términos originados por esos expertos que después se han extendido a la comunidad de hablantes. Son casos como el de ‘protón’ u otros términos técnicos similares. Aquí, el papel especial de esos expertos deriva del hecho que éstos son los introductores del término y, como tales, tienen una autoridad especial en términos del bautismo inicial. Pero estos términos técnicos no constituyen los casos paradigmáticos de términos de clase natural. Y, de todos modos, la determinación de la referencia de los términos de clase natural (sean o no términos técnicos) no puede hacerse depender de la existencia de expertos contemporáneos porque, en un momento dado, todos los expertos sobre un asunto, sobre protones – pongamos por caso-, podrían desaparecer (quedando en el mundo sólo gente con una idea vaga o equivocada acerca de la clase en cuestión), y eso no implicaría un cambio en el referente del término. La idea detrás de la teoría causal de la referencia es que el referente, como característica semántica relevante de nuestros términos, se preserva en diferentes circunstancias posibles, independientemente de la existencia o no de expertos en algún asunto. Entonces, ningún subgrupo lingüístico tiene un papel preponderante al determinar la referencia de nuestros términos de clase natural (y lo mismo para los nombres propios), pues, además, la comunidad lingüística, incluyendo a esos expertos, podría tener una concepción errónea sobre una clase natural (o sobre un individuo).

Creo que las consideraciones de Kripke son más fieles a la concepción externista básica de la teoría putnamiana que algunas interpretaciones que podrían hacerse de dicha teoría en una línea de tipo descriptivista. Al fin, lo que parece esencial en la teoría de Putnam, y afín al resto de sus ideas y de las ideas defendidas por otros autores externistas, es que los hablantes que desconocen lo que caracteriza una clase natural son capaces de referir a ella mediante una deferencia lingüística a otros hablantes. Es la concepción de fondo que también presenta la teoría causal histórica de Kripke: mediante la deferencia a otros hablantes de los que se ha tomado el término, se llega a los introductores del término y

al referente presentes al inicio de la cadena histórica. Es cierto que en la cadena de transmisión puede haber expertos que conozcan una descripción que singulariza la clase, pero la deferencia a esos expertos hay que entenderla, como dice Kripke, en tanto que introductores o usuarios previos del término, no en tanto que concedores de una descripción que es la que determina la referencia de nuestros términos de clase natural.

Capítulo 2. Las clases artificiales y los términos de clase artificial

En la primera sección de este capítulo, la sección **2.1**, expongo algunas diferencias y similitudes entre las clases naturales y las clases artificiales, puesto que éstas últimas se conciben contrastándolas con las primeras. A partir de la especificidad de las clases artificiales, discuto en la sección **2.2** las diferentes propuestas acerca de la naturaleza esencial de las clases artificiales. Para ello, partiendo del principio de realizabilidad múltiple para los artefactos, considero el punto de vista según el cual las clases artificiales tienen una esencia nominal. Tras criticar la idea de que las clases artificiales son clases meramente nominales (y los términos de clase artificial son términos meramente nominales), considero, repasando diferentes concepciones acerca de la esencia de los artefactos, la dificultad de proporcionar condiciones necesarias y suficientes de pertenencia a las clases artificiales.

Una vez presentada y discutida la cuestión metafísica de la naturaleza de las clases artificiales, en la sección **2.3**, expongo las consideraciones para rechazar que la teoría metafísica acerca de las clases artificiales determine la semántica apropiada para los términos para artefactos, como algunos autores han insinuado o reconocido. Asimismo, presento iguales consideraciones para el caso de las clases naturales y los términos de clase natural. Por último, en la sección **2.4**, me centro en el aspecto semántico de la determinación de la referencia de los términos de clase artificial. Para ello, presento la primera propuesta semántica para estos términos: la propuesta externista putnamiana, e intento evaluar su adecuación.

2.1. Diferencias y similitudes entre las clases naturales y las clases artificiales

2.1.1. Diferencias entre las clases naturales y las clases artificiales

Las clases artificiales y los objetos que las componen se conciben contrastándolos con las clases naturales y sus miembros. A continuación, expongo algunas diferencias intuitivas y ampliamente aceptadas entre las clases artificiales y las clases naturales, y después paso a comentarlas:

- (i) La existencia de las clases artificiales (como la clase de las sillas o la clase de los lápices) depende de la intervención humana²⁶; mientras que la existencia de las clases naturales es independiente de nuestra intervención.
- (ii) La naturaleza o propiedad característica de las clases naturales es una propiedad microfísica; en cambio, lo característico de las clases artificiales no parece ser una propiedad microfísica, sino, intuitivamente, su función.
- (iii) Nuestro acceso epistémico a la naturaleza de las clases naturales parece, de entrada, diferente al que tenemos respecto a las clases artificiales: mientras que la naturaleza de las clases naturales, entendida como la composición microfísica que comparten sus miembros, puede resultarnos cognitivamente inaccesible, éste no parece ser el caso con la función que tienen los artefactos que creamos.

Respecto al punto (i), parece obvio que la intervención humana es necesaria para la existencia de artefactos humanos. Esta intervención se da a diferentes niveles: los seres humanos no sólo crean o producen los artefactos sino que, además, los van perfeccionando a lo largo del tiempo. En cambio, las clases naturales existirían con total independencia de la intervención y existencia humanas. Además, otro factor relacionado con la intervención humana en la creación de artefactos es que otorgamos intencionalmente una finalidad a los miembros de las clases artificiales. Algunos autores, como Thomasson (2007), han caracterizado de manera general a las clases artificiales como el producto intencionado de las actividades humanas. Aunque esta caracterización no sirva como una definición estricta de las clases artificiales, pues incluiría productos intencionales de las actividades humanas que difícilmente calificaríamos de *artefactos* (como, por ejemplo, las consecuencias de nuestras acciones morales), sí incluye aspectos típicos de los artefactos: son productos humanos y su producción está guiada por una intención.

Otra diferencia entre las clases artificiales y las clases naturales destacada por los autores (Schwartz 1978; Kornblith 2009) es la que se recoge en el anterior punto (ii). Las clases naturales están caracterizadas, o así lo creemos, por una naturaleza microfísica. A veces, en vez de ‘naturaleza microfísica’, se emplea el término ‘naturaleza subyacente’. La idea es que esa naturaleza subyacente u oculta es de tipo

²⁶ O de algún ser inteligente no humano.

microfísico y hay que entenderla en términos de igualdad de composición química o molecular por parte de los miembros de una misma clase natural. Así, por ejemplo, todas las muestras de oro, comparten la propiedad de tener número atómico 79, pues están compuestas por átomos de oro, y todos los ejemplares de la especie tigre comparten un cierto genoma característico de la especie (aunque esto es más controvertido ya que parece que no siempre el criterio de similitud genética delimita la especie²⁷). En general, se puede afirmar que la propiedad característica de las clases químicas es su número atómico o composición molecular, y la propiedad característica de las clases biológicas es un cierto genoma. En cambio, es obvio que los miembros de una misma clase artificial muchas veces no comparten una misma propiedad estructural o composicional en este sentido microfísico. Puede haber mesas con diferentes composiciones materiales: mesas de metal, de madera, de plástico, etc. De entrada, la composición material de los artefactos no es el criterio que consideramos para su clasificación en las diferentes clases. Intuitivamente, la naturaleza compartida por los artefactos de una misma clase parece ser su función. La función de los artefactos sí parece ser un buen candidato para la identificación y distinción de cada clase artificial, aunque, a veces, podría considerarse que otras características, como la forma o el material, pueden jugar también un papel para clasificar a un objeto dentro de una determinada clase artificial. Por ejemplo, puede considerarse que las sillas y los taburetes comparten la misma función de servir como asiento y parece que lo que hace que los clasifiquemos en diferentes clases artificiales tiene que ver con sus diferentes formas. En el caso de los bolígrafos y los lápices, también podríamos decir que su función es básicamente la misma: servir para escribir. Y algo que los distingue es su diferencia en material; mientras un componente de los lápices es el grafito, los bolígrafos incorporan tinta para llevar a cabo la misma función. Pero de modo aproximativo, podríamos decir que la naturaleza de las clases naturales radica en una propiedad subyacente al nivel microfísico; mientras que la naturaleza de las clases artificiales radica principalmente en su función (más, quizás, otras características añadidas en algunos casos).

²⁷ El concepto moderno de *especie*, desde Mayr, considera que ésta se caracteriza por un acervo genético del que cada individuo de la especie sólo tiene una parte. La variabilidad intraespecífica que permite el acervo genético imposibilita que se pueda hablar de un genoma o conjunto de genes propios compartidos únicamente por todos los miembros de una especie.

La diferencia entre las clases artificiales y las clases naturales destacada en el punto (iii) tiene que ver con nuestro acceso epistémico a estas clases, respectivamente. La función de los artefactos es una propiedad pretendida por nosotros; en cambio las propiedades microfísicas de las sustancias naturales son poseídas por las sustancias naturales independientemente de nuestras intenciones. Este hecho parece marcar alguna diferencia en nuestro acceso epistémico a la naturaleza de las clases naturales y de las clases artificiales. Antes del desarrollo de la química moderna, nadie tenía acceso epistémico a la composición molecular del agua, por ejemplo. En cambio, en el caso de los artefactos, parece que siempre hay (o ha habido) un experto que sí conoce (o ha conocido) cuál es la función de los miembros de una clase artificial dada.

2.1.2. Similitudes entre las clases naturales y las clases artificiales

Pese a las diferencias intuitivas que hay entre las clases naturales y artificiales, también hay ciertas similitudes que habitualmente no se destacan en la literatura. Estas similitudes derivan del hecho que tanto la naturaleza de las clases artificiales como la naturaleza de las clases naturales no es algo superficial o directamente observable. En efecto, éste es el caso para la naturaleza microfísica de las clases naturales, pero también para la naturaleza funcional de las clases artificiales. Si se acepta que la función constituye, al menos en parte, la naturaleza de los diferentes tipos de artefactos, se trata de una característica que no es ni superficial ni obviamente accesible. Otra cosa es que, por la familiaridad que tenemos con muchos artefactos cotidianos, como los lápices o las sillas, identifiquemos de inmediato cuál es su función. Pero la función de los artefactos no es una característica obvia como su forma o color. Si nos encontráramos ante un artefacto nuevo o desconocido del que ignoráramos su función, tendríamos que llevar a cabo una investigación empírica para tratar de ver qué función tiene o puede tener. O sea que la tendencia a contraponer la naturaleza de las clases naturales a la naturaleza de las clases artificiales en términos de naturaleza subyacente *versus* naturaleza *superficial* y *accesible* (Schwartz 1978) es una tendencia expresada en términos incorrectos. La función de los artefactos no es superficialmente perceptible y esto hace que las clases naturales y las clases artificiales presenten ciertas similitudes²⁸.

²⁸ Una diferencia no explicitada antes entre las clases naturales y las clases artificiales es que estas últimas tienen un creador y ese creador, en principio, conoce su función. Pero se podría considerar que también las clases naturales tienen un creador que conoce su esencia. Si así fuera, las clases naturales estarían a la

Algunas de estas similitudes son:

- (i) Para descubrir la naturaleza (la función) de un artefacto, puede ser preciso llevar a cabo investigación empírica, como ocurre en el caso de los objetos naturales.
- (ii) El papel de los expertos, como en el caso de las clases naturales, puede ser importante o decisivo para determinar si un miembro dado pertenece a una cierta clase.
- (iii) Es posible hacer descubrimientos respecto a la naturaleza de los artefactos, tal como es posible hacer descubrimientos respecto a la naturaleza de las clases naturales; por ejemplo, es posible descubrir nuevas funciones no contempladas de un artefacto.

El punto (i) no pretende describir el caso más típico de nuestra relación epistémica con la naturaleza o la función de los artefactos. Pero ciertamente hay casos en que hay que investigar para descubrir la función de un artefacto. Como ya he dicho, podría ser necesario emprender una investigación empírica para descubrir la función de un artefacto desconocido, bien porque sea un artefacto nuevo para nosotros y no dispongamos en ese momento de la explicación de un experto, bien porque sea un artefacto de otra época o cultura cuyos expertos o usuarios ya no existen. Otra similitud entre las clases naturales y las clases artificiales es que, como recoge el punto (ii), en el caso de los artefactos, también puede requerirse un experto para determinar si un objeto dado es o no un cierto tipo de artefacto, puesto que la función, como la composición microfísica, no es una propiedad observable. Por ejemplo, podría requerirse el concurso de un experto en mecánica para determinar si una cierta pieza es o no el cilindro de un motor, del mismo modo que puede requerirse la opinión de un experto para determinar si cierta pieza es una pieza de oro o sólo lo parece.

Por último, igual que ocurre con las clases naturales, también parece posible hacer descubrimientos respecto a la naturaleza de los miembros de las clases artificiales; parece posible descubrir nuevas funciones de los artefactos ya conocidos. Así, por ejemplo, el cable eléctrico se introdujo en la primera mitad del siglo XX con la función de transmitir energía eléctrica y proporcionar luz, aprovechando su propiedad de transmitir señales de baja frecuencia. Décadas después, se averiguó que ese mismo tipo de cable también puede transmitir señales de alta frecuencia, lo que lo hace apto para la

par, en este asunto, con las clases artificiales. Retomaré más adelante, en el **capítulo 4**, este aspecto de un hipotético creador de las clases naturales y su acceso epistémico a la naturaleza de lo que ha creado.

transmisión de datos, obteniendo así una nueva aplicación para transmitir información en el campo de las telecomunicaciones. El caso del cable eléctrico muestra que podemos hacer descubrimientos acerca de las propiedades de los artefactos y estos descubrimientos pueden conllevar que cambiemos nuestros criterios clasificatorios. Así, el cable eléctrico es clasificado recientemente también como transmisor de datos, junto a otros artefactos ya conocidos e introducidos para esta función.

En resumen, respecto a las diferencias entre las clases naturales y las clases artificiales, los miembros de las primeras, a diferencia de los miembros de las segundas, parecen compartir una naturaleza subyacente microfísica que es independiente de nuestras intenciones. Respecto a las similitudes entre las clases naturales y las clases artificiales, aunque la función no sea una característica microfísica como lo es la composición molecular, la función de los artefactos tampoco es una característica obvia y directamente perceptible, posibilitando ciertos descubrimientos sobre las clases artificiales o dando un cierto papel a los expertos.

2.2. La esencia de los artefactos

2.2.1. El principio de realizabilidad múltiple y la esencia nominal de los artefactos

Un aspecto importante de la naturaleza de los artefactos es que son objetos físicos, materiales, a los que atribuimos una cierta función²⁹. Y una característica observada en los artefactos es que una misma función puede implementarse en diferentes disposiciones o estructuras físicas³⁰. Es lo que se ha llamado ‘el principio de realizabilidad múltiple para los artefactos’ y que ha sido considerado por diferentes

²⁹ En la introducción de esta tesis, aclaro que me centraré en los artefactos que son objetos materiales concretos, y que son, típicamente, objetos a los que atribuimos una función (y no, por ejemplo, artefactos como las obras de arte o como las instituciones abstractas).

³⁰ No queda siempre claro qué es lo que hay que entender exactamente por ‘estructura física’. A veces parece hacer referencia a la *composición material* (al tipo de materia que constituye el artefacto o incluso a la composición química) y, otras veces, a la *composición estructural*, algo más cercano a la *forma* del artefacto. De todos modos, hay ejemplos intuitivos de ambos casos, tanto de diferente composición material como de diferente estructura, con artefactos de un mismo tipo que realizan la misma función (como el caso de los relojes).

autores, por ejemplo, Dennett (1987), como un rasgo para comprender el mundo artificial. La idea del mencionado principio es que objetos con diferentes estructuras físicas pueden realizar la misma función y pertenecer a la misma clase artificial. Por ejemplo, un reloj digital y un reloj de sol, a menos que se considere que pertenecen a diferentes clases artificiales, son miembros de la clase *reloj*, pese a tener una estructura y composición material muy diferentes. O las sillas pueden estar hechas de diferentes materiales (de hierro, de madera, de mimbre, etc.) sin que por ello dejen de ser sillas ni dejen de tener la función que cumplen las sillas. Se pueden poner muchos ejemplos de este tipo que hacen pensar que el principio de la realizabilidad múltiple es un principio que, muchas veces, es verdadero aplicado a los artefactos.

Sin embargo, hay que notar que es un principio que no se cumple siempre o que, a menudo, presenta ciertas restricciones; en ocasiones, una determinada función puede exigir una determinada composición material o cierta estructura (disposición de la forma) general. Así, por ejemplo, al menos hasta ahora, los neumáticos de los coches han requerido del caucho para funcionar de la manera eficiente que se pretende. O las navajas, por ejemplo, parecen requerir de algún material metálico con ciertas cualidades para la constitución de su hoja. También puede suceder que el material que compone un artefacto de una cierta clase sea importante para clasificarlo como tal. Por ejemplo, los bolígrafos, para llevar a cabo su función y distinguirse de otros artefactos que también sirven para escribir o hacer trazos en algunas superficies, tienen que incorporar algún tipo de tinta. O, poniendo un caso en que una determinada función puede depender de una cierta estructura o disposición en la forma del artefacto, las ruedas de cualquier tipo parecen requerir de una cierta forma (redonda) para llevar a cabo eficientemente su función. En estos casos, se puede decir que, en alguna medida, la función determina una cierta estructura: la estructura sobreviene a la función. De hecho, para no pocos tipos de artefactos una cierta estructura *general*, aunque no necesariamente una composición material concreta, viene impuesta por la función. Por ejemplo, la función de conducción y transporte del agua, requiere una estructura sólida que impida que el agua escape; no sería viable una tubería hecha con un material poroso o poco resistente al peso para canalizar el agua. A menudo la función pretendida impone ciertos límites descartando ciertos materiales o estructuras.

Sin embargo, tenemos la intuición que la estructura composicional no es la característica principal que delimita las clases artificiales, a diferencia de lo que creemos que ocurre con las clases naturales. Parece, de entrada, que lo que delimita las clases artificiales es su función, algo impuesto por nosotros y no por la naturaleza. Esto ha llevado a algunos autores a pensar o a reforzar la idea que las clases artificiales, a diferencia de las clases naturales, no son clases *reales*³¹, sino meramente nominales, convencionales, cuyos miembros son agrupados siguiendo un criterio definicional que nosotros imponemos. A continuación, paso a exponer brevemente las razones que han dado dos autores relevantes, Wiggins (1980) y Schwartz (1978, 1980), a favor de esta postura.

Wiggins (1980) niega que las clases artificiales tengan esencias *reales* porque éstas vienen dadas, según él, por una estructura interna que da cuenta de lo que es el objeto perteneciente a la clase. En particular, la estructura interna o *principio de actividad*, como lo llama Wiggins, tiene que dar cuenta del origen de la existencia, de la existencia y del fin de la existencia de un objeto. Y las clases artificiales están determinadas o agrupadas, según él, por las funciones de sus miembros, no por sus constituciones internas o por una naturaleza intrínseca compartida por todos los miembros de la clase. No hay ningún principio *interno* en este sentido que permita aislar la esencia de un artefacto e identificar su pertenencia a la clase artificial.

Artifacts are collected up not by reference to a theoretically hypothesized inner constitution but under functional descriptions that have to be indifferent to specific constitution and particular mode of interaction with environment. A clock is any time-keeping device, a pen is any rigid ink-applying writing implement, and so on. The description gives what it is usually impossible to specify in the other cases, an explicit *nominal essence*: whereas a finite and determinate set of marks suitable for definitional purposes is precisely not what the members of natural kinds endowed with a scientifically palpable real essence have in common. (Wiggins 1980, 87)

³¹ Es difícil precisar qué entienden los filósofos por *clase real*. Como me parece que pueden entender diferentes cosas, espero que la cuestión se irá aclarando a medida que vaya exponiendo en las páginas que siguen las consideraciones de los diferentes filósofos. En particular, en 2.2.3, abordo directamente la cuestión.

Además, los artefactos, dice Wiggins, a diferencia de los objetos naturales, pueden persistir aunque sus partes cambien, aunque sean parcial o completamente desarmados y reconstruidos, o aunque sus funciones se interrumpan. Por ejemplo, un armario completamente desmontado y guardado hasta su próximo uso persiste como tal; en cambio, dice, esto no ocurre con los objetos naturales³². Y cuando algunos objetos naturales, tales como los órganos de algunos organismos, son agrupados también por su función, entonces, dice Wiggins, esta propiedad funcional sí delimita una clase real porque los diferentes tipos de órganos que realizan una misma función en organismos diferentes, sí tienen un mismo *principio de actividad*³³. Así, para Wiggins, mientras los artefactos no pueden constituir clases reales, sino sólo clases nominales, pues se rigen por el principio de la realizabilidad múltiple, los objetos naturales (también los órganos) sí forman clases *reales*.

Y, ¿qué entiende Wiggins por *clase real*? Entiende aquellas *constituciones internas* que la ciencia puede reconocer y descubrir, como, por ejemplo, la composición química de las sustancias, y parece que califica como *reales* a estas constituciones internas porque nos permiten hacer inducciones y formular leyes o regularidades en la naturaleza. Así, las clases reales son clases objetivas, delimitadas por la naturaleza independientemente de los intereses humanos. Esto no ocurre con los artefactos, cuya clasificación en diferentes clases responde a un criterio (funcional) nuestro. Sin embargo, no está claro si, más allá de estas apreciaciones, Wiggins quiere defender que los miembros de las clases reales son los únicos existentes, frente a los objetos de clases no reales (como los artefactos) que quedarían excluidos de un inventario ontológico del mundo.

³² En principio, parece cierto que un objeto natural no puede desmontarse *completamente* y volver a montarse sin pérdida de sus propiedades como tal. Ahora bien, es cierto que un organismo, por ejemplo, puede perder una de sus partes (un hombre puede perder una pierna, digamos), e incluso puede ser reemplazada por una parte artificial, sin que por ello deje de ser un organismo. Y parece difícil precisar dónde está el límite en el número de partes que puede perder.

³³ No está muy claro cómo hay que entender este *principio de actividad* compartido entre los diferentes órganos de un mismo tipo (por ejemplo, entre los corazones de las diferentes especies animales), pues no es obvio que los órganos de un mismo tipo compartan alguna estructura interna. Por lo que dice Wiggins (1980, 86-7 nota a pie de página), en el caso de los órganos, parece que habría que entender el *principio de actividad* común más como una conjunción de una cierta morfología parecida y el desempeño de un cierto papel común en los organismos de los que forman parte.

Otro autor que considera que la esencia de los artefactos es una esencia nominal es Schwartz. Sin considerar explícitamente el principio de realizabilidad múltiple, también piensa que las clases artificiales -a las que considera clases nominales- tienen una esencia nominal; esto es, agrupamos a los objetos en clases artificiales siguiendo un criterio definicional que contempla la función y la forma, no un criterio basado en una naturaleza subyacente o composicional que todos los artefactos de una misma clase comparten.

Perhaps the best examples of nominal kind terms are the names of common artefacts such as 'pencil', 'bottle' and 'chair'. Underlying traits play no role in the semantics of such terms. The essence of a nominal kind is not a natural essence, rather nominal kinds, at least in the ideal case, do have a linguistic essence. This linguistic essence is the definition of the nominal kind term. The extension of a nominal kind term *is* determined by an analytical specification of superficial features such as phenomenal properties, and/or form, function, or origin. (Schwartz 1980, 182)

Para Schwartz, la definición o descripción que poseemos de cada clase artificial expresa una esencia nominal o convencional de los tipos de artefactos. Y, en contraposición, la esencia *real* parece que hay que entenderla como una esencia delimitada por la naturaleza. Así que los artefactos no forman clases *reales*³⁴: la clasificación de los diferentes tipos de artefactos no viene determinada por alguna propiedad *natural* (esto es, por alguna propiedad al nivel de la composición material –para Schwartz, al nivel de la composición subyacente), sino que viene establecida *por nuestro criterio*. Según Schwartz, agrupamos a los diferentes artefactos siguiendo básicamente el criterio de la función y la forma, y éstas son características que nosotros imponemos y estipulamos en los artefactos que hacemos. Así, por ejemplo, atribuimos a los lápices la función de servir para escribir y tener una cierta forma, y estas propiedades son el único criterio, estipulado por nosotros, para aplicar el término 'lápiz'. Lo que tenga la función de escribir (junto con una cierta forma), será un lápiz, y lo que no tenga estas características, no lo será. Y la definición que poseemos de 'lápiz' es una definición

³⁴ Schwartz considera que las clases artificiales son clases nominales, y, en este sentido, se oponen a las clases reales en el sentido especificado. Pero no dice nada respecto a si considera que los artefactos (en tanto que tales, no en tanto que trozos de materia) son *reales* en el sentido de existentes.

analítica: expresa una verdad necesaria aplicada a los lápices y es conocida *a priori* (en virtud de ser usuarios competentes). La definición de ‘lápiz’ guía nuestro uso del término y es el único criterio que determina la correcta aplicación de la palabra. Los términos para artefactos son, en este sentido, de acuerdo con Schwartz, como los términos considerados tradicionalmente términos de clase nominal, como ‘soltero’ o ‘arquitecto’, de los que disponemos de una definición analítica.

2.2.2. Crítica a la idea de la esencia nominal de los artefactos

Antes de hacer una crítica a las razones aducidas por estos filósofos que defienden que las clases artificiales son clases nominales, conviene hacer una aclaración. La distinción entre *esencia nominal* y *esencia real* aplicada a los artefactos no es la misma distinción que se aplica a los objetos naturales (ver el apartado **1.3.1**, dedicado a Locke, del **capítulo 1**). En el caso de las clases naturales, se puede defender que éstas tienen una esencia nominal y una esencia real *à la Locke*; en el caso de los artefactos, la discusión se centra en si tienen una o la otra. Es decir, se trata de ver si las clases artificiales son clases nominales, como la clase de los solteros o la clase de los cazadores, cuyas características definatorias son cognitivamente accesibles *a priori* porque las hemos estipulado como definatorias, o son clases reales *en algún sentido análogo* a como se admite que lo son las clases naturales. En la discusión de los párrafos que siguen intento aclarar qué hay que entender por *esencia nominal* aplicada a los artefactos.

La idea de fondo, tanto de Wiggins como de Schwartz, es que las esencias *reales* de las clases sólo pueden corresponderse con algún tipo de composición estructural o subyacente (al nivel material) compartida por los miembros de la clase. Si aceptamos esto, entonces es claro que los artefactos no tienen una esencia *real*. Si tomamos el concepto de esencia *real* tal como fue introducido originariamente para las clases naturales, apuntando a una serie de propiedades microfísicas ocultas compartidas por todos los miembros de la clase, entonces es obvio que las clases artificiales no tienen una esencia *real*. Pero proponer un concepto de clase *real* que, dado el principio de la realizabilidad múltiple, ya excluye de entrada a los artefactos parece un argumento circular contra la idea que las clases artificiales sean clases *reales*³⁵. Hay que notar que

³⁵ Al final de esta sección, vuelvo sobre esta cuestión al exponer el argumento de Thomasson a favor de que las clases artificiales son clases reales.

si aceptamos hablar de *esencia real* y *esencia nominal* entendidas como en el caso de las clases naturales, y rechazamos que los artefactos tengan una *esencia real*, apoyándonos, como hacen Wiggins y Schwartz, en que los artefactos de una misma clase no comparten una estructura material, entonces propiamente tampoco tienen una *esencia nominal* (tal como es entendida en el caso de las clases naturales), pues la función no es una propiedad superficial o perceptible de las cosas (como tampoco lo es, por cierto, la soltería). Hay que aclarar, entonces, en qué sentido se entiende la *esencia nominal* aplicada a los artefactos. Considero que la interpretación más plausible de la idea de *esencia nominal* tal como es aplicada a los artefactos es la idea que la *esencia* de las clases artificiales es algo estipulado por nosotros, como la *esencia* de ser soltero, no una división que marque la naturaleza fuera de los intereses humanos, y algo a lo que tenemos un acceso cognitivo *a priori* en virtud de ser usuarios competentes de los términos para artefactos. Así lo dice Schwartz expresamente. Pero, si esta es la idea de *esencia nominal* que usan estos autores al hablar de los artefactos, tal como parece, y no la idea de *esencia nominal* que se usa al hablar de las clases naturales (como el conjunto de características perceptibles de los objetos naturales y que están causalmente relacionadas con su *esencia real*), entonces no parece sensato que se quiera excluir la *realidad* de las clases artificiales apelando a que su *esencia real* no es como la *esencia real* de las clases naturales.

Me centro ahora, entonces, en la mencionada idea de *esencia nominal* aplicada a los artefactos. Volviendo a Schwartz, éste afirma que los artefactos tienen una *esencia nominal* porque sus características esenciales, la función y la forma, no son características subyacentes microfísicas y disponemos de una descripción analítica, asociada con cada término para artefacto, que expresa una verdad necesaria y conocida *a priori* sobre cada tipo de artefacto. Esta descripción analítica funciona como una definición que conocemos sólo en virtud de ser hablantes competentes y nos proporciona el único criterio, estipulado, de aplicación del término de clase artificial. La primera idea que me parece cuestionable de la propuesta de Schwartz es el hecho que la descripción que asociamos a un término de clase artificial tenga que ser conocida *a priori* como lo es la descripción definicional que asociamos a ‘soltero’. Creo que puede parecer que, como comunidad lingüística, conocemos *a priori* la descripción que asociamos a un término para artefacto porque, una vez que el artefacto y su correspondiente término se han introducido, parece que al menos *alguien* (quizá el

creador del artefacto) tiene que conocer las características relevantes del artefacto nombrado en forma de una descripción. Pero este caso parece un caso diferente a aquél en que estipulamos un significado para un término mediante una descripción, como ocurre con ‘soltero’ o ‘médico’, pues para llegar a poseer la descripción correcta que especifica la forma y la función de un tipo de artefacto, el creador puede necesitar cierto tiempo de ensayo y error, así como hacer ciertos ajustes en sus conocimientos acerca de las características relevantes de un artefacto. O sea que, me parece que el creador de un artefacto puede estipular definitivamente en qué consiste el artefacto que tiene la intención de crear, pero el artefacto, una vez creado, podría no ajustarse a esa estipulación. En este sentido, aun cuando el creador acabe conociendo una descripción que especifica y selecciona sólo la clase artificial en cuestión, el acceso *a priori* a esa descripción no parece tan *inmediato* como el acceso que tenemos a la definición de ‘soltero’ o ‘cazador’. Además, como se ha dicho, es posible hacer descubrimientos sobre nuevas funciones de artefactos conocidos y, en estos casos, las especificaciones de las nuevas funciones no son conocidas *a priori* por nadie.

En definitiva, creo que las clases artificiales no son clases nominales porque los términos para artefactos no son términos de clase nominal como ‘soltero’ o ‘cazador’. El término ‘soltero’ sí tiene una esencia nominal; es, en terminología putnamiana, *one criterion word*³⁶: un término al que asociamos una propiedad estipulada que sirve como único criterio para la aplicación del término. El término tiene una definición analítica que da las condiciones necesarias y suficientes de pertenencia a la clase. Alguien es soltero si y sólo si es un hombre no casado. Pero los términos para artefactos como ‘lápiz’ no son *one criterion words*; pues como se verá a lo largo de este capítulo, puede haber (y hay de hecho en las discusiones de los filósofos) diferentes descripciones asociadas con estos términos. Creo que la diferencia entre el tipo de esencia de un artefacto y la esencia de ser soltero depende en parte de que, pese a imponer nosotros a los artefactos buena parte de las propiedades que presentan (pues los hacemos nosotros), no es tarea obvia decidir cuáles de entre esas propiedades son las que determinan la pertenencia a la clase (muchas de ellas pueden parecer *relevantes* para ser un tipo de artefacto). Nosotros imponemos muchas características en los artefactos: el color, la forma, la función, las medidas, el material, etc. Y no sólo hay que dilucidar cuál o

³⁶ Véase (Putnam 1979, 68).

cuáles de estas características son las que marcan los criterios de identidad de las clases artificiales que se corresponden con nuestras clasificaciones, sino en qué sentido entendemos características como la función³⁷. En el caso de ‘soltero’, estipulamos que ser soltero consiste únicamente en ser un hombre que no esté casado; a ser soltero le imponemos sólo una característica y ésta delimita la esencia. Pero en el caso de los artefactos, no es que estipulemos que un lápiz es todo aquello que sirva para escribir y tenga cierta forma, sino que el proceso parece consistir más bien en que primero hacemos lápices y les ponemos un nombre, sin hacer una estipulación del todo precisa ni tener necesariamente una descripción definitoria asociada y, luego, en algún momento, podemos llegar a una descripción que *capte* la esencia de los lápices. Si Wiggins y Schwartz admiten que la definición analítica asociada a ‘soltero’ proporciona las condiciones necesarias y suficientes para ser soltero, entonces, ‘lápiz’ no es como ‘soltero’ pues, como se verá más adelante, en las secciones **2.2.4** y **2.2.5**, dar las condiciones necesarias y suficientes de pertenencia a una clase artificial puede ser imposible o muy difícil (si no se quiere ir contra ciertas intuiciones). Y, en este sentido, los términos de clase artificial no son como ‘soltero’ o ‘cazador’.

Concluyo, entonces, que las clases artificiales no son como las clases consideradas tradicionalmente nominales, por las siguientes razones: (i) no parece que tengamos un acceso epistémico *a priori* a las características que determinan que un artefacto sea el tipo de artefacto que es como el que tenemos a la característica que determina que un individuo sea soltero, (ii) la descripción que asociamos a un tipo de artefacto no parece fruto de una mera estipulación y (iii) es muy difícil, si es que es posible, dar las condiciones necesarias y suficientes de pertenencia a una clase artificial, como se verá en las secciones que siguen, principalmente en la sección **2.2.4**, pero no así con las condiciones necesarias y suficientes para ser soltero o ser cazador.

2.2.3. La defensa de las clases artificiales como clases no nominales

Lo que más me interesa para los propósitos presentes es descartar que las clases artificiales sean clases nominales en el sentido especificado antes: clases definidas por

³⁷ Más adelante se verá que, para dar las condiciones de identidad de una clase artificial, si se incluye la función entre las características que determinan esas condiciones, entonces hay que precisar qué se entiende por *función*: ¿función actual?, ¿función pretendida?, ¿función potencial?...

una serie de propiedades especificadas en una descripción analítica. Sin embargo, como el debate en torno al tema del estatus de las clases artificiales ha girado en torno a la disyuntiva de si son clases nominales o clases *reales*, otra manera de argumentar contra la idea que los artefactos constituyen clases nominales es defender que forman clases *reales*.

Antes que nada, convendría aclarar qué se entiende por *esencia real* cuando se aplica al ámbito de los artefactos. Creo que se podrían entender al menos dos sentidos diferentes en que las clases artificiales son (o no) clases *reales*. Uno de ellos es considerar que las clases *reales* son aquellas cuyos miembros forman parte de un inventario serio del mundo. Entonces, según este punto de vista, los que defienden que las clases artificiales no son clases reales, defienden que los artefactos (en tanto que tales, más allá del pedazo de materia que son) no existen. Para ellos, en un sentido fuerte metafísico, no hay tales cosas como las mesas y las sillas. Creo que éste es el punto de vista de Wiggins. Pero también se podría considerar que, al decir que las clases artificiales no son *reales*, se está diciendo solamente que su esencia no está constituida por alguna propiedad estructural, objetiva, independiente de la mente, pero sin que esto implique que no haya artefactos. Quizá ésta podría ser la opinión de Schwartz, que está más interesado en destacar el aspecto convencional de los artefactos que en negar su *realidad*. La relación y confusión que se puede dar entre los dos sentidos contemplados de clase *real* seguramente tiene que ver con el hecho que a menudo se ha pensado que aquello que no existe independientemente de la mente no existe.

Hay, sin embargo, estrategias para defender que los artefactos tienen esencias *reales* y forman clases *reales* en los dos sentidos especificados. Así, por ejemplo, contra la idea que los artefactos no son *reales* (aunque sean existentes) porque sus esencias no son independientes de la mente humana, se puede argumentar, como Elder (2009), que los miembros de (al menos algunas) clases artificiales están caracterizados esencialmente por propiedades objetivas independientes de la mente: un cierto diseño, una función propia y un determinado lugar histórico. Aunque los artefactos existan por nuestra intervención, una vez que existen, sus características esenciales son independientes de la mente humana.

Respecto a la idea que los artefactos no forman parte de una ontología seria del mundo y que no existen como tales, Thomasson (2009a) cree que se debe a que, como he comentado antes, a menudo se confunde la dependencia de la mente humana con la no existencia real. Thomasson, a diferencia de Elder, no niega que la esencia de las clases artificiales depende de la mente, pero cree que esto no implica que los artefactos no tengan existencia. Según ella, no debemos aplicar en el campo de los artefactos el criterio realista (en un sentido existencial) tradicional para las clases naturales, el criterio de independencia de la mente. El ámbito natural y el ámbito artificial son ámbitos distintos, con criterios realistas distintos, dice. El criterio realista en general se ha basado, hasta ahora, teniendo como modelo a las clases naturales y, así, se ha estipulado que los únicos objetos reales, existentes, son los que existen con independencia de cualquier intención humana. Este criterio aplicado a los artefactos, obviamente, dará como resultado que éstos no tienen una existencia real y, por tanto, que las clases artificiales no son clases reales. Pero ¿por qué no proponer otro criterio realista para estas clases, defendiendo así el sentido común, que nos dice que sí existen? Thomasson propone un nuevo criterio que incluye tanto a los artefactos en tanto que tales como a los objetos naturales entre los objetos reales, considerando a las clases artificiales, por tanto, clases reales. Este nuevo criterio realista consiste, muy brevemente, en aceptar como real todo aquello cuyas condiciones de existencia parecen razonables por todo lo que sabemos. Así, por ejemplo, las condiciones de existencia de las sillas, según Thomasson, consisten en que algún o algunos creadores hayan querido hacer sillas y lo hayan conseguido. Parece, efectivamente, que no hay razones para pensar que esto no ha sido así³⁸.

Mi intuición es tanto que los artefactos son existentes como que su esencia o naturaleza, aunque dependiente de la intervención humana para su existencia, tiene un estatus objetivo. Respecto a la existencia de los artefactos, quiero notar que si el motivo por el que se considera que no hay artefactos es, como dice Thomasson, que no son independientes de la mente, entonces colocamos a los artefactos en el mismo saco que colocamos a los unicornios o a las brujas. Y parece que hay una diferencia intuitiva

³⁸ El criterio realista de Thomasson excluye a los unicornios o las brujas como objetos reales, pues ¿qué supondría que hubiese brujas, por ejemplo? Supondría que hubiese unas mujeres con ciertos poderes sobrenaturales por haber hecho un pacto con el diablo. ¿Es razonable pensar que se han dado o se dan estas condiciones de existencia? No.

importante entre las sillas, por ejemplo, y las entidades que claramente diríamos que no son reales o existentes. Así que, si queremos conservar ciertas distinciones que parecen que tendrían que mantenerse, la idea de la dependencia de la mente no es un buen criterio (como dice Thomasson) para la irrealidad de los objetos. Por otro lado, hay ciertas entidades que consideramos plenamente reales, existentes, como los lenguajes o ciertas propiedades secundarias, como el color, que tampoco serían reales según el criterio de dependencia de la mente. En definitiva, considerar que los artefactos no tienen una existencia real porque su existencia depende de nosotros mete en el mismo saco a los artefactos y a objetos que intuitivamente diríamos que no existen, y excluye también a otras entidades de la realidad cuando diríamos que son reales.

Respecto a la realidad de los artefactos en el sentido de su esencia objetiva, me parece que se podría defender que así como ciertas razas de perro dependen, para su existencia, de la intervención humana, pero hay criterios objetivos, independientes de la mente, para determinar si un cierto perro es o no de una raza, creo que también se podría considerar que los artefactos tienen una esencia o naturaleza objetiva, independiente de la mente. Intuitivamente, creo que esa naturaleza incluirá la función (quizá en el sentido de *función potencial*³⁹) y seguramente otras características (que quizá habría que especificar en cada caso). De todos modos, en el presente trabajo no me comprometo con una determinada caracterización metafísica de los artefactos. Más que destacar que los artefactos puedan constituir clases *reales* en algún sentido interesante, lo que me interesa es destacar que las clases artificiales no son clases nominales, o al menos no son clases nominales como las clases tradicionalmente consideradas así. Ahora bien, si el sentido en que considero que las clases artificiales no son clases nominales implica que son clases reales, entonces creo que se puede argumentar, siguiendo alguna de las estrategias sugeridas, que las clases artificiales forman clases reales.

2.2.4. Las condiciones necesarias y suficientes de pertenencia a las clases artificiales

Una diferencia importante entre las clases artificiales y las clases típicamente consideradas clases nominales, como la clase de los solteros o de los cazadores, es que de las segundas, a diferencia de las primeras, sí poseemos descripciones que nos

³⁹ Véase la sección 3.4.1 del capítulo 3.

proporcionan las condiciones necesarias y suficientes de pertenencia a la clase. Podría parecer que éste también es el caso para las clases artificiales, pero no es así. Voy a repasar a continuación y muy brevemente las principales propuestas que se han dado sobre la esencia o identidad de los artefactos en tanto que tales, y cómo esas propuestas no proporcionan las condiciones de identidad de las clases artificiales.

Dejando de lado el asunto de si los artefactos tienen esencias reales o nominales, hay dos aproximaciones básicas a la cuestión de qué constituye la esencia de los artefactos (sea una esencia nominal o una esencia real): la función y la intención. Para unos, la característica esencial más importante de los artefactos es su función; según otro punto de vista, la función juega un papel secundario y la esencia de los artefactos depende de nuestras intenciones. Empiezo con los filósofos que defienden una esencia funcional para los artefactos. Entre ellos están, como hemos visto, Wiggins y Schwartz. Es una opción que se corresponde con lo que diríamos intuitivamente sobre qué es lo que caracteriza a los diferentes artefactos. Pero ya Schwartz, sin desarrollar más el tema, se da cuenta de que hay que apelar a otras características aparte de la función; si no, ¿cómo distinguimos un lápiz de un bolígrafo, puesto que tienen la misma función?, ¿o una silla de un taburete? Parece, entonces, que hay que apelar también, para concretar en qué consiste la esencia de los tipos de artefactos, a otras características como la forma además de la función. Esto es lo que hace Schwartz. Pero si se pretenden dar las condiciones de identidad de las clases artificiales, y, de paso, dar cuenta de nuestro uso de los términos de clase artificial, una propuesta como la de Schwartz no bastará, pues, por ejemplo, un cuchillo roto sigue siendo, intuitivamente, un cuchillo y le aplicamos el término correspondiente. Así, la definición que asociamos a ‘cuchillo’, aunque incluya referencia a la forma y distinga los cuchillos de las navajas, no proporciona condiciones necesarias de aplicación del término, pues un cuchillo roto no tiene la función (quizás tampoco la forma) que especifica la supuesta definición. No sólo los miembros no funcionales de una clase artificial ponen en apuros a la definición que pretende Schwartz, sino también los miembros funcionales pero que pueden tener formas muy distintas de la estándar (como, por ejemplo, las grapadoras, que pueden tener formas muy distintas).

Abundando en la idea de añadir a la función de un artefacto otras características que den cuenta de sus condiciones de identidad, Carrara y Vermaas (2009) inspiran otra solución

al asunto. Como la función de un cierto tipo de artefacto es insuficiente para seleccionar la clase artificial en cuestión (la función de servir como asiento no distingue entre la clase de las sillas y la clase de los taburetes, por ejemplo), ellos proponen añadir otras características que singularicen la clase artificial. Esas otras características podrían entenderse de diferentes maneras, aunque incluyendo alguna propiedad de tipo *estructural* en algún sentido. Su propuesta concreta es que esas características que hay que sumar a la función pueden ser el *principio operacional* y la *configuración normal*. Dicho muy brevemente, el principio operacional es la manera en que funciona el artefacto y la configuración normal es la forma del artefacto. Poniendo un ejemplo de Lawler y Vega (2010) para ilustrar la idea, la función de un sacacorchos de tirabuzón es sacar corchos de botella, su principio operacional es enroscando el tirabuzón y su configuración normal es la forma que tiene y que identifica una determinada estructura causal. Si a la conjunción de estas características junto con la función del artefacto le damos el nombre de ‘función+’, entonces la idea es que la función+ sí singulariza cada clase artificial. Así, por ejemplo, especificando, junto a la función, el principio operacional y la configuración normal, la idea es que se pueda distinguir entre un cuchillo y una navaja, o entre una silla y un taburete. La esencia de los artefactos, entonces, es la función+, donde ‘+’ quiere incluir a las características que sumadas a la función sí singularizan las clases artificiales.

Pero, al seguir esta estrategia, se restringe en alguna medida el cumplimiento del principio de realizabilidad múltiple por parte de los artefactos, pues se evitaría que artefactos con una misma función, y que consideramos de la misma clase, pero con una diferente configuración normal, perteneciesen a una misma clase artificial. Lawler y Vega han destacado este problema de la propuesta de Carrara y Vermaas diciendo que quizás hace más distinciones que las que queremos admitir. Así, por ejemplo, el sacacorchos de láminas y el sacacorchos de tirabuzón, que funcionan con principios operacionales y tienen configuraciones diferentes, constituirían diferentes clases artificiales. Sin embargo, no parece que agrupemos a los diferentes tipos de sacacorchos en diferentes clases artificiales. Pero aun aceptando que, en algunos casos, puede tratarse de clases distintas que por comodidad agrupamos bajo el mismo término, está también el problema, dicen Lawler y Vega, de que cualquier artefacto de una clase dada con una forma lo *suficientemente original* o diferente constituiría una clase artificial con un solo miembro, pues tendría una configuración única. Según la propuesta de Carrara y

Vermaas, dependiendo de lo que entendamos por *configuración distinta*, no es difícil que proliferen las clases artificiales con un solo miembro. Y esta multiplicación innecesaria de las clases naturales no parece deseable ni corresponderse con el uso de nuestros términos para artefactos.

Además, otro inconveniente de la sugerencia de Carrara y Vermaas sobre la esencia de las clases artificiales es que, como en el caso de Schwartz, los artefactos no funcionales (por la causa que sea) no serían propiamente artefactos, pues no tendrían ninguna función. Si entendemos cualquier noción de función, sea función+ o no, como la función *actual* que tienen los artefactos, entonces, no podemos dar cuenta de que una silla rota sea una silla, o un reloj que ha dejado de funcionar (o que nunca haya funcionado) sea un reloj. La noción de *función* sin más, sin ulterior aclaración, no va a dar las condiciones de identidad de una clase artificial (aunque añadamos otras características). ¿Cómo hay que entender, entonces, el concepto de *función* para que, entre otras cosas, dé cuenta de los artefactos que, presumiblemente, pertenecen a la clase y no son funcionales?

Ante las carencias que presenta la noción de *función* entendida como función actual, Millikan (1989) propone entender el concepto de *función* como función *propia*. Brevemente, la *función propia* es la función que tiene un artefacto *porque (a causa de que)* ítems previos la han tenido. Es decir, es la función que habitualmente atribuimos a los artefactos pero considerada históricamente: los objetos tienen la función que tienen porque la *heredan*, son reproducidos o copiados a partir de objetos predecesores en el tiempo a causa de que éstos realizaban bien esa función. Los artefactos pueden *heredar* su función propia de la función pretendida para ellos por su creador (caso de los prototipos o artefactos nuevos) o pueden *hacerla* a partir de ítems previos que tenían esta función (caso de los artefactos copiados). La noción de función propia de Millikan supera algunas objeciones que se pueden hacer a la simple noción de función. Para empezar, la noción de función propia distingue entre *tener una función* y *ser usado para*, evitando así atribuir a las funciones accidentales o impropias que pueden tener los objetos el papel de función. Por ejemplo, un destornillador puede ser usado esporádicamente como un martillo, o un tronco como un remo, pero la función (propia) del destornillador no es servir como martillo ni la de un tronco servir como remo pues no han sido producidos a causa de que otros destornilladores y otros troncos realizaran

bien esas funciones. La noción de función propia también da cuenta de que los miembros rotos o malformados de una clase pertenecen a la clase porque, como los miembros funcionales, también fueron producidos a causa de que miembros anteriores históricamente realizaban aquella función. Otra cosa es que, por un accidente o por la imperfección de los mecanismos de reproducción de los artefactos, algunos acaben siendo no funcionales.

La noción de función propia de Millikan, pese a presentar las ventajas comentadas respecto a la noción *simple* de función, tiene el problema de no explicar convenientemente la aparición de funciones nuevas, tanto en el caso de los prototipos o artefactos nuevos con funciones genuinamente nuevas, como en el caso de nuevas funciones que se descubren en artefactos conocidos. En el caso de artefactos nuevos con funciones nuevas, la explicación de la función propia ya no puede apelar a la dimensión histórica de la función y debe decir que la función nueva es transmitida *a través de* la intención del creador. Aparte de que esto supone no disponer de una explicación unificada de la adquisición de las funciones por parte de los artefactos, tiene el inconveniente de que no tiene en cuenta que las intenciones de los usuarios al usar el artefacto pueden divergir de las intenciones del creador. En el caso del descubrimiento de nuevas funciones en artefactos ya conocidos, no está claro cómo la teoría puede explicar el fenómeno. Además, la noción de función propia no proporciona, tal como critica Thomasson (2009a), ni las condiciones necesarias ni suficientes de pertenencia a una clase artificial. La igualdad de función propia (la función pretendida) puede no ser suficiente para ser un miembro de una misma clase artificial, ya que otros factores como la forma, el origen, etc. también pueden ser necesarios. Y puede no ser necesaria porque algo puede ser una silla aunque la función pretendida por su creador no sea que la gente se siente encima.

La otra gran alternativa a la concepción funcional para explicar la esencia de los artefactos, es la concepción intencional sobre los artefactos. Thomasson, ante la insuficiencia de la noción de función, descarta cualquier noción de *función* como criterio para la pertenencia a una clase artificial porque no proporciona las condiciones necesarias y suficientes para ser un miembro de la clase. Thomasson (2003, 2009a)

ofrece un criterio alternativo que, según ella, al menos da las condiciones necesarias⁴⁰. La esencia de un tipo de artefactos no es la función actual ni la función pretendida por el creador para esos artefactos, sino una serie de características (que pueden incluir la función o no) que el creador ha impuesto en el artefacto con la intención (exitosa) de hacer un artefacto de un cierto tipo. Dicho de otra manera, para que, por ejemplo, un artefacto sea una silla es necesario que su creador haya tenido la intención de hacer una silla (ajustándose al concepto de los creadores previos de sillas o introduciendo un nuevo concepto de artefacto si se trata de la primera silla). El concepto del creador sobre las características generales que ha de tener un artefacto (si la función, la forma, el material, etc. o una conjunción de diferentes propiedades) sí específica, dice Thomasson, al menos, condiciones necesarias para que el artefacto sea el tipo de artefacto que es. Los conceptos de los creadores de los artefactos, especifiquen o no una función pretendida, dice Thomasson, son lo más relevante para determinar si su producto pertenece o no a una clase artificial dada.

Pero la propuesta de Thomasson tampoco parece que proporcione siempre las condiciones necesarias de pertenencia a una clase artificial. Supongamos que un creador produce un nuevo tipo de artefacto guiado por un concepto que especifica una determinada función, pero los usuarios le dan otro uso, de manera que esta segunda función es la única que los usuarios le atribuyen y la única que acaba desempeñando. En un caso así, el concepto del creador, que incluía una referencia a la función del artefacto en cuestión, no proporciona, intuitivamente, las condiciones necesarias de pertenencia a la clase. Y aunque un artefacto pueda tener dos funciones diferentes, la función pretendida por el creador y la función que le otorgan los usuarios, ésta última no es fruto de las intenciones del creador y puede ser que sea la que acabe determinando la pertenencia de un artefacto a su clase artificial. Otros contraejemplos a la teoría de Thomasson provienen de casos en que un creador acaba creando, sin pretenderlo, un cierto tipo de artefacto. Por ejemplo, un creador podría pretender hacer un reloj de juguete, una imitación de un reloj *auténtico*, pero que, sin quererlo, resultara que funciona como un reloj *de verdad*. En casos de este tipo, Thomasson tiene que decir que

⁴⁰ En el siguiente capítulo, el **capítulo 3**, al considerar las teorías descriptivistas más destacadas y desarrolladas sobre los términos de clase artificial, explicaré con más detalle la teoría de Thomasson, tanto en su vertiente metafísica, como epistemológica y semántica.

el creador no ha hecho un reloj (pues no ha tenido una intención, guiada por un concepto sustantivamente correcto, de hacer un reloj), pero creo que la intuición nos dice que el objeto creado azarosamente es un reloj. Por último, creo que, igual que en el caso de la propuesta de Millikan, la posibilidad de descubrir nuevas funciones en los artefactos hace que los conceptos de los creadores no proporcionen siempre condiciones necesarias para que un artefacto sea miembro de una clase.

Parece que un concepto apropiado de *función* para delimitar las clases artificiales, además de incluir a los artefactos rotos y excluir a aquéllos con usos accidentales para fines *impropios*, debería también dar cuenta de los artefactos genuinamente nuevos, así como de los artefactos ya conocidos con nuevas funciones descubiertas. La noción de *función propia* de Millikan, por ejemplo, explica bien la dimensión histórica que tienen los artefactos, pero no quizá su aspecto causal, que es lo que permite la novedad funcional. Thomasson también tiene problemas para dar cuenta de los prototipos o artefactos nuevos, pues aquí el concepto del creador no puede ajustarse al concepto de otros creadores previos, tiene que ser un concepto novedoso al que no está claro cómo accede el creador (no puede acceder antes de haber creado el artefacto, pero, por otro lado, ese concepto debe guiar el proceso de producción del artefacto). Y la intención del creador puede ser diferente de las intenciones de los usuarios, de manera que lo que determine la pertenencia a la clase de un cierto tipo de artefactos sea el concepto que tienen los usuarios antes que el concepto que tiene el creador.

2.2.5. Alternativa a las condiciones necesarias y suficientes

En conclusión, dar las condiciones necesarias y suficientes de la pertenencia a una clase artificial particular para todo objeto es una tarea muy difícil, si no imposible. Incluso especificar sólo las condiciones necesarias no es fácil. Aunque realmente el problema está sólo para aquéllos que postulan que los artefactos tienen una esencia nominal que nos resulta cognitivamente accesible o para aquéllos que defienden que, aunque lo que caracteriza a las clases artificiales no es una mera esencia nominal, tenemos que conocer las propiedades que determinan la pertenencia a la clase para cualquier artefacto.

Ante esta situación, podría sugerirse que lo mejor es renunciar a la idea que los artefactos se agrupan formando clases (*kinds*), si entendemos que pertenecer a una clase es cumplir una serie de condiciones necesarias y suficientes. Lawler y Vega (2010), por

ejemplo, proponen renunciar al lenguaje de los artefactos en términos de clases, y hablar meramente de parecidos de familia *à la Wittgenstein* entre los artefactos que consideramos de un mismo tipo. La idea wittgensteiniana sería que no hay ninguna propiedad⁴¹ o serie de propiedades compartidas por todos los artefactos de un tipo y que proporcionen condiciones necesarias y suficientes de pertenencia a la clase; por ejemplo, no todas las mesas presentan las mismas características (unas tienen cuatro patas, otras tienen tres patas, algunas son redondas, otras son cuadradas, etc.). Resulta, entonces, que aplicamos ‘mesa’ a objetos que se parecen entre sí en algunos rasgos, pero que también divergen entre ellos en algunos aspectos. Una mesa se puede parecer a otra en unas ciertas características que pueden ser diferentes de las características que comparte con una tercera mesa. En estos casos, sólo podemos hablar de un *parecido de familia* (sin poder proporcionar una definición) entre los diferentes objetos a los que aplicamos el término ‘mesa’. ¿Cuál es la guía, entonces, dado que carecemos de una definición, para aplicar el término? Según Wittgenstein (2004), aplicamos el término en la medida en que el objeto al que lo aplicamos se parece a un caso paradigmático. Así, por ejemplo, una mesa es aquello que está conectado por una cadena de propiedades con otras mesas hasta llegar a una mesa paradigmática. Según esta propuesta wittgensteiniana, la condición requerida para la aplicación de un término como ‘mesa’ es que exista una cadena de propiedades que conecte el objeto con un caso paradigmático.

Esto enlaza con la idea de Putnam, expuesta en la sección **1.5.3** del **capítulo 1**, que lo que sea una muestra de agua (o de cualquier otra clase natural) se define en relación a su similitud con un ejemplar paradigmático. Así que creo que se puede interpretar que la idea wittgensteiniana comparte un cierto elemento con el externismo defendido por Putnam, pues la aplicación de los términos considerados no está guiada por una descripción que tengamos en mente, sino por propiedades de las cosas que no necesariamente sabemos especificar (propiedades que, para Putnam, podrían no resultarnos cognitivamente accesibles). Siendo estrictos, este punto de vista también tendría que extenderse a los objetos naturales, pues no está claro que haya clases naturales delimitadas por condiciones necesarias y suficientes. Las descripciones que asociamos a los géneros naturales no proporcionan sus condiciones de identidad (pues,

⁴¹ Ninguna propiedad no trivial, pues se excluye la propiedad trivial que compartirían todas las mesas de ser llamadas ‘mesa’.

por ejemplo, un tigre puede tener tres patas o podría haber un robot con todo el aspecto de un tigre). Pero tampoco está claro que los criterios que proporciona la ciencia sean condiciones necesarias y suficientes para la pertenencia a una clase natural. Si algo es agua, entonces es H₂O; sin embargo no está claro que si algo es H₂O, entonces es agua (pues no está claro, por ejemplo, que una sola molécula de H₂O sea agua). Y con las clases biológicas el criterio científico aún está más alejado de proporcionar las condiciones de identidad de las especies: parece que tener un cierto genoma no es ni necesario ni suficiente para ser un tigre. La idea de los *parecidos de familia* entre los miembros de un mismo tipo de artefacto es que agrupamos a los artefactos siguiendo un criterio que no podemos especificar con precisión pues no hay condiciones necesarias y suficientes para ser un miembro del grupo (no hay ningún conjunto de características que compartan todos los individuos del grupo). Ante la falta de un tal criterio para agrupar los objetos y aplicar los términos correspondientes, una explicación de cómo contrastamos el *parecido de familia* es que tomamos algunos casos concretos como casos paradigmáticos, y todo lo que tenga un cierto aire de familia respecto a ese paradigma (en los aspectos *relevantes*) es clasificado en un mismo grupo.

Pero ahora la cuestión es respecto a qué propiedad de las cosas consideramos si presentan o no un parecido de familia o una cierta similitud. En definitiva, ¿qué aspectos son los *relevantes* para que consideremos que dos objetos dados son similares, dado que dos objetos cualesquiera se parecerán entre sí en muchos aspectos? La idea de Putnam es que la similitud entre dos objetos es una similitud según *nuestros intereses*. Donnellan (1983) profundiza en esta cuestión de los aspectos compartidos por los miembros de una misma clase y que son *relevantes según nuestros intereses*. Para Donnellan, los objetos naturales que consideramos de una misma clase se parecen entre sí en muchos aspectos y divergen en muchos otros, y, al clasificarlos como de la misma clase natural, estamos priorizando como más *importantes para nosotros* algunas de las características que comparten sobre otras. La idea de Donnellan es que podríamos haber elegido otras características como las *relevantes* para nuestros intereses clasificatorios⁴².

⁴² Esto no es decir que nuestras clasificaciones no responden a criterios objetivos, pues, como se verá, las propiedades que priorizamos son propiedades objetivas de las cosas, con poderes causales.

Para ver esto, propone un experimento mental. Consideremos dos culturas gemelas (con lenguas y ciencias idénticas) antes del surgimiento de la ciencia química. En un momento dado se desarrolla indistinguiblemente en ambas culturas la teoría atómica⁴³. En ambas culturas notan que un mismo elemento químico puede tener, y generalmente tendrá, varios isótopos diferentes y les dan nombres. Pero hay una diferencia de cantidad: en la Tierra o Cultura Gemela, además, es habitual que uno de los isótopos de un elemento particular constituya la mayor parte del elemento en la naturaleza, siendo los otros isótopos muy raros. Y es a menudo el caso que algunos isótopos de un elemento son radioactivos y otros no, y esto puede tener graves consecuencias. En tales circunstancias, podría ser una cuestión muy difícil, dice Donnellan, decidir si es más importante el número de isótopo o el número atómico. En la Tierra Gemela, quizá primando las graves consecuencias de la radioactividad, consideran *más importante* el número del isótopo de una sustancia química antes que su número atómico, e identifican la sustancia designada por alguno de sus términos de clase natural con el isótopo que constituye el grueso de lo que había sido llamado por ese término antes del desarrollo científico. Ellos, por ejemplo, no identifican el oro con el elemento que tiene número atómico 79, sino con un cierto isótopo teniendo un cierto número de isótopo. Así, la extensión de ‘oro’ es diferente en la Tierra y en la Tierra Gemela. Y la diferencia en la extensión del término parece que viene determinada, dice Donnellan, por una diferencia en los respectivos intereses a la hora de considerar las cualidades importantes de la clase natural (puesto que las propiedades objetivas de las cosas son las mismas).

Es concebible, entonces, otro criterio de clasificación de los géneros naturales. Así, de acuerdo con Donnellan, de la propuesta de Putnam se seguiría que la extensión de los términos de clase natural estaría determinada, no sólo por la naturaleza de los objetos naturales -por ciertas cualidades objetivas que la ciencia puede descubrir- sino también

⁴³ Según tal teoría, los átomos tienen un número particular de protones en el núcleo (su número atómico). Y, como también hay isótopos (núcleos atómicos con el mismo número de protones que otro átomo pero con un número diferente de neutrones), el número de isótopo es el número combinado de protones y neutrones en el núcleo del átomo.

por ciertos intereses⁴⁴ de los hablantes al considerar unas características más importantes que otras.

Resumiendo, una concepción alternativa, tanto para las consideradas clases naturales como para los artefactos, es considerar que los miembros de un mismo tipo comparten entre sí un cierto *parecido de familia* que no se puede especificar mediante condiciones necesarias y suficientes de pertenencia a la clase. Este parecido de familia se reduce, en última instancia, a una similitud con un cierto ejemplar paradigmático del tipo en cuestión. Pero *similitud*, ¿en qué aspecto? De acuerdo con Putnam, la similitud es respecto a una cierta característica, o características, que *nosotros* consideramos *importantes* según nuestros intereses. Así, la referencia de los términos de clase artificial y de clase natural vendría determinada tanto por ciertas propiedades objetivas de las cosas como por nuestros intereses (condicionados, a su vez, si se quiere, por nuestro modo de ser). Pero, entonces, dice Donnellan, la extensión de nuestros términos generales podría haber sido diferente, si hubiésemos tenido intereses diferentes. Yo creo que, si el experimento mental de Donnellan es coherente, el condicionamiento de nuestros intereses, en la determinación de la referencia de nuestros términos, no socava el externismo semántico. Pues el condicionamiento de nuestros intereses deriva del entorno en el que nos encontramos y cómo responde ante él nuestra naturaleza humana. Y todos estos factores pueden escapar a nuestro acceso cognitivo de manera que podemos carecer de un acceso cognitivo tanto a una descripción que especifique las condiciones necesarias y suficientes de pertenencia a una clase, como podemos carecer de acceso cognitivo a una descripción de nuestros intereses.

2.3. La relación entre la metafísica y la semántica

2.3.1. La relación entre la metafísica de los artefactos y la semántica de los términos de clase artificial

Para algunos autores descriptivistas respecto a los términos de clase artificial, como Schwartz y Thomasson, la metafísica de las clases artificiales condiciona o tiene

⁴⁴ Intereses que también tendrían que ver con la naturaleza, pero, en este caso, con la naturaleza o modo de ser de los hablantes, de quienes priorizan ciertas características objetivas sobre otras.

consecuencias sobre la semántica de los términos correspondientes. En el caso de Schwartz (1978, 1980, 1983) esta asunción es implícita, al asumir que los términos para artefactos responden a una misma semántica que los términos considerados tradicionalmente de clase nominal porque los artefactos de un mismo tipo no comparten una naturaleza subyacente; es decir, porque las clases artificiales no son como las clases naturales. En Thomasson, en cambio, la idea de la influencia de la metafísica de los artefactos en la semántica de los correspondientes términos es explícita:

[...] it is not just artifacts, but the natures of artifactual kinds themselves, that are, in some sense, human creations. This metaphysical point in turn has consequences for epistemology and reference—putting at least some humans in a much closer epistemic relation to the artifactual kinds they create and sustain than anyone can be in to the natural kinds they may study, and entailing that the reference of artifactual kind terms must, at some level, be determined by way of human concepts. (Thomasson 2009a, 73)

La idea muy a grandes rasgos de estos filósofos es que, tanto para existir como para ser lo que son, los artefactos dependen de nosotros, bien de ciertas intenciones (en el caso de Thomasson), bien de nuestras estipulaciones (en el caso de Schwartz). El siguiente paso es derivar una consecuencia de tipo epistemológico: como lo que son los artefactos depende de nosotros, tenemos un acceso cognitivo privilegiado a su esencia, a aquello a lo que pretendemos referir con el correspondiente término. Por tanto, concluyen estos autores, cuando usamos un término de clase artificial, necesariamente tenemos un acceso cognitivo a las propiedades que determinan la extensión del término, y el externista se queda sin las posibilidades de construir argumentos de ignorancia y error respecto a la naturaleza de los artefactos.

En el caso de una postura como la de Schwartz o Wiggins, es claro que la metafísica que proponen para los artefactos condiciona el acceso epistémico que tenemos a ellos, pues si la esencia de los artefactos consiste en lo que convencionalmente decidimos que son, entonces tenemos un acceso *privilegiado*⁴⁵ a lo que son. Por ejemplo, lo que sea un lápiz es algo estipulado por nosotros y, por tanto, será un lápiz sólo aquello que se

⁴⁵ Este acceso *privilegiado* supone un acceso infalible.

ajuste a la definición de 'lápiz'. Tenemos un acceso cognitivo privilegiado, incorregible, a lo que es ser un lápiz o a lo que es ser una silla, del mismo modo que tenemos un acceso cognitivo privilegiado a lo que es ser soltero. Así, concluye un descriptivista como Schwartz, no podemos desconocer la esencia o propiedades relevantes de los artefactos que hacemos y, de este modo, los argumentos de ignorancia y error se vuelven imposibles (imposibilitando o dificultando mucho el externismo semántico), del mismo modo que es imposible, para todo hablante competente, ignorar o estar equivocado sobre lo que es ser soltero.

En el caso de Thomasson, el salto de la metafísica al campo epistémico es parecido: lo que son los artefactos depende de intenciones humanas (las intenciones del creador), de manera tal que el creador de un artefacto posee un concepto infalible sobre las características relevantes para que el producto de su creación sea el tipo de artefacto que es. Aquí también hay un privilegio epistémico, al menos por parte de algunos miembros de la comunidad lingüística (los productores de artefactos) que imposibilita la ignorancia o el error *masivo*, por parte de estos hablantes, sobre las características esenciales de los artefactos. Esto, creo que impide que el externismo pueda dar cuenta de cómo se determina la referencia de nuestros términos para artefactos, pues el externismo semántico para tales términos depende de que nuestro acceso cognitivo a las características de los artefactos que determinan la referencia de los correspondientes términos pueda estar sujeto a la ignorancia o el error sobre esas características. Y si estas posibilidades están descartadas de entrada, el externismo pierde al menos buena parte de su motivación. El externismo, para mostrarse plausible ante el descriptivismo, necesita tomar la descripción que según el internista asociamos con el término y que nos conduciría hasta el referente del término y mostrar que es posible que esa descripción esté equivocada o sea insuficiente para seleccionar la extensión del término.

El problema básico de las dos posturas metafísicas sobre los artefactos y del paso que dan, a partir de aquí, a la infalibilidad epistémica respecto a lo que son los artefactos, es que consideran que atribuir una función a un artefacto equivale a que el artefacto realmente tenga esa función. Pero, desde mi punto de vista, podríamos equivocarnos sobre las propiedades que atribuimos a un artefacto. Ya he señalado antes que, por ejemplo, la descripción que un creador puede tener del artefacto que quiere crear no tiene por qué corresponderse con la descripción a la que realmente se ajusta el producto

final que ha creado. El razonamiento que parece estar detrás de la exclusión del error o la ignorancia sobre la función o las características relevantes de los artefactos es que éstos tienen necesariamente la función que les atribuimos porque así lo hemos estipulado. Pero este razonamiento incurre en petición de principio, pues precisamente se trata de ver si efectivamente conocemos o no de manera infalible su naturaleza (aquello a lo que el término refiere). Desde mi punto de vista, el externista no niega que pueda haber descripciones asociadas a los términos para artefactos, sino que, como he destacado en el **capítulo 1**, niega que sean descripciones inmunes a la ignorancia o el error sobre las propiedades que caracterizan a los artefactos (sean éstas las que sean). Entonces, el reto del externista es mostrar que, en alguna medida, podríamos ignorar o equivocarnos respecto a la naturaleza (que, presumiblemente, incluye la función) que atribuimos a los artefactos. En el **capítulo 4** desarrollo esta cuestión y explico el alcance y el tipo de argumento que puede construir un externista para debilitar la posición descriptivista respecto a los términos para artefacto.

En definitiva, lo que creo que imposibilita el externismo semántico es la tesis epistemológica que dice que nuestro acceso cognitivo a los artefactos es tal que no podemos ignorar ni equivocarnos sobre su naturaleza. Es importante notar que, sin afirmar una tesis equivalente, me parece que, al menos las principales posturas metafísicas sobre los artefactos, no van a excluir, *per se*, el externismo semántico para los términos de clase artificial, pues el externismo es una teoría semántica que, de entrada, no se compromete con una postura metafísica determinada. Para ver esto basta ver que el externismo es compatible con diferentes concepciones metafísicas sobre los artefactos: puede aceptar una naturaleza funcional (en cualquiera de las versiones de esta concepción) de los artefactos, o aceptar una naturaleza determinada por las intenciones humanas (asumiendo, eso sí, que esas intenciones no son inmunes a la ignorancia o el error).

Lo crucial desde un punto de vista externista para los términos de clase artificial es que la referencia de tales términos se determine por características del mundo, características de los artefactos, que podrían resultarnos cognitivamente inaccesibles. Sea la función (entendida del modo que se considere apropiado) o sean las características que intencionalmente pretende atribuir el creador aquello a lo que

pretendemos referir con nuestros términos de clase artificial, lo importante, para el externista, es que podríamos equivocarnos o ser ignorantes sobre esas características.

Considero que la única postura incompatible con el externismo semántico sobre los términos de clase artificial es la postura que defiende que las clases artificiales son clases nominales. En efecto, los que defienden esto, como Schwartz o Wiggins, van más allá de afirmar una naturaleza funcional de las clases artificiales y, como expuse en **2.2.1**, suponen o afirman también la tesis de tipo epistemológico según la cual tenemos un conocimiento infalible sobre la naturaleza funcional de los artefactos (pues son creaciones nuestras) y conocemos una definición analítica de los términos para artefactos. Si ‘los lápices son los objetos (*con una cierta forma*) que tienen la función de servir para escribir’ es una verdad analítica (verdadera sólo en virtud de su significado y que expresa una verdad necesaria y cognoscible *a priori*), entonces, creo que esta postura acerca de las clases artificiales es incompatible con el externismo, ya que, desde esta concepción, no veo cómo son posibles los argumentos de ignorancia o error sobre los lápices. Así que me parece que el externismo para los términos para artefactos es incompatible con la postura metafísica que defiende que las clases artificiales tienen una esencia nominal, siempre y cuando se entienda que eso sólo significa que los términos de clase artificial son términos de clase nominal como ‘soltero’.

Más allá de esto, no creo que el externismo se comprometa con ningún tipo de esencialismo⁴⁶ o antiesencialismo respecto a las clases artificiales, y es compatible con ambos. Así, un externista respecto a los términos para artefactos podría defender, siguiendo su postura metafísica favorita, que lo característico de cada clase artificial es su función, pongamos por caso, y que la función constituye la esencia de los artefactos: si los lápices tienen la función de servir para escribir, entonces necesariamente tienen la función de servir para escribir. Lo característico de una postura externista es la posibilidad del error respecto a la función que les atribuimos; en este caso, la posibilidad del error respecto a la función de servir para escribir. Y esta posibilidad de error o ignorancia respecto a lo característico de un tipo de artefacto es posible, aunque

⁴⁶ No voy a distinguir entre esencialismo *de clase* y esencialismo *individual* (distinción que, entre otros hace Devitt (2005) y que he mencionado en **1.5.3**, p. 44, nota a pie de página), pues creo que lo que digo se aplica indistintamente a las dos nociones de esencialismo.

se trate de características esenciales a los artefactos, en la medida en que sea posible *epistémicamente* una tal situación de ignorancia o error⁴⁷. Pero un externista también podría optar por un antiesencialismo⁴⁸ respecto a las clases artificiales: podría defender que los lápices se caracterizan por su función de servir para escribir, pero esta propiedad es contingente a los lápices (es decir, los lápices no tienen necesariamente la función de servir para escribir; en otro mundo posible, podrían tener otra función). Le basta con que esa propiedad contingente, a la que apunta nuestro término ‘lápiz’, pueda ser desconocida o nos podamos equivocar sobre ella.

Incluso el internismo descriptivista respecto a los términos de clase artificial también es independiente de la teoría metafísica sobre los artefactos. Como se ha visto, el descriptivismo no obliga a adoptar una determinada postura metafísica: diferentes autores descriptivistas han optado por diferentes explicaciones metafísicas sobre los artefactos (algunos han optado por una explicación intencional; otros, por una explicación metafísica funcional). Con lo que sí creo que se compromete el descriptivista semántico, a nivel metafísico, es con proporcionar *alguna* explicación sobre la naturaleza de los artefactos, pues afirma que tenemos un acceso cognitivo a tal naturaleza mediante una descripción. Por otro lado, el descriptivismo para los términos de clase artificial también es compatible tanto con posturas esencialistas como con posturas antiesencialistas sobre la naturaleza de los artefactos. Ciertamente, un descriptivista semántico para los términos de clase artificial puede defender un esencialismo de clase para los miembros de una clase artificial: los lápices tienen necesariamente, en todo mundo posible, la función de servir para escribir. Esto es lo que defiende, por ejemplo, Schwartz. Lo que necesita el descriptivista es que esa propiedad

⁴⁷ Véase la sección 1.5.1 para una discusión acerca de la posibilidad epistémica *versus* la posibilidad metafísica.

⁴⁸ Una cuestión diferente es que, tradicionalmente, se ha relacionado el externismo semántico con el esencialismo. El motivo, me parece, es que, en el contexto del estudio de los términos de clase natural, el externismo defiende que podemos desconocer las propiedades de las clases naturales *relevantes* para la referencia de los correspondientes términos, y podemos desconocerlas porque, típicamente, están *ocultas*, son subyacentes, y, ya desde antiguo, la naturaleza oculta, subyacente, supuesta en los objetos naturales se ha asociado o identificado con su esencia (con aquello que hace que la cosa sea lo que es). Pero creo que la asociación entre externismo y esencialismo no tiene que darse necesariamente.

de servir para escribir que tienen los lápices sea siempre cognitivamente accesible (mediante una descripción). O un cierto tipo de descriptivismo, un descriptivismo de fijación de la referencia⁴⁹, también podría optar por una postura antiesencialista: los lápices no tienen necesariamente la propiedad de servir para escribir; en otro mundo posible podrían tener otra función (y seguir siendo lo que son). Así, en cada mundo posible, la propiedad *relevante* de nuestros términos de clase artificial sería una propiedad contingente diferente pero a la que, en cada situación, tendríamos acceso epistémico.

En resumen, creo que, contra lo que se afirma algunas veces, por lo menos buena parte, si no todas, de las consideraciones metafísicas sobre los artefactos no determinan una semántica para los términos de clase artificial, a menos que, al hacer metafísica, hagamos también semántica y afirmemos que los términos de clase artificial son términos de clase nominal, como ‘soltero’, porque las clases artificiales son clases con una esencia nominal. Pero creo que, si se defiende cualquier otra postura metafísica, tanto el externismo como el internismo serán compatibles con ella, pues las teorías meramente semánticas son independientes, en principio, de otras teorías sin componentes semánticos.

Sin embargo, tanto el externismo como el internismo no son del todo independientes de la teoría epistemológica acerca de las clases artificiales, pues la motivación del externismo depende de la posibilidad de la ignorancia y el error sobre la naturaleza de los artefactos, y la plausibilidad del internismo depende de la imposibilidad de esa ignorancia o error. Así que, se defienda la postura metafísica que se defienda sobre los artefactos, es importante no dar un paso de tipo epistemológico injustificado. Y esto precisamente es lo que me parece que ocurre cuando se afirma que el hecho que nosotros produzcamos los artefactos (cosa que podría tener su incidencia metafísica) hace que tengamos inmunidad epistémica respecto a lo que son los artefactos. Esta última afirmación, por razonable que parezca, está injustificada partiendo sólo de la idea que nosotros creamos los artefactos. Precisamente, lo que está en cuestión en la

⁴⁹ Esto es, un descriptivismo que defiende que los términos de clase artificial no son sinónimos de las descripciones asociadas a ellos por los hablantes.

discusión entre el externismo y el internismo semánticos respecto a los términos de clase artificial es precisamente si existe necesariamente esa inmunidad epistémica.

2.3.2. El papel de la metafísica en la determinación de la referencia de los términos de clase natural

Para remarcar la independencia entre los aspectos metafísicos y los aspectos semánticos, creo que también puede ser útil considerar el caso de las clases naturales y sus correspondientes términos. Me parece que, en el caso de los términos de clase natural, tampoco hay un compromiso metafísico derivado sólo de la teoría semántica; la metafísica y la semántica se presentan también como ámbitos separados. Es importante notar que, por ejemplo, el esencialismo que defiende Kripke respecto a las clases naturales es independiente del externismo como teoría semántica sobre la determinación de la referencia⁵⁰. Kripke acepta el esencialismo (la tesis de que al menos algunas propiedades de algunas entidades son esenciales, son propiedades tales que esas entidades no podrían no tenerlas sin dejar de ser las entidades que son) por cuestiones que no tienen nada que ver con el externismo, ni éste implica el esencialismo. Un motivo, independiente de toda teoría semántica, para aceptar el esencialismo es que se corresponde con la intuición ampliamente compartida que algunas propiedades son esenciales a sus portadores. La intuición nos dice que algo que no es un animal no podría ser un gato o algo que no está constituido por moléculas de hidrógeno y oxígeno no podría ser agua.

Pero hay que notar que el esencialismo, como tesis metafísica, es independiente del externismo como teoría semántica sobre la determinación de la referencia de los términos. Así, por ejemplo, nada impide que un externista respecto a la determinación de la referencia de los términos de clase natural defienda un antiesencialismo (la no existencia de propiedades esenciales), o que un internista pueda defender el esencialismo. Así, un antiesencialista respecto a las clases naturales puede aceptar muy bien la teoría semántica kripkeana respecto a los términos de clase natural y admitir que son posibles los argumentos de ignorancia y error respecto a las clases naturales. Simplemente, respecto a estos argumentos, no distinguiría entre posibilidad metafísica y posibilidad epistémica; siempre se trataría, para él, de la mera posibilidad metafísica. O

⁵⁰ Salmon (1979) deja claro que el esencialismo no puede ser derivado de la nueva teoría de la referencia.

dicho de otro modo, un antiesencialista sí podría aceptar la posibilidad de descubrir que, por ejemplo, los gatos son robots y que hemos estado equivocados sobre su naturaleza. Lo que sucedería es que tomaría esta posibilidad por una posibilidad no meramente epistémica, sino también metafísica (dado que, como antiesencialista, asumiría que ser un animal no es una propiedad esencial a los gatos). Por otra parte, el esencialismo también puede ser defendido por un simpatizante del internismo semántico para los términos de clase natural, pues son teorías independientes de ámbitos distintos. Para apreciar esta compatibilidad sólo hay que notar que defender que algunas propiedades son esenciales a los miembros de una clase natural no implica defender que lo que determina la referencia de los correspondientes términos es alguna propiedad esencial. Así, por ejemplo, un esencialista podría defender que *estar constituido por moléculas de H₂O* es una propiedad esencial a las muestras de agua, pero aceptando a la vez que lo que determina la referencia de ‘agua’ son propiedades cognitivamente accesibles, como el color, el sabor, el olor del agua, etc., que puede considerar propiedades contingentes de las muestras de agua.

Otra cuestión metafísica importante acerca de las clases naturales y que, a menudo, se ha ligado al externismo semántico, es la cuestión de la naturaleza subyacente de los miembros de las clases naturales. En general, los filósofos externistas, como Kripke y Putnam, admiten que suponemos que los miembros de las clases naturales tienen una naturaleza común subyacente, que es a lo que queremos referir con nuestros respectivos términos de clase natural. Y que exista tal naturaleza subyacente, sin duda, hace más plausible el externismo semántico para los términos de clase natural pues si aquello a lo que pretendemos referir está oculto, es claro que podemos equivocarnos o ser ignorantes sobre ello⁵¹. Pero es importante destacar que tal suposición podría estar equivocada: las muestras de agua podrían no tener en común más que una serie de características observables. Así que el externismo semántico respecto a los términos de género natural no depende de que la naturaleza sea subyacente o superficial, sino de que podamos ignorar o equivocarnos sobre esa naturaleza (aunque sea superficial). Me parece que también la cuestión metafísica de la naturaleza superficial *versus* la naturaleza

⁵¹ Mostrando un caso, posible o actual, de naturaleza subyacente de una clase natural (con las posibilidades inherentes de ignorancia y error respecto a esa naturaleza), el externismo se vuelve plausible.

subyacente es independiente de la teoría semántica respecto a cómo se determina la referencia de los términos. Lo que me interesa destacar es que, al menos buena parte de las principales cuestiones metafísicas, son independientes de la semántica (en el caso de las clases naturales). La semántica de los términos de clase natural no depende de que la propiedad sea esencial o no, o de que sea subyacente o no. Entonces, si esto es así, me parece que pierde plausibilidad la idea que la semántica de los artefactos depende de su metafísica (y de sus diferencias con la metafísica de las clases naturales).

A veces, autores como Schwartz, que aceptan que la nueva teoría de la referencia explica bien cómo funcionan los términos de clase natural, han intentado justificar una semántica diferente para los términos de clase artificial apelando de alguna manera a que las clases naturales y las artificiales responden a metafísicas diferentes. Pero, tras lo expuesto, contra este razonamiento cabe objetar lo siguiente:

- (i) La concepción metafísica que uno tenga de las clases naturales no determina la semántica de los términos de clase natural (diferentes concepciones metafísicas son compatibles a la vez con diferentes tesis semánticas).
- (ii) Ciertamente, se puede asumir que las clases naturales y las clases artificiales responden a explicaciones metafísicas diferentes, pero esto, por sí mismo, no justifica que respondan a semánticas diferentes, al menos si no se da una ulterior explicación.
- (iii) La metafísica de los artefactos (que, por otro lado, no está clara) tampoco parece determinar su semántica (diferentes concepciones metafísicas son compatibles a la vez con diferentes tesis semánticas).
- (iv) Por último, intentar fundamentar la semántica, tanto de los términos de clase natural como de clase artificial, en sus respectivas metafísicas tiene el problema que, ni en el caso de las clases naturales ni en el caso de las clases artificiales, está clara cuál es la explicación metafísica apropiada.

2.4. Una primera propuesta semántica para los términos de clase artificial

2.4.1. El argumento putnamiano de los lápices-organismos

Putnam (1975) es uno de los primeros autores, si no el primero, que hace explícitamente una propuesta semántica para los términos de clase artificial. A partir de sus

consideraciones sobre los términos de clase natural, que le llevan a una semántica de tipo externista, propone también el externismo semántico para los términos para artefactos. Esta conclusión putnamiana parece guiada en alguna medida por la intuición, y los argumentos a favor de la misma no están tan desarrollados ni son tan concluyentes, para muchos, como en el caso de su análisis de los términos de género natural. En la última sección del presente capítulo, intentaré comprobar la adecuación del modelo externista semántico de Putnam aplicado a los términos para artefactos, y en qué medida la intuición putnamiana de que tales términos responden a una semántica externista puede ser correcta.

Las consideraciones de Kripke y Putnam acerca de los términos de clase natural convencieron a muchos de que la semántica correcta para términos como ‘oro’ o ‘gato’ tenía que ser una semántica de tipo externista puesto que el uso de estos términos es corregible -también dentro de una comunidad lingüística- debido a la posible ignorancia o equivocación de los usuarios de los términos respecto a la naturaleza de las clases naturales. Putnam propone extender su propuesta semántica externista a muchos otros términos del lenguaje aparte de los nombres propios y los términos de clase natural. Su intuición es que muchos términos generales responden también a una semántica externista; entre ellos, los términos de clase artificial. A primera vista, la propuesta putnamiana no parece ajustarse tan bien a los términos para artefactos porque parece razonable pensar que a la hora de usar o aplicar los términos para artefactos sí disponemos (al menos, como comunidad lingüística) de algún tipo de acceso cognitivo a las diferentes clases artificiales. Además, este acceso cognitivo que nos permite discriminar correctamente entre los tipos de artefactos parece que tiene, en particular, la forma de una descripción o una definición que, al menos algunos hablantes de la comunidad, asocian al término. Parecería entonces que tal descripción que posee, si no cada hablante particular, al menos la comunidad lingüística en su conjunto, es la que determinaría la referencia del término, haciendo de mediadora entre el término para artefacto y el artefacto. Así, por ejemplo, parece que, cuando usamos ‘lápiz’ y lo aplicamos correctamente a un objeto, estamos usando como criterio una descripción o definición, algo así como *artefacto con una forma estilizada fácilmente asible que contiene grafito y sirve para escribir*, una descripción que podríamos dar fácilmente pues apela a características observables y a características conocidas. En el caso de términos como ‘bujía’, por ejemplo, la definición correspondiente sería conocida por

ciertos expertos de la comunidad lingüística a los que el resto de hablantes podría deferir al usar el término.

Pero Putnam cree que 'lápiz' funciona semánticamente de manera análoga a como funciona 'agua'. Es decir, la extensión de los términos para artefactos, como 'lápiz', 'silla' o 'reloj', está determinada, según Putnam, por alguna propiedad que comparten los objetos que caen en la extensión del término y no por una descripción que asocien individual o colectivamente los hablantes. Ahora bien, para descartar el descriptivismo para los términos artificiales, esa naturaleza que comparten los artefactos de una misma clase tendría que poder ser inaccesible cognitivamente a los hablantes, tendría que ser susceptible de descubrirse. Esto es, para defender el externismo aplicado a los términos para artefactos, hay que argumentar que podríamos desconocer cuáles son las propiedades importantes que determinan la referencia de estos términos o que podríamos hacer descubrimientos sobre ellas. Igual que en el caso de los términos de clase natural, es crucial a un modelo externista como el modelo putnamiano que sea posible construir argumentos basados en los problemas de la ignorancia o el error que afecten a la posición internista.

Putnam es consciente de que habitualmente, aunque no siempre, las características diferenciadoras de cada clase artificial, típicamente la función y, a veces, la forma, son características conocidas por los hablantes competentes por el mero hecho de serlo. Y, por tanto, son a menudo características especificables en términos de una descripción al alcance de cualquiera. Generalmente, parece que tenemos acceso a la naturaleza de algunos de los artefactos más sencillos o comunes. Así, por ejemplo, Putnam dice que la naturaleza de los lápices es ser artefactos hechos para escribir. Parece que al término 'lápiz' va asociada una descripción como 'artefacto hecho para escribir'. Y el enunciado 'Los lápices son artefactos hechos para escribir' es, según Putnam, una afirmación metafísicamente necesaria (si es verdadera), pero epistémicamente corregible, exactamente igual que 'El agua es H₂O'. Ahora bien, que exista una descripción disponible asociada al término no significa que la referencia del término se determine mediante lo que tenemos en la mente. Ahora disponemos de una descripción asociada al término 'agua': *sustancia constituida por moléculas de H₂O*, pero tal especificación funciona sólo como un criterio para discriminar entre lo que es agua y lo que no es agua. No funciona como una definición que proporciona un significado sinónimo al

significado de ‘agua’ porque es, en principio, corregible. Además, el experimento mental de la Tierra Gemela muestra que el significado de ‘agua’ no puede ser algo equivalente al significado de alguna descripción que asociemos al término.

Putnam presenta un argumento, un experimento mental, que intenta mostrar que los términos para artefactos se usan de manera externista. Esto es, que los términos para artefactos no se usan como dice el descriptivista, sino que los hablantes intentan referir a lo que tiene la misma naturaleza que los artefactos locales en el mundo actual. Antes de exponer el experimento, conviene aclarar su motivación. Putnam propone el experimento mental para determinar cómo los hablantes usan los términos para artefactos porque, guiándonos por el comportamiento actual de los hablantes, es imposible determinar si usan el término para artefacto de manera descriptivista o de manera no descriptivista. El término ‘lápiz’, por ejemplo, se aplica a objetos con una cierta forma y función, tanto si lo que determina la referencia del término es la descripción asociada o la similitud con los lápices locales paradigmáticos. Entonces, la única manera de distinguir los usos es determinando cómo se comportarían en ciertas situaciones contrafácticas dos hablantes que usan el término para artefacto de diferentes maneras, uno de manera descriptivista y el otro de manera externista.

El experimento mental que propone consiste en imaginar que descubrimos que los lápices son organismos. Un día los abrimos, los miramos bajo el microscopio electrónico y vemos que tienen órganos, tejidos, nervios, etc. Esto es epistémicamente posible, dice Putnam; podría resultar que los lápices son organismos⁵². Para los que son contrarios a afirmar esta posibilidad epistémica, Putnam dice que los artefactos tienen, o creemos que tienen, una naturaleza artificial, no biológica o química. Y que tengan una naturaleza artificial es una cuestión de hecho, no de lenguaje. Y, así, contrariamente a lo que sugeriría la teoría descriptivista, sí podemos imaginar que descubrimos que los lápices son organismos. En la situación imaginada, el usuario descriptivista de ‘lápiz’ no aplicará el término a estos organismos, ya que ahora sabe que no satisfacen su especificación asociada (*artefacto hecho para escribir*) al término. Tal usuario no podría descubrir que los *lápices* eran organismos, ya que simplemente no serían lápices. En cambio, el usuario externista de ‘lápiz’ sí aplicará el término a tales organismos

⁵² De nuevo, se trata de una posibilidad epistémica, no metafísica (puesto que, para Putnam, si los lápices son artefactos, entonces son necesariamente artefactos).

como hacía cuando creía que eran artefactos. Lo que dirá es que ahora sabe cuál es la naturaleza de los lápices. De ahora en adelante rehusará aplicar ‘lápiz’ a candidatos que no sean de esta clase de organismos (a menos que se haga un ulterior descubrimiento sobre la naturaleza de los lápices). Así, concluye Putnam, como *nosotros* podemos imaginar que los lápices son organismos, debemos de usar el término de manera externista o no descriptivista. De manera análoga, en el famoso argumento de los gatos-robots también propuesto por Putnam, si descubriéramos que todos los gatos son en realidad robots con apariencia de gatos, seguiríamos aplicando el término ‘gato’ a los mismos individuos, dice Putnam. Simplemente ahora conoceríamos su verdadera naturaleza, sobre la que estábamos equivocados.

Schwartz (1978) presenta una posible réplica descriptivista al argumento de Putnam de los lápices-organismos consistente en decir que la especificación que el hablante asocia con ‘lápiz’ no es la que propone Putnam. La descripción asociada no es ‘artefacto hecho para escribir’, sino más bien una descripción como ‘*objeto* hecho para escribir’ u ‘*objeto* con la función de servir para escribir’. Es decir, el descriptivista puede apelar a que la especificación asociada a ‘lápiz’ no incluye que sea un artefacto y, por tanto, en el experimento mental de Putnam, el usuario descriptivista también podría descubrir, contra lo que afirma Putnam, que los lápices son organismos. Es decir, el descriptivista, para dar por buena la intuición que sería posible descubrir que los lápices son organismos, necesita que en la descripción asociada a ‘lápiz’ no se haga alusión al origen de los lápices, sino únicamente a su función o a su función y su forma.

Ahora bien, si Schwartz está en lo cierto, y la especificación asociada a ‘lápiz’ es la que él propone, ha mostrado, a lo sumo, que la descripción que mantiene Putnam como asociada a lápiz es la descripción equivocada y que, suprimiendo la noción de *artefacto* de dicha especificación, el experimento mental putnamiano no es concluyente. Pero la razón que ofrece Schwartz para justificar que la descripción asociada a ‘lápiz’ no incluye ninguna referencia al origen artificial de los lápices no parece tampoco concluyente. Para él, las descripciones asociadas a los términos de clase artificial son descripciones analíticas, descripciones verdaderas por definición y que expresan verdades necesarias predicadas de los artefactos en cuestión. Por tanto, la descripción asociada a ‘lápiz’ tiene que ser una descripción verdadera y necesaria sobre los lápices. Y, según Schwartz, no es necesario (metafísicamente) que los lápices sean artefactos;

podrían crecer en los árboles, por ejemplo, dice Schwartz y, así, no serían ni artefactos (pues tendrían un origen natural) ni organismos (pues no estarían vivos). Esto es extremadamente improbable, pero no metafísicamente imposible, concluye Schwartz.

El problema de esta argumentación es la idea de que no es necesario metafísicamente que los lápices sean artefactos. Esto no parece obvio. Putnam tiene la intuición contraria: que los lápices, si son artefactos, son necesariamente artefactos. Otra cosa es que podamos imaginar, como en el experimento mental propuesto, que no son artefactos, sino organismos. Pero se trata sólo de una posibilidad epistémica, no metafísica. La intuición de Putnam es que los lápices son necesariamente artefactos y ‘Los lápices son artefactos hechos para escribir’ es una afirmación necesaria (si es verdadera), pero epistémicamente corregible. Que exista la posibilidad de que la afirmación sea falsa sólo quiere decir que podríamos desconocer la verdadera naturaleza de los lápices, pero no que los lápices podrían ser de otro modo esencialmente diferente a como en realidad son. De todas maneras, si el posicionamiento sobre esta cuestión acerca de la naturaleza necesaria de los lápices es un asunto de tener o no determinadas intuiciones, entonces la réplica de Schwartz a Putnam tampoco es concluyente. Además, suponiendo que Schwartz tenga razón y la descripción que proporciona Putnam esté equivocada, esto no excluiría *per se* que la semántica de ‘lápiz’ sea una de tipo externista. La extensión de ‘lápiz’ aún podría determinarse mediante una relación de similitud con respecto a la función (no la naturaleza original) de los lápices paradigmáticos. Que haya una descripción singularizadora de una clase artificial no es suficiente para mostrar que la semántica del término de clase artificial es de tipo descriptivista. Por otro lado, la dificultad patente en consensuar una descripción asociada a un término para un artefacto tan *sencillo* y habitual como los lápices parece ir en contra de la propia idea descriptivista.

2.4.2. Adecuación del modelo putnamiano a los términos para artefactos

A continuación, expongo los principales elementos del modelo putnamiano para evaluar su adecuación aplicados a los términos de clase artificial.

Un primer elemento del modelo putnamiano es, naturalmente, su externismo: la extensión de los términos de clase artificial no está determinada por alguna entidad cognitivamente accesible, sea una descripción o no, sino por características externas de

las cosas que podrían no resultarnos cognitivamente accesibles. Es decir, podría ser que estas características tuviesen que ser descubiertas. Presumiblemente, cuando Putnam extiende su propuesta semántica a los términos para artefactos hay que entender que, también para estos términos, los significados *no están en la cabeza*.

En principio, entonces, también en el caso de la aplicación de los términos de clase artificial son posibles los problemas de la ignorancia y del error. Esto es, de acuerdo al modelo putnamiano, es posible que usemos un término para artefacto a pesar de desconocer la naturaleza de la clase artificial a la que nos referimos o que nos equivoquemos al juzgar esa naturaleza, y, no obstante, logremos referir con el término o aplicarlo con éxito. Las características de los artefactos que determinan la referencia de los términos para artefactos pueden no resultarnos cognitivamente accesibles: pueden ser desconocidas o podemos equivocarnos sobre ellas. En el **capítulo 4**, exploraré la posibilidad de construir argumentos basados en la ignorancia y el error sobre la naturaleza de las clases artificiales y el alcance que tienen contra el descriptivismo.

Siguiendo con el modelo putnamiano, ¿cómo se determinaría, entonces, la referencia de un término como ‘lápiz’? Hay que suponer que de manera análoga a como se determina la referencia de ‘agua’. La extensión de ‘lápiz’, entonces, quedaría determinada de la siguiente manera: una entidad x , en cualquier mundo posible, es un lápiz si y sólo si mantiene la relación *mismo A (mismo artefacto que)*, entendida como una relación a través de los mundos posibles, con los objetos a los que *nosotros* llamamos ‘lápiz’ en el mundo actual. La relación *mismo artefacto que* hay que entenderla como una relación a través de los mundos posibles por la que si un artefacto en el mundo W_1 tiene las mismas propiedades importantes que un artefacto en W_2 , entonces el primer artefacto mantiene la relación *mismo A* con el último artefacto.

Las *propiedades importantes* o *relevantes* compartidas por los miembros de la clase artificial a las que alude Putnam parece que hay que entenderlas, si la situación es análoga a la de los miembros de una clase natural, como propiedades esenciales a la clase artificial (esencia de clase); esto es, como propiedades que necesariamente los miembros de la clase tienen en tanto que miembros de la clase. En términos más formales, la definición de Putnam parece implicar que:

$\Box \forall x (x \text{ es un lápiz} \leftrightarrow x \text{ es el mismo } A \text{ que los lápices } T)$

(donde 'lápicesT' está por *lápices de la Tierra* o *lápices en el mundo actual*)

Esto significa que necesariamente es el caso que x es un lápiz si y sólo si es el mismo tipo de artefacto que los objetos a los que nosotros llamamos 'lápiz' en la Tierra. O sea que, en todo mundo posible, si x es un lápiz, entonces x mantiene la relación *mismo A (mismo artefacto que)* con los lápices locales, y, en todo mundo posible, si x mantiene la relación *mismo A (mismo artefacto que)* con los lápices locales, entonces x es un lápiz. Hay que notar que, si esto es lo que dice la definición de Putnam, de su definición no se sigue que los objetos que de hecho son lápices en la Tierra sean necesariamente el tipo de artefacto que son. Es decir, de

Necesariamente $\forall x (x \text{ es un lápiz} \leftrightarrow x \text{ es el mismo } A \text{ que los lápices } T)$

no se sigue

$\forall x (x \text{ es un lápiz} \leftrightarrow \text{Necesariamente } x \text{ es el mismo } A \text{ que los lápices } T)$

Por tanto, la definición de Putnam no implica nada respecto a las esencias individuales de los lápices. Es compatible con su definición que este lápiz que tengo delante existiese en otro mundo posible sin tener las *propiedades importantes* de los lápices locales; es decir, que este lápiz podría existir en otro mundo posible sin tener la función de los lápices locales. Ahora bien, que de 'Necesariamente los lápices tienen la función de escribir' no se siga 'Este lápiz tiene necesariamente la función de escribir', no quiere decir que no sea deseable sostener este último enunciado, pues parece verdadero intuitivamente (con la adecuada interpretación de *función*). Simplemente, se necesitaría otro argumento diferente para defender que cada lápiz, en tanto que el objeto concreto que es, tiene necesariamente la función de escribir.

Putnam enfatiza el carácter externista de la semántica de los términos apelando tanto al componente indéxico como al componente social presente en la determinación de la referencia. Así, respecto al componente indéxico, cuando usamos 'lápiz', por ejemplo, la idea putnamiana sería que pretendemos que el término refiera a todo lo que comparte las características importantes con los lápices paradigmáticos. El componente indéxico se manifestaría en la introducción de un nuevo término, en la que señalando una muestra paradigmática aplicaríamos el término y pretenderíamos que refiriera a todo lo

que comparte las mismas características relevantes. Sin embargo, un dato contra esta intuición putnamiana es que la introducción de un nuevo término para un tipo de artefacto típicamente no es ostensiva (aunque puede haber casos en que sí lo sea). Generalmente, se introduce el término junto con una descripción que da cuenta de algunas propiedades relevantes del artefacto (habitualmente, una descripción de su función). Otra cosa es que esa descripción introductoria proporcione las condiciones necesarias y suficientes de pertenencia a la clase artificial.

En cambio, la determinación social de la referencia de los términos de clase artificial es un fenómeno muy extendido. Hay multitud de artefactos, cuyos términos correspondientes son empleados a menudo, de los que los hablantes competentes no conocen las características distintivas. Parece plausible, entonces, suponer que deferimos a los expertos en estos casos en virtud de una división del trabajo lingüístico en nuestra comunidad que también abarcaría al menos algunos términos para artefactos. Del mismo modo, también puede darse una cadena causal histórica a través de la cual recibamos los términos de clase artificial de otras personas intentando usarlos de la misma manera, aunque no conozcamos ninguna descripción singularizadora de la clase artificial.

En resumen, el modelo de Putnam, de entrada, parece que no se adapta tan bien a los términos para artefactos como a los términos de clase natural. La principal dificultad está, desde mi punto de vista, en ofrecer un argumento análogo al de la Tierra Gemela para un término de clase artificial; el argumento de los lápices-organismos no ha resultado convincente a muchos autores por una razón aducida por Schwartz: lo definitorio de una clase artificial –y aquello a lo que nuestros términos para artefactos pretenden referir- es, en principio, su función, no su origen artificial. Por tanto, el externismo debe presentar un argumento de ignorancia y error respecto a la función de un tipo de artefactos (lo que consideramos su esencia), no respecto a su origen. Ante la falta de un argumento así, parece que la carga de la prueba está del lado del externista, pues el descriptivista parte de la plausibilidad inicial del descriptivismo para los términos de clase artificial. En el **capítulo 4**, discuto la posibilidad de presentar argumentos de este tipo.

Capítulo 3. Las réplicas descriptivistas al modelo putnamiano para los términos de clase artificial y objeciones a estas réplicas

En este capítulo, presento las críticas que, desde posturas descriptivistas, se han hecho a la propuesta putnamiana de extender la semántica externista a los términos de clase artificial. Quiero destacar que estas alternativas internistas al externismo putnamiano son teorías *descriptivistas* en algún grado; no he encontrado en la literatura ninguna alternativa semántica a la propuesta de Putnam que se apoyara en un internismo no descriptivista. El internismo semántico para los términos para artefactos supone que nuestro acceso cognitivo a los artefactos se produce mediante el acceso a una descripción que singulariza su naturaleza. Así que dejaré de lado el tratamiento de la posibilidad de un internismo no descriptivista para los términos de clase artificial.

Empiezo por considerar, en la sección **3.1**, la crítica de Schwartz a la propuesta putnamiana. Su oposición se basa, principalmente, en consideraciones metafísicas sobre la naturaleza de los artefactos y en consideraciones sobre nuestro acceso epistémico a las clases artificiales. La alternativa semántica que Schwartz presenta es lo que llamaré ‘descriptivismo individualista’, para distinguirlo de un descriptivismo *social* o *comunitario*. En la sección **3.2**, planteo las críticas que pueden hacerse a la propuesta de Schwartz, criticando tanto su metafísica como los argumentos que presenta a favor de su teoría descriptivista para los términos para artefactos.

A continuación, en la sección **3.3**, expongo otra crítica al modelo putnamiano: la crítica de Thomasson. Como Schwartz, Thomasson también parte de consideraciones metafísicas para llegar a conclusiones semánticas de tipo descriptivista en algún grado. Sin embargo, su alternativa semántica para los términos de clase artificial, aunque se opone al externismo de Putnam, no pretende ser *puramente* descriptivista, como lo es la propuesta de Schwartz. Ella defiende una *teoría híbrida* de la referencia, a la que también llamaré ‘teoría causal descriptiva’, que incorpora elementos descriptivistas y externistas. En la última sección, **3.4**, expongo diferentes críticas que pueden plantearse a una teoría causal descriptiva como la que sostiene Thomasson.

3.1. La crítica del descriptivismo *puro* a la propuesta de Putnam

3.1.1. Consideraciones metafísicas

La propuesta putnamiana de extender una semántica de tipo externista para los términos de clase artificial ha encontrado oposición también por parte de autores que, como Schwartz, aceptan el externismo semántico para los términos de género natural. Schwartz sí acepta el modelo externista putnamiano para los términos de clase natural porque, según él, los miembros de una clase natural comparten una naturaleza oculta o subyacente. Así, por ejemplo, todas las muestras de agua tienen una naturaleza subyacente común: estar constituidas por H₂O. Y esta naturaleza común subyacente hay que descubrirla. Y si hay que descubrirla, resulta que, antes de llevar a cabo tal investigación, podríamos no disponer de una especificación que singularice la clase natural. Esto es, podríamos ignorar o equivocarnos sobre la naturaleza de cierto tipo de objetos naturales y, no obstante, referir con éxito al usar el correspondiente término de clase natural. Así, en el caso del agua, antes del descubrimiento de su naturaleza subyacente, nuestra descripción asociada a ‘agua’ podría no determinar la extensión del término, tal como muestra el experimento mental de la Tierra Gemela. El externismo es, entonces, plausible, para Schwartz, aplicado a los términos de clase natural.

En cambio, los artefactos de una misma clase, dice Schwartz, no tienen en común ninguna naturaleza oculta o subyacente. La categoría de miembro de una clase artificial está determinada, según él, por características *superficiales*, tales como una cierta forma y función. Como expondré en la siguiente sección, al presentar la crítica a la propuesta de Schwartz, la función no es una característica *superficial* en el sentido de ser una característica *directamente accesible a través de los sentidos*, como lo es la forma. Aunque, ciertamente, podemos *inferir* la función de un artefacto viendo cómo es usado por otros, ante un artefacto desconocido y que nunca hemos visto utilizar, tendríamos que investigar cuál puede ser su función. Creo que el modo más plausible de interpretar a Schwartz cuando califica a la función de los artefactos de característica *superficial* es interpretar que Schwartz quiere decir que se trata de una característica *no microfísica*. Otra interpretación que creo que también concuerda con lo que defiende Schwartz es considerar la función de los artefactos como algo cognitivamente accesible –aunque no sea perceptible– para nosotros en virtud de nuestra relación epistémica con los artefactos.

Partiendo entonces de esta diferencia metafísica entre la naturaleza subyacente o microfísica de los objetos naturales y la naturaleza no microfísica y necesariamente accesible de los artefactos, Schwartz pretende llegar a conclusiones semánticas sobre los términos de clase artificial. Contrastando los términos ‘lápiz’ y ‘agua’, dice que, si ‘lápiz’ fuera como ‘agua’ y, por tanto, su significado estuviera determinado por algo externo no necesariamente accesible cognitivamente, los lápices tendrían una naturaleza oculta común, y nosotros tendríamos la intención de que ‘lápiz’ refiriera a lo que sea que comparte esta naturaleza. Pero no es el caso que los lápices tengan una naturaleza oculta o subyacente por descubrir ni existe tampoco la presunción de tal naturaleza. Lo que hace a algo un lápiz son características *superficiales* como una cierta forma y función.

The big difference between artifact kinds and water is that we do not presuppose that there is any underlying nature that makes something to be the kind of artifact that it is. [...] When I say that there is no common nature that all pencils have I mean that there is nothing apart from some set of superficial characteristics that all pencils have in common. (Schwartz 1978, 572)

Y estas características *superficiales*, cognitivamente accesibles, de los lápices ya proporcionan, según Schwartz, condiciones necesarias y suficientes para ser un miembro de la clase artificial de los lápices. A partir de aquí, Schwartz infiere que ‘lápiz’ no puede tener una semántica externista tal como propone Putnam pues no hay ninguna naturaleza oculta por descubrir. Como expuse en la sección 2.2.2 del capítulo anterior, la esencia o naturaleza de los artefactos es, para Schwartz, una esencia nominal o lingüística. De acuerdo con Schwartz, podemos dar una definición o especificación nominal para los términos de clase artificial igual que para los términos ‘soltero’ o ‘cazador’. Según Schwartz, además, esta especificación o definición asociada a cada término de clase artificial predicada de los miembros de la clase expresa una verdad analítica: un enunciado verdadero sólo en virtud del significado del término para artefacto. Es una verdad necesaria y conocida *a priori*. Y es incorregible: no está sometida a refutación por ulteriores descubrimientos. Schwartz ilustra este punto proporcionando algún ejemplo de definición para un término de clase artificial. Así, la especificación analítica del término ‘sloop’ es ‘A boat having a single mast with a mainsail and jib’ (Schwartz 1980, 183).

En este caso, la definición del término para artefacto, ‘sloop’, incluye otro término de clase artificial, ‘boat’. En otros casos, la definición de un término de clase artificial no incluye ningún otro término para artefacto que lo englobe, y la especificación asociada solamente hace referencia a la forma y la función de los objetos que caen en la extensión del término, como en el caso de ‘lápiz’ (ver sección 2.4.1). Pero hay que notar que, en ambos casos, ninguna definición asociada a un término para artefacto incluye ninguna referencia al origen artificial de los artefactos en cuestión. La especificación asociada a ‘lápiz’ no es ‘artefacto que sirve para escribir’, sino, como vimos, ‘objeto *con tal y tal forma* que sirve para escribir’. Y ‘boat’, igualmente, se define sólo haciendo referencia a su forma y función, y, por tanto, en la definición de ‘sloop’ tampoco hay ninguna referencia explícita ni implícita al hecho que sea un artefacto. En la sección 2.4.1 del **capítulo 2**, expliqué que Schwartz se ve obligado a prescindir de cualquier alusión al origen de los lápices en la definición de ‘lápiz’ porque admite como posible el experimento mental de Putnam consistente en imaginar que descubrimos que los objetos a los que llamábamos ‘lápices’ son, en realidad, organismos. Lo que Schwartz critica del experimento mental de Putnam es que éste considera que la descripción asociada a ‘lápiz’ hace referencia a que es un artefacto. Así, según una variante del mismo experimento mental, si, en una Tierra Gemela, los objetos que pasan por ser lápices fueran en realidad organismos (mientras aquí, en la Tierra, son artefactos), estos *lápices* miméticos, al ser indistinguibles, en términos de la forma y la función, de los lápices de la Tierra, serían también lápices para Schwartz. La extensión de ‘lápiz’ incluiría tanto a los lápices de la Tierra como a los de la Tierra Gemela. En cambio, para Putnam, no serían realmente lápices porque su naturaleza sería diferente a la de los lápices terrestres. La crítica de Schwartz a Putnam es que se equivoca al pensar que cualquier cosa que no fuera un artefacto no podría ser un lápiz. *Ser un artefacto* no forma parte, pues, según Schwartz, de la especificación analítica asociada a ‘lápiz’; no es necesario que los lápices sean artefactos: podrían crecer en los árboles, por ejemplo, dice Schwartz.

En resumen, la crítica de Schwartz a la propuesta semántica putnamiana para los términos de clase artificial consiste en decir que la definición analítica asociada a ‘lápiz’ es ‘objeto (*con cierta forma*) que sirve para escribir’. Esta definición, en tanto que es analítica, expresa una verdad necesaria de los lápices y es una definición conocida *a priori* por los hablantes competentes del lenguaje.

3.1.2. Un criterio para distinguir entre términos de clase natural y términos de clase artificial

Schwartz aún presenta otro argumento contra la teoría de Putnam: proporciona un criterio, el criterio de las generalizaciones estables, para distinguir entre los términos de clase natural y los términos de clase artificial. Resaltando, de este modo, las diferencias entre las clases naturales y las clases artificiales, Schwartz pretende concluir que los términos de clase natural y los de clase artificial deben de responder a teorías semánticas diferentes. Antes de exponer el mencionado criterio, Schwartz parte de algunas observaciones. En primer lugar, Schwartz observa que las oraciones que expresan la naturaleza o esencia de una clase -artificial o natural- son metafísicamente necesarias. Así, por ejemplo, tanto ‘los lápices son objetos (*con tal y tal forma*) que sirven para escribir’ como ‘el agua es H₂O’ son oraciones metafísicamente necesarias; esto es, no están sujetas a ser refutadas por un contraejemplo. Si descubriéramos que una aparente muestra de agua no está constituida por H₂O, no diríamos que tenemos un contraejemplo a la afirmación de que el agua es H₂O, sino que la muestra en cuestión no es agua. La afirmación, entonces, supera lo que Schwartz llama ‘el test del contraejemplo’. Y ‘los lápices son objetos que sirven para escribir’ también pasa el test del contraejemplo: algo que no tuviera la función y forma de los lápices no podría ser un lápiz, dice. La afirmación pasa el test del contraejemplo porque, para Schwartz, se trata de una verdad analítica y, entonces, obviamente, no puede ser refutada por un contraejemplo.

La diferencia entre las oraciones que expresan la naturaleza de una clase artificial, como ‘los lápices son objetos (con una cierta forma) que sirven para escribir’, y las oraciones que expresan la naturaleza de una clase natural, como ‘el agua es H₂O’, es que las primeras son incorregibles, mientras que las segundas son corregibles. Así, aunque las oraciones acerca de la esencia de las clases naturales son metafísicamente necesarias, son corregibles. Podríamos descubrir que, en realidad, la naturaleza del agua no es estar constituida por H₂O, sino por otra propiedad. En cambio, ‘los lápices son objetos que sirven para escribir’ es una verdad incorregible de acuerdo con Schwartz.

Pues bien, a partir de estas consideraciones, Schwartz llama ‘generalizaciones estables’ a las oraciones que tratan sobre la naturaleza de una clase y cumplen dos condiciones: (i) son metafísicamente necesarias (pasan el test del contraejemplo) y (ii) son corregibles o epistémicamente contingentes. A partir de aquí, el criterio que proporciona

para distinguir entre los términos de clase natural y los términos de clase artificial es el siguiente: los términos de género natural pueden ser los sujetos de generalizaciones estables, mientras que los términos de género artificial no. Así, ‘El oro es el elemento con número atómico 79’ es una generalización estable, pues cumple (i) y (ii), pero ‘los lápices son objetos (con una cierta forma) que tienen la función de servir para escribir’ no lo es porque aunque cumple (i), dado que es una verdad analítica, no cumple (ii) por el mismo motivo. Y ‘los lápices son *artefactos* que sirven para escribir’ (la especificación propuesta por Putnam) tampoco es una generalización estable, ya que, aunque cumple (ii), no pasa el test del contraejemplo (un único lápiz que no fuera un artefacto sí falsearía la generalización, dice Schwartz).

Para Schwartz, esta diferencia entre los términos de clase natural y los términos de clase artificial respecto a las generalizaciones estables es un indicio de que la semántica de los términos para clases naturales y para clases artificiales no puede ser la misma pues estamos en situaciones epistémicas diferentes en cada caso respecto a la naturaleza de las clases. En definitiva, para Schwartz, la descripción asociada a cada término de clase artificial está asociada analíticamente y expresa las condiciones necesarias y suficientes para la pertenencia a la clase artificial. Y la extensión de los términos de clase artificial quedaría determinada por la satisfacción, por parte de los artefactos, de las condiciones expresadas por esa descripción.

3.2. Objeciones al descriptivismo *individualista* de Schwartz

3.2.1. Crítica a la concepción metafísica de Schwartz

La primera objeción que se le puede hacer a Schwartz es sobre su noción de *características superficiales*, cuando habla de las características definatorias de las clases artificiales. Como remarqué en el **capítulo 2**, con ‘*características superficiales*’ no puede hacer referencia a características *directamente observables o accesibles a través de los sentidos*, pues es obvio que la función de un artefacto no es una propiedad perceptible. Ante un artefacto desconocido para nosotros, tendríamos que investigar cuál es su función si nadie nos informa al respecto. Como dije antes, una interpretación plausible de lo que dice Schwartz es considerar que, con ‘superficial’, quiere decir *no subyacente*, pues opone las características *superficiales* de los artefactos a las

características *subyacentes* u *ocultas* de los objetos naturales. Y por característica *subyacente* parece que hay que entender característica *microfísica*. Ahora bien, si por característica *superficial* hay que entender característica *no microfísica* o *no subyacente*, hay que notar dos cosas. La primera es que tal diferencia entre las condiciones de pertenencia a una clase natural (compartir cierta naturaleza microfísica) y de pertenencia a una clase artificial (compartir ciertas propiedades no microfísicas como la forma y la función) no justifica, en principio, una explicación semántica diferente del funcionamiento de los términos de clase natural y los de clase artificial. Defender que diferentes análisis metafísicos conducen a diferentes semánticas exige una justificación que Schwartz no proporciona. En segundo lugar, hay que notar que no es tan obvio que todos los miembros de una clase natural compartan una misma naturaleza *microfísica* (si es que hay que entender así la naturaleza *subyacente* a la que apela Schwartz). De hecho, es dudoso que haya una tal naturaleza subyacente al nivel microfísico compartida por los miembros de, por ejemplo, una misma especie biológica⁵³. Y hay otros casos, como el caso de la artritis. Dejando a un margen si ‘artritis’ es un término de clase natural o no, hemos descubierto que, contrariamente a lo que se podía suponer, la artritis sólo consiste en una serie de síntomas superficiales, principalmente la inflamación de las articulaciones, sin que haya nada más en común (ningún proceso o ninguna causa común oculta) entre las diferentes instancias de la enfermedad. Pero este hecho no apoya por sí mismo que la extensión de ‘artritis’ esté determinada por una descripción que exprese ese conjunto de características superficiales, pues a tal descripción hemos llegado por investigación empírica, no por introspección o competencia lingüística.

Incluso respecto a los términos de sustancias químicas, como ‘agua’, Putnam dice que con el uso de tal término va implícita la presuposición de una naturaleza oculta microfísica compartida por las muestras de agua, pero tal presuposición no tiene por qué ser verdadera. Podríamos descubrir que, en realidad, todas las muestras de agua no tienen en común más que algunas propiedades observables y nada más. Pero esto sería un descubrimiento que mostraría que estábamos equivocados. Así que el hecho de poder aplicar el externismo semántico a un tipo de términos generales, como los términos de clase natural, no depende de que los miembros de la clase natural en cuestión compartan

⁵³ Véase la nota 1 del **capítulo 2**.

una misma naturaleza microfísica (pues, quizá, no la compartan) sino que, desde mi punto de vista, la plausibilidad del externismo depende de que podamos hacer descubrimientos acerca de la naturaleza *relevante* de los miembros de la clase. Poniendo un ejemplo que da el propio Schwartz, cuando descubrimos que las ballenas eran en realidad mamíferos y no peces -como pensábamos-, se mostró, desde mi punto de vista, que la descripción ‘pez de grandes dimensiones con *tales y tales características*’ no podía proporcionar el significado de ‘ballena’. Creo que la viabilidad del externismo no depende tanto de que la extensión esté determinada por una naturaleza subyacente o no, como de que esté determinada por alguna propiedad que podríamos desconocer.

En definitiva, creo que no se puede concluir, como hace Schwartz, que los términos de clase artificial no responden a una semántica externista como la propuesta por Putnam porque los artefactos no tienen una naturaleza subyacente entendida en un sentido microfísico. Es cierto que la composición molecular compartida por las sustancias de una misma clase natural parece dar fuerza al externismo para los términos de clase natural, pero tal composición molecular común podría no darse y esto no excluiría necesariamente el externismo semántico para el correspondiente término de clase natural. Igualmente, entonces, que no exista una tal naturaleza química compartida por los artefactos de una clase artificial no excluye *per se* el externismo para los términos para artefactos. Lo decisivo a favor de una postura externista, y que socavaría el internismo semántico, sería exponer argumentos de ignorancia y error que mostraran que podríamos desconocer en qué consiste la naturaleza *relevante* de los artefactos, sea la que sea. En el siguiente capítulo, el **capítulo 4**, consideraré la posibilidad de tales argumentos, intentando dilucidar su alcance en la disputa entre el externismo y el descriptivismo semánticos para los términos de clase artificial.

Por otro lado, se podría considerar que Schwartz, al hablar de la forma y la función de los artefactos como características *superficiales* no está simplemente diciendo que se trata de características no microfísicas, puesto que es obvio que esas propiedades no son microfísicas. Sino que quiere decir que son *superficiales* en el sentido que son accesibles cognitivamente solamente en virtud de nuestra competencia lingüística, independientemente de si son inmediatamente accesibles a los sentidos o no. Pero entonces es claro que esto no se puede aplicar a todos los hablantes competentes de una comunidad lingüística. Muchos hablantes ignoran, por ejemplo, las propiedades

superficiales definitorias (la forma y la función) de las bujías y, sin embargo, usan el correspondiente término de manera exitosa. Por tanto, si Schwartz quiere defender que las propiedades relevantes de los artefactos son cognitivamente accesibles a los hablantes en virtud de ser hablantes competentes, un descriptivismo *social* o *comunitario* en el que sólo algunos hablantes tienen acceso cognitivo a la forma y la función de un cierto tipo de artefacto es más plausible que el descriptivismo *individualista* que defiende. En una propuesta de descriptivismo *social*, los hablantes que desconocen la forma y la función de un tipo de artefacto consiguen referir al usar el correspondiente término de clase artificial a través del mecanismo de la deferencia a los expertos (en ese tipo de artefactos) de la comunidad. Creo que Schwartz tendría que distinguir entre hablantes expertos en un tipo de artefactos, o hablantes que conocen sus características distintivas, y hablantes que no las conocen, pues parece claro que no todos los hablantes competentes de una comunidad lingüística poseen una descripción singularizadora asociada con cada término para artefacto que usan. Desde un descriptivismo de este tipo⁵⁴, las características relevantes de cada clase artificial se considerarían accesibles cognitivamente sin ulterior investigación empírica *para la comunidad lingüística en su conjunto*. Sería la descripción que posee la comunidad lingüística, en virtud de que la poseen algunos de sus miembros, la que determinaría la referencia de los términos de clase artificial⁵⁵.

Sin duda, este descriptivismo comunitario es preferible al de Schwartz y es de suponer que, en alguna medida, sería suscrito por el propio Schwartz si quiere admitir que los hablantes que desconocen las características importantes de un tipo de artefactos pueden

⁵⁴ Considero que esta explicación es descriptivista ya que, aunque la deferencia a los expertos introduce un elemento *externo* en la determinación de la referencia de los hablantes no expertos, apela, en última instancia, a una descripción conocida por alguien como aquello que selecciona el referente de los términos para artefactos.

⁵⁵ Este descriptivismo *social* o *comunitario* es muy diferente del descriptivismo también considerado *social* de la teoría *cúmulo* (*cluster*) de Searle. En el primero, la referencia de un término está determinada por una descripción conocida por algún o algunos expertos de la comunidad a la que defieren los usos no expertos. En cambio, en la teoría *cúmulo* de Searle, la referencia de un término está determinada por un número *suficiente* de descripciones del cúmulo social que el conjunto de los hablantes asocia con el término en cuestión.

usar con éxito el término correspondiente⁵⁶. Para intentar rebatir este descriptivismo *social*, más plausible de entrada, creo que hay que mostrar que puede estar equivocado en su idea de fondo; esto es, habría que mostrar que hay argumentos de ignorancia y error respecto a las propiedades relevantes de los artefactos por parte de toda una comunidad lingüística, incluyendo a los expertos. Lo que está en cuestión, en la discusión entre el externismo y el internismo semánticos respecto a los términos para artefactos, es si es posible que se determine la referencia de uno de estos términos sin apelar al conocimiento de una descripción por parte de alguien (sea cada hablante particular sea un experto dentro de comunidad lingüística). En el capítulo siguiente, el **capítulo 4**, exploro la posibilidad y el alcance de tales argumentos externistas de ignorancia y error.

3.2.2. Crítica a la analiticidad de los términos de clase artificial

Otro punto que enfatiza Schwartz para descartar el externismo para los términos para artefactos es que, según él, la teoría descriptivista puede dar cuenta de su significado. Según Schwartz, los hablantes asociamos descripciones con los términos de clase artificial que ya proporcionan las condiciones necesarias y suficientes para la aplicación de los términos. Y estas descripciones suelen especificar la forma y la función de cada tipo de artefacto. Pero, como apunté en el **capítulo 2**, la especificación de la forma y la función de un tipo de artefactos no siempre proporciona las condiciones necesarias y suficientes de pertenencia a una clase artificial y de aplicación del término, pues los miembros rotos de una clase artificial y los miembros funcionales con formas *originales* también parecen pertenecer a la clase correspondiente. En este sentido, creo que es claro que los términos de clase artificial, como ‘lápiz’, no son como los términos de clase nominal, términos como ‘soltero’ o ‘pescador’.

Por otro lado, incluso si entre los miembros de una clase artificial sólo admitimos a los miembros funcionales y con formas más *convencionales*, no parece tan *fácil* o inmediatamente accesible proporcionar una definición singularizadora del tipo de artefactos en cuestión, quizá ni siquiera para los expertos. Para empezar, como vimos,

⁵⁶ A menos que Schwartz quiera defender que los hablantes que usan, por ejemplo, ‘bujía’ sin conocer las características distintivas de las bujías no son hablantes competentes, y, al usar el término, no consiguen referir.

no hay acuerdo entre los filósofos sobre si la definición asociada a un término de clase artificial debe especificar si se trata de un artefacto o no. Además, la especificación que da el propio Schwartz asociada a 'lápiz', *objeto con cierta forma que sirve para escribir*, no distingue entre lápices y bolígrafos, por ejemplo. Esto no quiere decir que no se pueda dar una especificación mejor asociada a 'lápiz', pero quizá no está claro que sólo haya que apelar a la forma y la función de los lápices; quizá podría ser necesario apelar a otro tipo de característica como el material. Y también podría muy bien suceder que cada clase artificial requiera apelar a un tipo diferente de características, de modo que no sea correcto decir que las características definitorias de las clases artificiales son, en todos los casos, solamente la forma y la función.

Hay que notar, incidentalmente, que el hecho que los hablantes poseyeran una descripción singularizadora de las diferentes clases artificiales, no descartaría automáticamente el externismo semántico para los términos de artefactos. En el caso de los nombres propios, muchos hablantes sí asocian una descripción singularizadora -o serían capaces de proporcionar una- con muchos de los nombres propios que usan. Pero basta que sean posibles casos como los proporcionados por Kripke con 'Gödel' o 'Feynman', para que el descriptivismo aplicado a los nombres propios pierda crédito. De nuevo, lo determinante, desde mi punto de vista, para la validez del descriptivismo no es que poseamos descripciones singularizadoras asociadas con los términos que usamos, sino que no sea posible⁵⁷ construir argumentos de ignorancia y error respecto a esas descripciones. Da la impresión que Schwartz quiere excluir el externismo para los términos de clase artificial basándose en que poseemos descripciones singularizadoras de los artefactos. Aun asumiendo que esto es así, hay que notar que el hecho de conocer una descripción singularizadora de las clases artificiales no excluye *per se* el externismo para los correspondientes términos. El razonamiento de Schwartz parece ser el siguiente: si (en algún momento) no disponemos de una descripción que seleccione la extensión de un término, entonces la semántica internista para el término queda descartada. Pero de esto pasa a concluir que si disponemos de tal descripción sobre una clase, entonces la semántica del correspondiente término es de tipo internista y el externismo queda descartado. Y este último razonamiento no se sigue del primero. Podemos, de hecho, asociar a un término una descripción que selecciona su extensión,

⁵⁷ La posibilidad implicada es de tipo epistémico, como explico en 1.5.1.

pero esto no excluye que esa descripción *podría* haber estado equivocada o haber sido insuficiente para seleccionar esa extensión. Y es esta posibilidad, creo, la determinante para argumentar a favor del externismo semántico, o en contra cuando tal posibilidad no exista. Y esto también se aplica a los términos para artefactos. En el capítulo siguiente retomaré esta cuestión.

Otro problema de la teoría de Schwartz, derivado de la analiticidad que presentan, según él, las definiciones asociadas a los términos de clase artificial, es que esas descripciones o definiciones asociadas a los términos para artefactos tienen que ser conocidas *a priori*. Esto es problemático, sobre todo, en cuanto a la *primera* adquisición del concepto que caracteriza al artefacto. Supongamos, como parece plausible en muchos casos, que el primer individuo en adquirir el concepto sobre un cierto tipo de artefacto sea el creador de ese tipo de artefactos. ¿En qué sentido el conocimiento que posee sobre ese artefacto es *a priori*? Parece que el propio proceso de fabricación de un artefacto nuevo conlleva ciertas pruebas y ajustes hasta poder conocer cuáles son las características definitorias del artefacto. La especificación de tales características definitorias no parece algo análogo a la estipulación de las características definitorias asociadas con ‘soltero’ o ‘cazador’, por ejemplo. Y si el primer hablante que conoce la descripción que selecciona el referente del término es un usuario diferente al creador tampoco se ve cómo la descripción que asocia es conocida *a priori* por él. Por otro lado, la posibilidad de descubrir nuevas funciones para artefactos ya conocidos hace que la especificación de las funciones descubiertas no sea conocida *a priori* por nadie.

3.2.3 Crítica al argumento de las generalizaciones estables

El criterio que proporciona Schwartz para distinguir entre términos de clase artificial y de clase natural, diciendo que los segundos, pero no los primeros, pueden ser sujetos de generalizaciones estables, no es concluyente. Si consideramos, como propone Putnam, que la descripción que asociamos a ‘lápiz’ es ‘artefacto con una cierta forma hecho para escribir’, entonces, ‘los lápices son artefactos con una cierta forma hechos para escribir’ sí podría ser una generalización estable porque *podría* cumplir las dos condiciones que pide Schwartz para las generalizaciones estables: expresar una verdad (i) metafísicamente necesaria y (ii) epistémicamente corregible. Para Schwartz, ‘los lápices son artefactos con una cierta forma hechos para escribir’ no es una generalización estable porque incumple la condición (i). Como expuse en el **capítulo 2**, su motivo para

defender esto es que un solo *lápiz* que no fuera un artefacto, porque ha crecido en un árbol, por ejemplo, refutaría la afirmación de que los lápices son artefactos hechos para escribir. Pero no parece obvio que los lápices no sean necesariamente artefactos. Ni es obvio que un único *lápiz* con un origen natural refutaría la afirmación. Putnam, como se ha visto, tiene la intuición contraria: si los lápices son artefactos, entonces son necesariamente artefactos. Quien comparta la intuición putnamiana, podrá concluir que ‘los lápices son artefactos hechos para escribir’ es una verdad metafísicamente necesaria. Y, como también es epistémicamente contingente, tal como admite Schwartz, sí será una generalización estable.

Por otro lado, si consideramos, como hace Schwartz, que la descripción asociada con ‘lápiz’ es ‘objeto con una cierta forma que sirve para escribir’, sin especificar si los lápices son artefactos o no, tampoco está claro que no sea una generalización estable. Podríamos considerarla, como hace Schwartz, una verdad metafísicamente necesaria pero, sin embargo, epistémicamente corregible (contra lo que dice Schwartz). La razón para considerarla epistémicamente corregible podría venir de la posibilidad de algún argumento de ignorancia o error respecto a la función de los lápices. Schwartz presupone que no existe un tal argumento y, por eso, considera la oración epistémicamente incorregible. Desde mi punto de vista, si cabe la posibilidad o no de tales argumentos es algo que hay que explorar. Si fuera imposible o incoherente presentar este tipo de argumento, Schwartz tendría razón en que ‘los lápices son objetos con una cierta forma que sirven para escribir’ expresa una verdad incorregible. Pero habrá que ver si esto es así. De todas formas, hay que notar que la discrepancia en opiniones acerca de cuál es la descripción que asociamos con lápiz, discrepancia que se da entre Putnam y Schwartz, se basa en diferentes intuiciones metafísicas sobre los lápices y supone una dificultad para el descriptivista que, como Schwartz, afirma que conocemos su naturaleza.

En resumen, considero que el descriptivismo de Schwartz no es viable, en primer lugar, por su carácter individualista. Pero aun dejando esto a un lado, Schwartz tiene que asumir otros problemas. Primero, tiene que explicar por qué su análisis metafísico de los artefactos, en el que éstos presentan una naturaleza muy diferente a la de los objetos naturales, justifica una semántica de tipo diferente a la que acepta para los términos de clase natural. Segundo, como expuse en el **capítulo 2**, creo que no se sostiene que los

términos para artefactos son términos de clase nominal como ‘soltero’, pues, como vimos, las definiciones de tales términos no proporcionan condiciones necesarias y suficientes de aplicación de los términos correspondientes. También es difícil justificar para Schwartz que las descripciones que asociamos con los términos de clase artificial son, como dice, conocidas *a priori* y expresan verdades analíticas. Y, por último, su argumento de las generalizaciones estables, para distinguir entre términos de clase natural y términos de clase artificial no es concluyente.

3.3. La teoría híbrida de la referencia de Thomasson

3.3.1. Descriptivismo *sui generis*

Thomasson también presenta una alternativa a la propuesta semántica externista putnamiana para los términos de clase artificial. Desde mi punto de vista, la propuesta de Thomasson se enmarca dentro del descriptivismo semántico, pues, como se verá, sin el papel que juegan las descripciones asociadas a los términos para artefactos, no podría determinarse la referencia de estos términos. Como expondré a continuación, se trata de un descriptivismo de tipo *social o comunitario* en el que no todos los hablantes competentes de la comunidad lingüística asocian al término una descripción correcta de la clase artificial en cuestión, sino sólo ciertos expertos (los creadores de los artefactos). Thomasson incorpora, así, en su teoría, la deferencia lingüística como un mecanismo *externo* a los hablantes -mecanismo tomado de la teoría putnamiana⁵⁸ - a la hora de determinar la referencia de determinados usos no expertos de los términos de clase artificial. Sin duda este factor hace que la teoría de Thomasson sea más plausible y atractiva que una teoría como la de Schwartz, pues Thomasson admite sin problemas que algunos hablantes puedan equivocarse o ser ignorantes sobre la naturaleza de un cierto tipo de artefactos y, no obstante, puedan referir al usar el correspondiente término de clase artificial. No hay problema en que la referencia exitosa se produzca gracias a la deferencia a los expertos en el tipo de artefactos en cuestión.

⁵⁸ Como se verá, la diferencia con la propuesta putnamiana, respecto a la deferencia a los expertos, es que para Thomasson, a diferencia de lo que piensa Putnam, el conocimiento que tienen los expertos sobre las características distintivas de las clases es incorregible.

Ahora bien, según Thomasson, su teoría no es una mera teoría descriptivista de tipo social pues, como veremos, pretende incorporar otros elementos *externos* en la determinación de la referencia de estos términos. Incluso en la determinación de la referencia de los términos de clase artificial, cuando son usados por los expertos, aparte de la descripción asociada por el experto, hay elementos *externos* a la mente del hablante (y que pueden ser cognitivamente inaccesibles) que también juegan un papel a la hora de determinar la referencia de estos términos. Así, la descripción que los expertos asocian a una clase artificial puede no ser suficiente (aunque sí es necesaria) para determinar la referencia del correspondiente término; otros factores que se escapan al conocimiento de los expertos pueden jugar un papel al determinar la referencia. Por este motivo, Thomasson considera que su teoría para la determinación de la referencia es una teoría *híbrida*: una teoría que mezcla un elemento descriptivista y un elemento externista. En las secciones que siguen, expongo su teoría.

3.3.2. La propuesta metafísica

Thomasson, al igual que Schwartz, parte de la metafísica de los artefactos para llegar a conclusiones semánticas internistas sobre los términos de género artificial. Primero, considero su propuesta metafísica y, después, reviso cómo extrae consecuencias epistemológicas y semánticas sobre los términos para artefactos a partir de la naturaleza de los objetos artificiales. En una primera aproximación, Thomasson considera que una característica destacable de los artefactos, en contraste con las clases naturales, es que son productos de las actividades humanas. Pero, estrictamente, se trata de productos *intencionados* de las actividades humanas, pues hay productos humanos, como, por ejemplo, la polución, que no son intencionales, sino más bien una mera consecuencia de nuestras actividades y no pueden considerarse propiamente artefactos. Continuando con su análisis, lo que constituye la esencia de los artefactos no es su forma y función características, sino las intenciones y los conceptos de los creadores de los artefactos. Un artefacto es el tipo de artefacto que es porque así lo ha determinado la intención de su creador, intención que ha estado guiada por un concepto sustantivo que tiene el creador sobre qué es ser ese tipo de artefacto. Este concepto, aclara Thomasson, especifica al menos las condiciones necesarias⁵⁹ para que el artefacto pertenezca a una

⁵⁹ Más adelante, intento aclarar por qué quizás el concepto del creador no especifica las condiciones suficientes de pertenencia a la clase artificial.

determinada clase. Thomasson define del siguiente modo lo que es para un objeto x pertenecer a una clase artificial K :

Necessarily, for all x and all artifactual kinds K , x is a K *only if* x is the product of a largely successful intention that (Kx) , where one intends (Kx) *only if* one has a substantive concept of the nature of Ks that *largely* matches that of *some group* of prior makers of Ks (if there are any) and intends to realize that concept by imposing K -relevant features on the object.
(Thomasson 2003, 600)

En la definición de Thomasson, para que un objeto x sea un artefacto de una determinada clase K , es necesario que se den dos condiciones: (i) la existencia de la intención por parte del creador de producir un K , y (ii) que tal intención sea llevada a cabo con éxito. Y para que esto último pueda ocurrir, el creador debe tener un concepto sustantivamente correcto de lo que es ser un K . Cuando la clase artificial K existe previamente, el concepto del creador debe ajustarse, si no completamente, al menos de manera *sustantiva* al concepto de los creadores previos de Ks . No es suficiente que tenga una intención desnuda de hacer meramente *uno de estos artefactos* porque, dice Thomasson, tal intención podría no ser necesaria, ya que los prototipos también son miembros de una clase artificial y gente de diferentes culturas puede coincidir en producir un artefacto de la misma clase. Y tal intención, añade Thomasson, podría no ser suficiente porque un artefacto tiene que ser el producto de un proceso controlado de fabricación que implica imponer un número de características pretendidas en el objeto. Así que el creador no puede sólo pretender hacer *uno de éstos* sin tener ninguna idea de qué características son relevantes para ser *uno de éstos*. Si el creador que pretende hacer un K conoce la existencia de Ks previos, su concepto de qué es un K debe ajustarse *amplia y sustantivamente* al concepto de los creadores precedentes. De otro modo, su intención de hacer un K es fallida. Si no conoce la existencia previa o simultánea de otros Ks , está en la misma situación que el creador de un prototipo y, si su concepto coincide en lo esencial con el concepto de los otros creadores de Ks , podemos considerarle un creador de Ks coincidentalmente. Así es como miembros de diferentes culturas pueden crear artefactos de la misma clase (por ejemplo, cuchillos) guiados por el mismo concepto sustantivo.

Pero, ¿qué sucede en el caso de los prototipos o artefactos nuevos? Cuando el creador está haciendo un prototipo de una clase artificial de su invención, no puede, obviamente, tener una intención de hacer *uno de éstos* previamente existentes. Pero debe tener, dice Thomasson, ciertos objetivos a los que dirigir su actividad, incluyendo algunas características que imponer al objeto creado. Debe tener alguna idea *sustantiva* de qué tipo de artefacto pretende crear. En este caso, las características relevantes para la pertenencia a la clase artificial son estipuladas por el creador. Y si consigue imponerlas y alcanzar sus objetivos, podemos considerarle un creador de una clase artificial.

Así, según Thomasson, su definición de qué es para un objeto x ser de una clase artificial K , incluye tanto a los prototipos como a los artefactos producidos cuando ya existen otros de la misma clase. También incluye a miembros rotos o estropeados de una clase artificial, pues la intención exitosa del creador al hacer el artefacto no excluye que éste se pueda romper o estropear con el tiempo⁶⁰. En resumen, la idea de Thomasson es que, para que un objeto determinado sea un artefacto de una cierta clase, es necesario que sea el producto de una intención exitosa –que implica un concepto sustantivamente correcto–, por parte de su creador, de hacer un artefacto de esa clase. Así, para que un objeto sea una silla es necesario que sea el producto de la intención (exitosa) de su creador de hacer una silla (donde la intención exitosa ha sido guiada por un concepto sustantivamente correcto de lo que son las sillas).

La naturaleza de los artefactos está determinada, entonces, de acuerdo con Thomasson, por la intención del creador y no por su función, como por ejemplo, defiende Schwartz. Thomasson rechaza la función de los artefactos como lo esencial a las clases artificiales porque, como dije en el capítulo anterior, en la sección 2.2.4, la función, incluso entendida de diferentes maneras, no proporciona ni las condiciones necesarias ni las condiciones suficientes para que un artefacto sea miembro de una clase artificial. Sin embargo, Thomasson no se opone a que, a menudo, la función es lo definatorio de los artefactos, pero sólo cuando la función es una característica contemplada en el concepto que el experto tiene en la mente y quiere imponerla al artefacto de su creación. Pero esto no siempre será así. Es decir, a veces, el concepto del artefacto que el creador tendrá en

⁶⁰ Sí excluye, en cambio, a los artefactos que son producidos defectuosamente por su creador, pues no se ajustan al criterio de realización razonablemente *exitosa* que Thomasson exige.

la cabeza no incluirá su función, o incluirá la función con una serie de otras muchas características. Como la esencia de las clases artificiales está determinada por las intenciones de sus creadores respectivos, es razonable pensar que estas intenciones supondrán conceptos diferentes para cada clase artificial que, a veces, según la clase en cuestión, incluirán la función del artefacto, a veces sólo características físicas, en otros casos una conjunción de propiedades físicas y de la función, etc.

Por tanto, según Thomasson, decir que la función (ya sea la función pretendida por el creador o la función que acaba realizando el objeto en un cierto contexto) es la esencia de los artefactos no es correcto. Que algo sea una silla no depende de que tenga de hecho la función de las sillas o que su creador haya querido que se use como silla. Ya vimos en la sección 2.2.4 del **capítulo 2** que hay claros contraejemplos a estas situaciones: algo puede ser una silla aunque no se use para sentarse o aunque su creador no haya querido que se use para sentarse sino como objeto decorativo. Para que algo sea una silla, según Thomasson, lo necesario (aunque quizás no suficiente) es que su creador haya tenido la intención (llevada a cabo exitosamente) de hacer una silla, independientemente de si después el objeto es usado o no para sentarse.

Es importante apreciar este punto de la teoría de Thomasson para entender que no está diciendo lo mismo que una propuesta que defienda que lo definatorio de un tipo de artefactos es su función pretendida, la función que pretendió imponerle su creador⁶¹ (ver sección 2.2.4). Según este último punto de vista, si algo es un cierto tipo de artefacto, entonces tiene una cierta función pretendida (no necesariamente una cierta función actual); esto es, su creador tuvo la intención de imponerle una cierta función (que luego sea usado con otra finalidad es otra cuestión). Así, por ejemplo, para la explicación millikiana de la función pretendida, si algo es una silla, entonces su creador quiso imponerle la función histórica de las sillas (servir como asiento). Es decir, la intención del creador de imponer a un objeto la función de servir como asiento es una condición necesaria para que el objeto sea una silla. Pero, como critica Thomasson, esta explicación tiene contraejemplos. La función pretendida no es una condición necesaria para que un artefacto sea el tipo de artefacto que es: algo puede ser una silla aunque su

⁶¹ Esto es lo que defiende Millikan en última instancia, según la interpretación de Thomasson.

creador no quiera que la gente se siente, sino que sólo pretende, por ejemplo, que sirva como objeto decorativo.

En cambio, según la propuesta de Thomasson, para que algo sea una silla sólo es necesario que haya sido creada por un creador con la intención (llevada a cabo exitosamente) de hacer una silla. Como vimos, la intención de hacer una silla, por parte del creador, implica la posesión de un concepto *sustantivamente* correcto de lo que es una silla. Quizás este concepto incluya, entre otras características, la función de servir como asiento. Si es así, entonces también será una condición necesaria, para que algo sea una silla, que su creador haya querido imponerle la función de servir como asiento. Pero la propuesta de Thomasson excluye que la naturaleza de un tipo de artefacto esté determinada necesariamente por la función que haya querido imponerle el creador, pues es claro que algo puede ser una silla aunque su creador tenga otros planes muy distintos para el objeto que el servir para que la gente se siente encima. Para Thomasson, lo único necesario para que un objeto sea una silla es que su creador haya tenido la intención (guiada por un concepto correcto y ejecutada con éxito) de hacer una silla. Lo que no puede ser de ninguna manera, de acuerdo con Thomasson, es que el creador haya tenido la intención (exitosa) de hacer una silla guiado por un concepto sustantivamente correcto de lo que son las sillas, y, pese a todo, no haya conseguido hacer una silla. Ahora bien, la intención del creador de *hacer uno de estos*, guiada por un concepto sustantivo, siendo condición necesaria para que un objeto sea el tipo de artefacto que es, puede no ser condición suficiente. Thomasson cree que, en algunos casos, otros factores, como la intención de los usuarios para la aceptación o perfeccionamiento del artefacto en cuestión, podrían también ser necesarios para que un objeto sea el tipo de artefacto que es. Es decir, Thomasson no descarta que, para que algo sea un cierto tipo de artefacto, además de lo que aporta el creador, pueda ser necesaria la aceptación colectiva o algún otro factor social que no especifica demasiado.

3.3.3. Derivación de conclusiones epistemológicas

Partiendo de la naturaleza necesariamente intencional de los artefactos, Thomasson llega a una conclusión epistemológica: el creador del artefacto no puede estar equivocado ni ser ignorante, de manera *masiva*, sobre la naturaleza del artefacto que ha creado. Como creador del artefacto tiene que conocer cuáles son las características esenciales del artefacto que ha creado. El tipo relevante de intenciones para hacer una

cosa de la clase artificial *K* ha de implicar un concepto sustantivamente correcto de qué es un *K*, y una intención, realizada con un éxito suficiente, de imponer muchas de las propiedades relevantes de los *Ks* en el objeto creado. Por tanto, si la intención del creador es llevada exitosamente a cabo, su concepto sobre la clase artificial en cuestión no puede estar equivocado, pues es el que guía el proceso de producción del artefacto. Este concepto o material descriptivo proporciona al fabricante un cierto tipo de inmunidad ante la ignorancia y el error sobre la naturaleza de los artefactos. Naturalmente, el resto de los hablantes de la comunidad sí pueden ser ignorantes o estar equivocados sobre esta descripción de la naturaleza de una clase artefactos.

Thomasson, a diferencia de Schwartz, no dice que las descripciones o los conceptos que conocen los creadores de los artefactos sean accesibles *a priori* ni que estén analíticamente asociadas con el término para artefacto. Thomasson defiende un privilegio epistémico de los creadores respecto a los artefactos que crean, pero este conocimiento tiene unos límites, dice. Por ejemplo, el conocimiento de cualquier creador particular está limitado a su época y tradición; los artefactos pueden cambiar gradualmente a lo largo del tiempo y el creador de un artefacto no tiene por qué conocer cómo evolucionará el artefacto que creó en su día. Thomasson admite que el concepto *final* que posee un creador o creadores acerca de una clase de artefactos puede haberse pulido con el paso del tiempo, pues puede ser preciso cierto tiempo para que el artefacto se ajuste bien a una cierta finalidad o a los gustos de la gente, etc. Los creadores también pueden ser ignorantes o equivocarse sobre qué condiciones físicas, por ejemplo, son nomológicamente necesarias (o suficientes) para que los artefactos tengan las características superficiales que presentan. Igualmente, los creadores también pueden ignorar muchos otros hechos *colaterales* sobre los artefactos que crean, como su valor cultural o sociológico, por ejemplo. Pero el punto más importante, que desarrollaré al hablar de la propuesta semántica de Thomasson, es que los creadores pueden desconocer ciertas características que podrían jugar un papel a la hora de determinar la referencia del término para artefacto. Sin embargo, de acuerdo con Thomasson, el desconocimiento de estas características no puede en ningún caso calificarse de ignorancia o error *sustanciales* sobre la naturaleza de la clase artificial. En la siguiente sección, intento explicar este punto.

¿En qué consiste este conocimiento *sustancial* y privilegiado que tienen los creadores sobre los artefactos que han creado? Como vimos, los creadores de un tipo de artefacto poseen un concepto (sustantivamente correcto) que proporciona condiciones al menos necesarias de pertenencia a la clase artificial en cuestión. Es decir, los creadores conocen una serie de características necesarias de los artefactos que han creado y, como se trata de características necesarias a los artefactos, no es posible la ignorancia o el error sobre ellas por parte de los creadores. ¿Cuáles son estas características necesarias que el creador conoce (necesariamente) y sobre las que no es posible la ignorancia o el error? Thomasson no las explicita completamente, ni pone ningún ejemplo sobre ello, pero habla de estas características como de *características generales* o *de primer nivel de conocimiento*. En un primer nivel de conocimiento, el creador de un artefacto conoce qué *tipo de* características son relevantes para esa clase de artefacto en cuestión: *si* la forma, la función, la conjunción de la forma y la función, etc. Respecto al conocimiento de estas propiedades *generales*, especificadas en el concepto del creador, no es posible la ignorancia o el error.

Pero hay un segundo nivel de conocimiento sobre la naturaleza de los artefactos, en el que se especifican las características *particulares* o concretas de la clase (una determinada forma, una función bien definida, etc.) que no es inmune a la ignorancia y al error (al menos, en un primer momento) por parte del creador. La idea de Thomasson es que el creador, cuando se trata de un artefacto nuevo, conoce de manera infalible qué tipo de características son relevantes para ese tipo de artefacto (si la forma más la función, o sólo la función, o la función y el material, etc.), pero puede desconocer qué forma *exacta* necesita un artefacto de esa clase o qué función *precisa* tiene. Llegar a este nivel de conocimiento más concreto sobre las propiedades de una clase artificial puede requerir un cierto tiempo, y el conjunto de la sociedad puede contribuir a decidir estas características con más precisión. Pero, con el paso del tiempo, los conceptos de los creadores -conceptos pulidos, acabados- también recogerían estas características de segundo nivel, y sus conceptos, según Thomasson, seleccionarían las clases artificiales proporcionando las condiciones necesarias y también suficientes de pertenencia a ellas.

A continuación, en la siguiente sección, al exponer la teoría semántica de Thomasson, espero que quede más clara esta cuestión de las características de *primer nivel* y de *segundo nivel* de concreción.

3.3.4. Conclusiones semánticas: teoría híbrida de la referencia

A partir del cierto privilegio epistémico del creador, el siguiente paso que da Thomasson es de tipo semántico: la descripción que conoce el creador de un artefacto y que proporciona las condiciones necesarias de pertenencia a la clase es la que determina⁶² la referencia del correspondiente término de clase artificial. Como siempre que hay un artefacto, hay un creador que necesariamente conoce la descripción de su naturaleza (al menos, en un nivel *general*), esa descripción es la que determina la referencia del término de clase.

[...] would-be grounders of the reference of a general term must have some high-level concept of what category of kind they intend to refer to, thereby establishing the general *sorts* of features relevant to unifying members of that kind. This gives those who ground (and re-ground) the reference of the term some privileged knowledge about what sort of kind (if any) their terms refer to, and what sorts of features unify its members. This much, of course, applies equally to any general term (not just artifactual kind terms).

(Thomasson 2009a, 64)

Ahora bien, si esa descripción hace referencia sólo a las características más *generales* de la clase artificial, al menos en un momento inicial de introducción del tipo de artefacto, sin concretar las características más *particulares*, entonces sería en principio posible que esa descripción general no seleccionara, por sí sola, la extensión del término pues ciertas propiedades más *particulares* del artefacto podrían jugar un papel al determinar la referencia del término. Thomasson hace la siguiente consideración respecto a los términos generales (cosa que incluye a los términos de clase artificial):

[...] general terms may pick out their referents independently of any human concepts about the particular nature (though not the category) of the kind involved; regarding that nature, everyone may remain ignorant or in error.

(Thomasson 2009a, 65)

Así, propiedades del mundo, que podrían escapar al acceso cognitivo incluso del creador, podrían jugar un papel también a la hora de determinar la referencia de los

⁶² Al menos, parcialmente, como veremos.

términos para artefactos. Por ello, Thomasson propone lo que llama una ‘teoría híbrida’ de la referencia para los términos de clase artificial. Es *híbrida* porque, en un *primer nivel* de concreción sobre las características *generales* que determinan la pertenencia a la clase artificial, la referencia de los términos para artefactos viene acotada o parcialmente determinada por los conceptos o descripciones que poseen los creadores sobre qué tipo de características *generales* son relevantes para ser un miembro de la clase artificial. Pero como estos conceptos no especifican (al menos, no necesariamente) las características *particulares* de la clase (esta forma concreta, esta función concreta...), en un segundo nivel, los términos generales podrían seleccionar sus referentes independientemente de cualquier concepto humano sobre la naturaleza *particular* (aunque no sobre la naturaleza *general*) de la clase. Respecto a esta naturaleza particular todo el mundo podría permanecer equivocado o ignorante cuando se introduce el término. Con el paso del tiempo, dice Thomasson, los conceptos específicos de los fabricantes también seleccionarían la referencia de los términos para artefactos en este segundo nivel.

Además de su concepción metafísica y epistemológica sobre los artefactos, Thomasson encuentra una fuerte motivación para defender una teoría híbrida de la referencia en lo que se ha llamado ‘el *qua problem*’. El *qua problem* se presenta en el momento de la introducción de un nuevo término en el lenguaje y consiste, según aquellos que lo plantean, en la absoluta indeterminación de la referencia cuando se introduce el término por primera vez. Para entender el problema, paso a exponerlo como lo hacen Devitt y Sterelny (1987), de los que Thomasson lo acepta como un problema genuino al que la teoría externista causal de Kripke y Putnam no da una solución satisfactoria. Devitt y Sterelny exponen el *qua problem* de la siguiente manera⁶³. En la introducción de un término, un nombre propio, digamos, hay, según Devitt y Sterelny, un requerimiento de un conocimiento mínimo sobre el referente: el introductor debe pensar en el objeto bautizado bajo algún término categorial *general* como ‘animal’ o ‘persona’, por ejemplo⁶⁴. Parece, dicen, que el introductor del término singular *debe* conocer qué clase

⁶³ Tal problema, según ellos, se aplica tanto a los nombres propios como a los términos de clase natural y de clase artificial. Thomasson, al final, defiende lo mismo.

⁶⁴ El enlace causal fallará si la causa de la experiencia perceptual no se ajusta a los términos categoriales usados para conceptualizarla.

de objeto tiene delante (si es un animal, una persona, un objeto material, etc.). En el caso de la introducción de los términos generales, también debe de haber algún conocimiento implícito de que la muestra es un miembro de una clase natural o de una clase artificial. Si se trata de un miembro o muestra de una clase natural, por ejemplo, el concepto categorial *clase natural* asociado al término establece que las propiedades relevantes para pertenecer a la clase y determinar la futura referencia del término tienen que ver con la estructura interna de la muestra. Es decir, no sólo se trata de asociar la categoría *general* al término, sino que la categoría lleva asociada (se supone que en virtud de ser hablantes competentes) el tipo de características que son relevantes para que algo sea un objeto natural de una cierta clase (el tipo de característica, en este caso, es la estructura interna o microfísica). La idea que anima a defender esto es que una cierta muestra a la que categorizamos como *objeto natural* representa *múltiples muestras*: un objeto con una cierta estructura interna, una cierta configuración de la materia, una superficie de un cierto color, etc. Pero cuando concebimos un objeto bajo la categoría de *objeto de una clase natural*, la propiedad que nos parece *relevante* para que pertenezca a una determinada clase natural, de entre todas las propiedades que ejemplifica el objeto, es la de tener una cierta composición microfísica. Es a esta característica de tipo *general* a la que pretendemos referir cuando introducimos el correspondiente término. Según Devitt y Sterelny, lo que selecciona la muestra como un miembro de una cierta clase natural, con la estructura interna como propiedad relevante para pertenecer a esa clase, es el estado mental del introductor del término que está en contacto causal con la muestra.

Obviamente, las mismas observaciones, respecto al *qua problem*, se pueden hacer con la introducción de los términos de clase artificial. Tanto para Devitt y Sterelny, como para Thomasson, al introducir un nuevo término para artefacto, la referencia de éste quedará indeterminada si el introductor no tiene un concepto mental asociado sobre la categoría *general* a la que pertenece ese tipo de cosa y qué tipo de características son relevantes para ser ese tipo de cosa. Para Devitt y Sterelny, las propiedades relevantes asociadas con la categoría *artefacto* son la forma y la función; para Thomasson, son propiedades que, en cada caso, determina el creador del artefacto y nadie más (aunque, previsiblemente, estará incluida la función la mayoría de las veces). Así, los introductores, ante el artefacto en cuestión, asocian un concepto categorial, el concepto *artefacto*, que determina que sean su función y su forma -y no, pongamos por caso, su

estructura interna o su estatus legal- las características seleccionadas para referir a ese tipo de artefactos. Si no se produce una desambiguación respecto a qué tipo de cosa (con sus respectivas cualidades relevantes), entre todas las que ejemplifica la muestra, pretende referir el término, la referencia no puede llegar a determinarse. En resumen, el introductor de un término general, sea de clase artificial o natural, asocia⁶⁵ con ese término: (i) alguna descripción que clasifique el término como de clase natural o de clase artificial, y (ii) algunas descripciones que determinan qué naturaleza de la muestra (la estructura interna, o la función) es relevante para la referencia del término.

Una vez expuesto el *qua problem*, me centro en la teoría híbrida de la referencia. Tanto Devitt y Sterelny como Thomasson consideran que el *qua problem* es un problema genuino que da lugar a, si no una teoría descriptivista *pura*, al menos a una teoría híbrida de la referencia, quedando descartada una mera teoría causal como la que pretende Putnam⁶⁶. Con el fin de aclarar al máximo cómo funciona una teoría híbrida de la referencia, y puesto que Thomasson no desarrolla demasiado cómo funciona la que ella propone, expongo, en primer lugar la que proponen Devitt y Sterelny, y, después, la de Thomasson. Para los términos de clase artificial, Devitt y Sterelny no se deciden entre una teoría descriptivista pura o una teoría híbrida (con un componente descriptivo y otro causal)⁶⁷, pero exploran esta última opción. ¿Cómo funcionaría, según ellos, esta teoría causal descriptiva para los términos de clase artificial? Proponen que el elemento descriptivo es la descripción categorial, junto con el tipo de características relevantes asociadas: la clasificación del objeto como artefacto y, por tanto, como algo perteneciente a una clase delimitada por la forma y función. Como la forma de un

⁶⁵ Para Devitt y Sterelny, esta asociación puede ser consciente o inconsciente.

⁶⁶ Al menos en lo que respecta a la introducción del término o fijación inicial de la referencia; respecto a la transmisión del término u otros usos posteriores, podrían funcionar diferentes teorías, incluida una teoría como la de Putnam, según estos autores.

⁶⁷ Afirman que para los términos de clase artificial, quizás, podría funcionar un descriptivismo tradicional o un descriptivismo *cluster*, pero no descartan un descriptivismo causal o teoría híbrida. Para lo que ellos llaman ‘términos de clase artificial no básicos’ (los que se definen en términos de otros términos de clase artificial y de la forma), como ‘yate’, apuestan por un descriptivismo tradicional. En cambio, para los términos básicos para artefactos, como ‘silla’, ‘lápiz’ o ‘mesa’, los que se definen apelando a la forma y la función, podría funcionar una teoría híbrida de la referencia.

artefacto es una característica superficial, es algo que los introductores conocen y forma parte, también, del material descriptivo asociado con el término a la hora de introducirlo en el lenguaje. Así, por ejemplo, el elemento descriptivo presente en la fijación de la referencia de ‘silla’ hace referencia al tipo de objeto que son las sillas (artefactos y, por tanto, objetos cuyas características relevantes para la referencia son la forma y la función), más una serie de características superficiales, perceptibles (una base plana con respaldo que se apoya en el suelo con patas).

Entonces, ¿cuál es el elemento externo, causal, según Devitt y Sterelny, que también juega un papel en la fijación de la referencia del término y que emerge de la relación causal ostensiva en la introducción del mismo? La función. Mientras en una teoría descriptivista *pura*, la función de los artefactos, que determina la referencia, se fijaría por una descripción de la función, dicen Devitt y Sterelny, en la teoría causal descriptivista que ellos exploran, la función queda fijada por los objetos presentes en la muestra cuando se introduce el término, y no por una descripción de la función. Así, el nuevo término para artefacto refiere a cualquier artefacto con las mismas características físicas *relevantes* (especificadas en una descripción a la que los hablantes-introductores tienen acceso cognitivo) y con la misma función que el objeto, u objetos, de la muestra (función que queda determinada ostensivamente). Como la función no forma parte de la descripción asociada al término, sino que es referida de manera ostensiva, esto, en principio, permitiría que el introductor de un término de clase artificial pueda desconocer la función de la clase artificial en cuestión. Sin embargo, admiten que esto no es probable y que es muy difícil encontrar este tipo de ejemplos en que los introductores de un término para artefacto desconocen o están equivocados sobre la función que tiene⁶⁸.

Thomasson, a diferencia de Devitt y Sterelny, no duda que una teoría híbrida de la referencia es la apropiada para explicar cómo se fija la referencia de los términos para artefactos en el momento de su introducción. La referencia, entonces, de los términos de clase artificial se determinaría mediante un elemento descriptivo, y un elemento causal.

⁶⁸ Sin desarrollar más esta cuestión, apuntan que el mejor sitio para encontrar este tipo de ejemplos de desconocimiento de la función de un tipo de artefacto es en casos donde las creencias religiosas o sociales mistifican la función de ciertos artefactos.

El elemento descriptivo, como en la teoría de Devitt y Sterelny, estaría constituido por una descripción que especifica el tipo de cosa a la que los introductores pretenden referir, junto con el tipo de características, a las que Thomasson llama ‘generales’, que son relevantes para que el objeto sea de la clase artificial que es. Así, el introductor de un término para artefacto, desambigua la referencia mediante una descripción como, por ejemplo, ‘artefacto cuyas características relevantes son la forma y la función’⁶⁹. Si esas características son superficiales (como la forma), en principio, la descripción del introductor del término (el creador, típicamente) incluirá también, en un nivel más concreto, la especificación de esas características superficiales, como la forma.

Entonces, ¿cuáles son las características que en la teoría híbrida de Thomasson pueden determinarse de una manera causal? Desde mi punto de vista, no queda claro. Considero que hay dos candidatos para ser el tipo de características que pueden determinarse causalmente y que podrían no estar especificadas en el concepto del creador del término. El primer tipo de características serían las características no superficiales, a las que el creador no tiene un acceso a través de los sentidos; esto incluiría la función. E implicaría la posibilidad de que el creador pueda ser ignorante o estar equivocado sobre la función del artefacto que ha creado. Pero, a diferencia de Devitt y Sterelny, Thomasson no dice explícitamente que la función, al no ser una característica superficial, sea una característica que se pueda determinar de manera ostensiva, señalando a un artefacto y pretendiendo referir con el término correspondiente a todo artefacto que tenga la misma función que esa muestra *de ahí*. De hecho, Devitt y Sterelny, al explorar cómo sería una teoría causal descriptiva para los términos de clase artificial, admiten que la función podría ser algo desconocido para los introductores de un término, pero consideran que es una posibilidad muy remota, y, a fin de cuentas, no afirman rotundamente que una teoría híbrida sea la adecuada para los términos para artefactos. Yo no creo que Thomasson quiera admitir que la función es una propiedad que se determina causalmente, permitiendo la ignorancia o el error por parte del creador, pues entonces el concepto del creador sobre un cierto artefacto sería demasiado impreciso, especificando sólo que se trata de un artefacto (cuyas características

⁶⁹ Recordemos que, según el tipo de artefacto, el tipo de características generales relevantes para la pertenencia a la clase son unas u otras. Esto es algo que conoce (porque lo estipula) el primer creador de un tipo de artefacto.

generales distintivas respecto a un objeto natural son tales y tales) con tales y tales características superficiales. Por ejemplo, creo que un concepto sustancialmente correcto de los lápices, por parte del introductor del término ‘lápiz’, que guíe la producción del primer lápiz, no puede limitarse a especificar que los lápices son artefactos, cosas para las que lo importante, a la hora de referir con el correspondiente término, no es su naturaleza microfísica, sino su función y forma, digamos. El concepto del creador debe especificar algo acerca de esa función y esa forma concretas.

Realmente, si el creador de un artefacto pudiera estar *completamente* equivocado respecto a la función del artefacto que ha creado (pero no respecto a que el objeto es un artefacto, que la función juega un papel importante en que sea el tipo de artefacto que es y que tiene *tales* propiedades perceptibles), entonces no se ve dónde estaría el privilegio epistémico del creador, que Thomasson quiere defender, sobre otros usuarios del artefacto⁷⁰.

Esto nos lleva, al segundo tipo de características que considero como candidatos más probables, dentro de la propuesta thomassoniana, para determinar causalmente, y no descriptivamente, la referencia de los términos para artefactos. Por lo que ella comenta respecto al conocimiento del creador de un artefacto, parece que aquello que se le puede escapar cognitivamente al creador tiene más bien que ver con un nivel *alto* de concreción de las características distintivas de la clase artificial. Así, por ejemplo, parece que el concepto del creador de un artefacto incluye el tipo de cosa que es (un artefacto), junto con el *tipo* de características relevantes para ser el tipo de artefacto que es (digamos, la forma y la función, en un cierto tipo de artefacto), y *algún* nivel de concreción sobre en qué consisten ese tipo de características relevantes (qué forma y función *genéricas* tiene el artefacto). Por ejemplo, el concepto del primer creador de un cuchillo especifica que el objeto es un artefacto de un tipo tal que su forma, función y material son lo relevante para que sea el tipo de artefacto que es. Pero también parece que tiene que especificar, aunque sea de una manera ruda, cómo es esa forma

⁷⁰ De todos modos, aunque la propuesta de Thomasson defendiera lo mismo que la de Devitt y Sterelny, no sería una propuesta semántica externista como la de Putnam, por motivos que expondré después al hacer la crítica a la teoría de Thomasson. Aunque sí admitiría, de entrada, como reconocen Devitt y Sterelny, la posibilidad de *algunos* argumentos de ignorancia y error respecto a la naturaleza de los artefactos: aquellos argumentos que planteen la ignorancia y el error sobre la función del artefacto (y no, por ejemplo, sobre el hecho de ser un artefacto).

(alargada), el material (duro, cortante) y la función (cortar). Esta descripción del creador, sin embargo, podría no seleccionar únicamente la referencia del término porque podría haber otros artefactos parecidos físicamente y con la misma función genérica (cortar). Así que la especificación *bien concretada* de la función, el material, etc. puede ser algo que se determine ostensivamente ante la presencia de la muestra, sin que el creador tenga un conocimiento completo sobre ello.

Creo que esta segunda interpretación de cómo funciona su teoría híbrida de la referencia se ajusta más a lo que quiere defender Thomasson. De todos modos, la conclusión que a ella le interesa destacar es que los términos para artefactos no pueden referir a las clases independientemente de *todas* las creencias y los conceptos humanos sobre la naturaleza de las clases artificiales. En algún nivel, los conceptos humanos sobre la naturaleza de una clase artificial juegan un papel crucial en la referencia de los términos para artefactos.

3.4. Objeciones a la teoría thomassoniana

3.4.1 Crítica a la concepción metafísica y epistemológica de Thomasson

Una objeción que se puede plantear a la idea de fondo de la propuesta metafísica de Thomasson es que, aunque se admita que la naturaleza de las clases artificiales está conectada de alguna manera con las intenciones humanas, ¿por qué tienen que ser las intenciones de los creadores y no las de los usuarios lo que determina la esencia de los artefactos? Especialmente en el caso de un artefacto nuevo, el proceso de creación puede estar guiado por una determinada intención que incluye -como parece plausible- alguna noción sobre la función para ese artefacto, pero los usuarios de ese tipo de artefacto pueden tener una intención diferente al usarlo, otorgándole otra función. En casos así, la intuición nos dice que las intenciones que parecen conectadas con la naturaleza del artefacto son las intenciones de los usuarios. Dennet, por ejemplo, considera que el inventor o creador de un tipo de artefacto no es el árbitro final de lo que es un artefacto o de cuál es su función.

The inventor is just another user, only circumstantially and defeasibly privileged in his knowledge of the functions and uses of his device. If others

can find better uses for it, his intentions, clearheaded or muddled, are of *mere historical interest*. (Dennett 1990, 186)

Entonces, según Dennett, son los usuarios los que deciden cuál es la función de un artefacto y los que acaban determinando su naturaleza. Dennett pone el ejemplo del ordenador creado por Turing y Von Newman. Estos inventores podrían desaprobado el uso de su ordenador como procesador de textos, pero eso no cambiaría el hecho que el artefacto que crearon es un perfecto procesador de textos. El propósito particular de un creador en un momento dado es un hecho histórico que no tiene porqué establecer las funciones que puede y actualmente tiene el artefacto. En definitiva, Thomasson cae, de acuerdo con Dennett, en la llamada ‘falacia intencional’⁷¹. La falacia intencional, aplicada a los artefactos, consiste en atribuir a la intención que hay detrás de los artefactos -la intención del creador, para Thomasson- el papel de determinar qué son los artefactos o qué funciones tienen.

Pero, para Dennett, su crítica no alcanza sólo a la intención de los creadores, sino también a la intención de los usuarios. Pues aunque, según Dennett, son los usuarios, antes que el creador, los que tienen la tarea de determinar qué es o para qué sirve un artefacto, tampoco ellos son omniscientes respecto a su naturaleza, pues la utilidad de los artefactos puede ir cambiando a lo largo del tiempo. Si consideramos, como hace Dennett, que la naturaleza de un tipo de artefacto está constituida por algo así como su función *potencial*, o todas las funciones que podría llegar a desarrollar, está claro que nadie tiene un completo conocimiento de su naturaleza. Esto acaba con cierto privilegio epistémico, tanto por parte del creador como por parte de los usuarios, sobre la naturaleza de los artefactos. Así, dice Dennett, aunque el conjunto de los usuarios a lo largo de la historia va definiendo en cada momento histórico en qué consiste ser un determinado tipo de artefacto, ningún usuario individual tiene un privilegio epistémico sobre un tipo de artefacto, pues desconoce la historia completa de ese tipo de artefacto. Y lo mismo se aplica al conjunto de la comunidad de usuarios de un artefacto: tampoco ésta tiene un conocimiento privilegiado sobre un tipo de artefacto (quizá sí lo tiene sobre el uso de un artefacto en un momento histórico determinado), pues puede

⁷¹ El término fue pensado para aplicarse, entre otras cosas, en el ámbito natural y pretende denunciar la atribución de una intención finalista, o un diseño *inteligente*, a los procesos naturales. Dennet extiende el término al campo de los artefactos y de otras obras humanas (obras de arte, textos, etc.).

desconocer todo el potencial que tiene un artefacto determinado.

Sin embargo, Thomasson admite, como dije, que el privilegio epistémico que ella defiende que tiene el creador sobre el artefacto que ha creado no es total y tiene sus límites. En particular, reconoce que el creador puede desconocer las funciones posibles o futuras del artefacto que ha creado (pues puede desconocer todas las relaciones causales que el artefacto es capaz de establecer). Es decir, Thomasson podría estar de acuerdo con Dennett respecto a este límite del conocimiento del creador (o de toda la comunidad) respecto a un artefacto determinado. Ahora bien, lo que está aquí en cuestión es en qué consiste la naturaleza de las clases artificiales. Para Thomasson esa naturaleza no incluye de ninguna manera las funciones potenciales de un artefacto y, por tanto, no es problemático, en su explicación, que el creador las desconozca. No es preocupante para ella que el privilegio epistémico del creador no alcance hasta ahí, pues las funciones posibles de un artefacto no constituyen su esencia, de acuerdo con Thomasson. Pero hay otras concepciones metafísicas acerca de la naturaleza de los artefactos, concepciones que no ponen el acento en el aspecto intencional, como, por ejemplo, la de Dennett. Y, para quien considere, como él, que la naturaleza de los artefactos abarca toda su potencialidad, el desconocimiento por parte de todos de esa naturaleza global sí es relevante, pues indica que nadie conoce la esencia de los artefactos.

Otro problema relacionado con la función potencial que puede presentar la metafísica de Thomasson tiene que ver con la posibilidad de descubrir nuevas funciones para artefactos ya conocidos. En estos casos, la pertenencia a la (nueva) clase estaría determinada por alguna característica desconocida por el introductor o introductores del artefacto y descubierta después. Así, por ejemplo, retomando el caso que consideré en la sección 2.1.2 del **capítulo 2**, el cable eléctrico fue introducido para proporcionar luz en los espacios públicos y privados, pero décadas después se descubrió que también tenía la propiedad de transmitir voz y datos. Así, la pertenencia del cable eléctrico a la (nueva) clase artificial *transmisor de datos* no está determinada por la intención y el concepto sustantivo del creador del cable eléctrico. Naturalmente el cable eléctrico sigue teniendo la función para la que fue introducido, transmitir energía eléctrica. Pero, respecto a su nueva función descubierta, el creador no tuvo la intención de imponer una función y unas características que desconocía y, de este modo, las características

relevantes para la pertenencia a una clase no son siempre necesariamente, contra lo que dice Thomasson, objeto de estipulación, sino que pueden tener que ver con los poderes causales de los artefactos. En estos casos, cuando se descubre que un artefacto también pertenece a otra clase artificial por tener otra función, el concepto del creador no puede determinar la referencia del término para artefacto.

Resumiendo, creo que la metafísica para los artefactos que propone Thomasson tiene problemas tanto con los artefactos nuevos o prototipos a los que los usuarios dan otra función que se desvía de la que contempla el concepto del creador, como con los descubrimientos de nuevas funciones en artefactos conocidos. En estos casos, el concepto del creador no especificaría condiciones necesarias de pertenencia a la clase, contra lo que ella defiende.

Aún otro inconveniente de su teoría metafísica, respecto a la producción por parte de un creador de un tipo de artefacto ya conocido, es la cuestión de cuánto pueden apartarse las intenciones de un creador de las intenciones de los creadores precedentes del mismo tipo de artefacto. Recordemos que Thomasson exige que, para que un objeto sea un cierto tipo de artefacto (ya existente), es necesario que el concepto que tiene su creador (y que ha guiado el proceso de producción) se ajuste de manera *sustancial* al concepto de los creadores previos. Pero esto podría evitar, al menos en ocasiones, la evolución y los cambios que vemos que padecen los artefactos de un mismo tipo a lo largo de la historia. Algunas veces, los cambios que padecen un tipo de artefactos son tan graduales que Thomasson podría acomodar el hecho que el concepto de cada nuevo creador en la cadena no se ha apartado *sustancialmente* del concepto que tenían sus predecesores. Éste puede ser el caso, quizás, de lo que ha ocurrido históricamente con los ordenadores. Los primeros ordenadores no se parecen mucho a los ordenadores actuales; pero si se presentara este caso como uno en que el concepto que los creadores asocian con ‘ordenador’ ha cambiado *sustancialmente*, Thomasson podría decir que el cambio se ha producido en los conceptos de unos creadores a otros de manera muy gradual. Pero creo que también podría darse el caso de un cambio *sustancial* en el concepto de un tipo de artefacto conocido cuando un creador introduce de golpe alguna modificación o innovación sustancial (digamos, por ejemplo, en la forma del artefacto). En casos así, la intuición nos puede decir que estamos ante el mismo tipo de artefacto, para el que usamos el mismo término de clase artificial, pero que presenta una

modificación importante. Así que Thomasson no podría apelar convincentemente a que el último creador ha producido un nuevo tipo de artefacto precisamente por apartarse demasiado del concepto precedente. Creo que en estas situaciones posibles Thomasson tiene un problema, aunque quizá menor, para ajustar su teoría.

Finalmente, respecto a sus consideraciones epistemológicas, Thomasson dice que el creador necesariamente tiene que conocer al menos las características más *generales* que determinan la pertenencia del artefacto a la clase artificial, pues tiene que estar en posesión de un concepto que especifica qué características imponerle. Como vimos, no está del todo claro cuáles son estas características que el creador del artefacto conoce necesariamente. Algunas son características que tienen que ver con la naturaleza más *general* de cada tipo de artefacto, como la característica de *ser un artefacto* (junto con las características distintivas que ello implica para cada clase de artefacto). Otras son características plausiblemente más concretas (como qué forma, función, etc.), aunque quizás no bien concretadas *totalmente*. En ambos casos hay que notar que su concepción excluye ciertos argumentos de ignorancia y error respecto a la naturaleza de los artefactos, pues defiende que al menos los creadores conocen necesariamente estas características esenciales de la naturaleza de los artefactos que crean. Así que, sin duda, la mejor manera de intentar rebatir su propuesta epistemológica es exponer argumentos de ignorancia y error respecto a la naturaleza de los artefactos. Uno de ellos es el argumento de los lápices-organismos de Putnam, que expuse en la sección 2.4.1 y que vuelvo a considerar en la sección que sigue del presente capítulo. Presentaré otros argumentos de ignorancia y error, junto con un análisis de su alcance, en el **capítulo 4**.

3.4.2. Crítica a la teoría híbrida de la referencia

Empiezo considerando el *qua problem*, que es seguramente la motivación principal de quien adopta una teoría híbrida de la referencia. En primer lugar, cabe preguntarse si el *qua problem* es un problema genuino que demanda una solución como la considerada por Thomasson y Devitt & Sterelny. Según los formuladores del problema, éste radica en la absoluta indeterminación de la referencia al introducir por primera vez un término en el lenguaje. Como vimos, la solución propuesta es que, para determinar la referencia del término, es necesario que el introductor asocie algún concepto categorial; si no, la referencia queda indeterminada pues la muestra ejemplifica muchas clases diferentes. Creo que, sin duda, es cierto que la mayoría de las veces los introductores de un término

general, ya sea de clase natural o artificial, asocian algún tipo de información *general* sobre aquello a lo que pretenden referir. Ahora bien, también creo que la intuición nos dice que esa información podría estar *sustancialmente* equivocada. Así, por ejemplo, volviendo al experimento mental putnamiano de los gatos-robots, podría resultar que estuviésemos muy equivocados acerca de la naturaleza de los gatos, con un concepto categorial *general* sobre los gatos muy equivocado. Sin embargo, aunque los introductores del término ‘gato’ hubiesen estado muy equivocados sobre la verdadera naturaleza de los gatos, diríamos que el término ‘gato’, no obstante, consiguió referir desde un principio. Lo mismo ocurre con el experimento mental de los lápices organismos: el concepto categorial asociado a ‘lápiz’ por los introductores del término, *artefacto*, podría haber estado equivocado, pues podríamos haber descubierto que los lápices eran en realidad organismos. Los que consideran el *qua problem* como un problema genuino y le dan el tipo de solución propuesta por Thomasson no pueden admitir que sea posible equivocarse en la categoría general asociada a los lápices, la categoría de *artefacto*, sin que el término no consiga referir⁷². Sin embargo, creo que un término, de clase natural o artificial, puede referir aunque el concepto categorial asociado al término esté sustancialmente equivocado (como en los casos de los experimentos mentales mencionados). No creo que el *qua problem* plantee una dificultad que conduzca inevitablemente a una teoría descriptivista en alguna medida. Simplemente, la respuesta a la pregunta ‘¿Cómo se determina la referencia de un nuevo término en el lenguaje?’ puede ser ‘por la conexión causal entre el hablante –o el uso del término por parte del hablante- y la cosa presente con sus propiedades relevantes – sean las que sean.’ Los que postulan una teoría causal descriptiva admiten una parte *externa*, causal, perteneciente al entorno, que también contribuye a determinar la referencia del término. Pero, entonces, la cuestión es por qué no podría determinar *toda* la referencia.

Por otro lado, una complicación del *qua problem*, que reconocen Devitt y Sterelny, es que es un problema que puede reproducirse en niveles más concretos. Para que la referencia empiece a determinarse, es necesario, según ellos, que el introductor asocie

⁷² Hay que notar el contraste, en este caso, con la postura de Schwartz que, desde un descriptivismo clásico, admite – al menos, en principio- que el experimento mental de los lápices organismos es posible (incluso metafísicamente), pues el concepto *artefacto* no forma parte, para él, de la definición asociada a ‘lápiz’.

algún concepto de primer nivel o concepto general, como *objeto natural*, *artefacto*, etc. (junto con lo que implica ese concepto categorial: que las propiedades relevantes para la referencia del término son subyacentes, o funcionales, etc.). Pero es obvio que esta información asociada será insuficiente para seleccionar únicamente el referente del término, pues será una información igualmente aplicable a objetos de otras clases. Cualquier muestra de una clase es también una muestra de muchas otras clases. Así, por ejemplo, un tigre es un animal que, además, es vertebrado, mamífero, etc. Los introductores del término ‘tigre’ asocian a él la categoría general *animal* y pretenden referir a todos los objetos que tengan la misma naturaleza subyacente que la muestra, pero ¿qué naturaleza subyacente? Los ejemplares comparten muchas naturalezas subyacentes... En el caso de los términos de clase artificial, sucede lo mismo: el *qua problem* se puede ir reproduciendo hasta llegar a la referencia del término, pues una muestra puede ejemplificar muchas clases artificiales. Por ejemplo, un artefacto concreto, podría ser tanto un tope de puerta como un pisapapeles. Así que el *qua problem* puede reproducirse en niveles cada vez más concretos. Y esto es problemático porque no habría motivo para dejar de plantearlo en un cierto punto, llegando así hasta la determinación del referente mediante una descripción que ya seleccionaría únicamente la extensión del término. Pero en este caso, estamos ante un descriptivismo clásico, y no ante una teoría híbrida de la referencia.

Sin embargo, aun pasando por alto estos inconvenientes derivados de considerar el *qua problem* un reto serio, hay una objeción que se ha hecho a una teoría causal descriptiva como la de Thomasson. Es una original crítica que ha hecho Dodd⁷³ (2012) apoyándose en su interpretación de la teoría de Evans⁷⁴. Dodd concede que el *qua problem* es un problema al que hay que responder, pero, aun así, dice que una teoría híbrida como la de Thomasson no es la mejor solución al problema. Según Dodd, en la teoría de Thomasson hay presentes dos tesis semánticas diferentes⁷⁵, que se entremezclan. La

⁷³ La crítica que hace Dodd a Thomasson se centra en los términos para obras de arte (que, plausiblemente, podrían considerarse otro tipo de términos para artefactos). Aquí, yo extendiendo su tratamiento al caso de los términos de clase artificial que interesan aquí.

⁷⁴ La obra de Evans a partir de la que Dodd hace su interpretación y propuesta es ‘The causal theory of names’.

primera, (i), es que la referencia de un término sólo se puede fijar si el introductor asocia algún contenido conceptual mínimo sobre el referente. Como se vio, esta tesis permite a Thomasson evitar el *qua problem*, pues, para ella, una relación meramente causal entre el hablante y la cosa en el mundo no fijará la referencia del término correspondiente. La segunda tesis, (ii), la tesis clave en la crítica de Dodd, es que el mecanismo por el que el concepto del introductor del término fija la referencia del término es el *ajuste* del referente a lo que especifica el contenido conceptual. El concepto amplio que tienen los introductores del término funciona como un patrón (*template*) al que debe ajustarse el referente o los objetos de la extensión del término. El concepto establece, por tanto, unas condiciones que deben satisfacer las entidades a las que se aplica el término.

Dodd concede (i), pues, para él, el *qua problem* es un problema genuino que exige que, para que se determine la referencia de un término, los introductores de éste tengan un concepto básico de la clase de cosa a la que pretenden referir. El problema de la teoría de Thomasson, según Dodd, es (ii), pues si aceptamos esta segunda tesis, excluimos que la información asociada por los introductores pueda estar equivocada respecto a la verdadera naturaleza de la clase. En el caso de las clases artificiales, (ii) obliga a Thomasson a creer que la ignorancia o el error masivos acerca de la naturaleza de una clase artificial no son posibles por parte de, al menos, el creador o creadores de la respectiva clase. Para Thomasson, los miembros de una clase artificial tienen que satisfacer la concepción de los introductores de los respectivos términos. Dodd piensa que sostener (ii) es problemático por dos razones principales. La primera es que hay la posibilidad que el concepto asociado por los introductores a un término contenga inconsistencias. Y de la postura de Thomasson se sigue que si los introductores/creadores de un término tienen una concepción ontológica de sus referentes que implica inconsistencias, entonces no puede haber referentes. Y, como dice Dodd, éste es un resultado indeseable. El segundo motivo por el que rechazar (ii) es que, para Dodd, podría suceder que el referente no se ajustara a la concepción que los introductores asocian con el término, aunque tal concepción sea consistente. Así, por ejemplo, en el caso del término de clase natural ‘gato’, podríamos descubrir, como

⁷⁵ Estas dos tesis también estarían presentes en la teoría de Devitt y Sterelny.

propone Putnam, que los gatos son robots marcianos. En este caso, los introductores del término se habrían equivocado sustancialmente sobre la naturaleza de aquello a lo que pretendían referir con el respectivo término. Pero, según la teoría de Thomasson, en este caso, el término ‘gato’ no referiría, ya que sus introductores pretendieron que fuera un término de especie biológica⁷⁶. Y, para Dodd, esto parece falso.

En el caso de los términos de clase artificial, la posibilidad de la ignorancia o el error por parte de los creadores sobre los artefactos que han creado depende de la plausibilidad de los argumentos de este tipo que se puedan presentar (algo que analizaré en el siguiente capítulo), pero, ya de entrada, Thomasson, al aceptar (ii), ya rechaza esta posibilidad.

Dodd, entonces, rechaza (ii), pues cree que el concepto general que poseen los introductores sí podría estar equivocado, o ser inconsistente, acerca del referente. Pero, entonces, ¿cómo quiere defender (i)? Como rechaza (ii), el mecanismo por el que el concepto del introductor fija la referencia del término no puede ser que el concepto especifique unas condiciones que el referente del término tiene que cumplir. Y propone un mecanismo de fijación de la referencia que, de acuerdo con Dodd, ya había propuesto Evans en 1973. Según Dodd, Evans propone una teoría causal que adopta (i), pero rechaza (ii). Tal explicación estaría de acuerdo con que la referencia de un término está determinada por el contenido conceptual que los introductores asocian con él, pero no mediante el mecanismo de la satisfacción por parte del referente de lo que especifica ese contenido conceptual. Este contenido conceptual o cuerpo de información asociado al término fija la referencia del término, no en virtud de actuar como un *patrón* al que el referente debe ajustarse, sino por medio de su origen causal. En el caso de los términos singulares, el ítem determinado como el referente del término es la cosa que es la fuente (*source*) dominante⁷⁷ del cuerpo relevante de información, no el ítem entre las

⁷⁶ En el caso de los gatos-robots, Thomasson también podría decir que, con el descubrimiento, ‘gato’ puede sufrir un cambio de significado, en que se reemplaza la concepción antigua acerca de los gatos por la nueva (en que son concebidos como robots). Pero Dodd cree -pienso que con razón- que esta respuesta no es satisfactoria porque en la situación imaginada parece que descubrimos algo sobre los *gatos* (sobre aquellas cosas de las que pensábamos que eran animales).

⁷⁷ Como el hablante puede asociar diferentes informaciones provenientes de diferentes ítems o *diferentes cosas* presentes en la misma muestra (por ejemplo, puede asociar a un nombre propio la información que

diferentes cosas presentes que se ajusta a esa información. En el caso de los términos generales: un término general refiere a la clase de entidad que figura en la muestra que constituye la fuente (origen) causal dominante del cuerpo de información asociado al término.

Como el contenido conceptual asociado con un término de clase fija la referencia del término por el origen causal y no por el *ajuste*, se sigue que este contenido puede contener una inexactitud sustancial respecto a la naturaleza de la clase. Para Dodd, una propuesta como la de Evans tiene la ventaja, sobre la de Thomasson, de ser compatible con la posibilidad que el introductor de un término tenga un concepto inconsistente sobre aquello a lo que quiere referir con el término. Como que un término refiera a una clase de cosa está determinado por que esa clase de cosa sea la fuente causal dominante del cuerpo de información asociado a él, el término puede referir incluso si el contenido conceptual asociado es inconsistente. Y también es compatible con la posibilidad de descubrir que estamos sustancialmente equivocados o somos ignorantes sobre la naturaleza del referente de un término⁷⁸. Como el contenido conceptual asociado con el término determina la referencia *vía* su origen causal, y no actuando como un patrón o plantilla al que se debe ajustar un candidato a referente, es posible que un término pueda

su portador es una persona, por un lado, y, por otro, puede asociar al mismo nombre propio que el portador es un trozo de materia), el ítem que es el referente del término es el que es la fuente causal *dominante* de la información *relevante*. No hay que entender que la fuente *dominante* es necesariamente el ítem que es la causa de la *mayoría* de nuestra información, sino más bien que es aquel ítem que es la fuente de información que más nos importa a nosotros por el tipo de criaturas que somos. Así, en el caso de la introducción de un nombre propio para un ser humano, por el tipo de criaturas que somos y los intereses que tenemos, la fuente causal dominante de la información que asociamos es el ítem que es una persona que está ante nosotros, y no el ítem (ejemplificado por la misma cosa) que es un pedazo de materia.

⁷⁸ Dodd reconoce que hay un cierto grado de distorsión en la concepción asociada con un término de clase que, si se sobrepasa, hace que falle la referencia. Si toda la información relevante que asociamos con un término estuviera radicalmente equivocada (por ejemplo, si aquello que hemos bautizado con un nombre y que, desde lejos, nos parecía un perro, resultara ser realmente una sombra producida por una máquina, la intuición, según Dodd, es que la referencia ha fallado, no que el nombre propio refiera en realidad a la sombra). Pero mientras este umbral de distorsión no se sobrepase, el término refiere a la clase de entidad representada en la muestra que es el origen causal dominante de la concepción que los introductores asocian con el término.

tener un referente incluso si la concepción asociada resulta estar equivocada sustancialmente. En el caso del experimento mental de los gatos-robots, ‘gato’ selecciona esas cosas con bigotes y que maúllan, aunque hayamos estado equivocados sobre su naturaleza, pues esas cosas son el origen dominante del cuerpo de información (equivocado) que asociamos con el término. De acuerdo con Dodd, una teoría como la de Evans explica bien nuestras intuiciones, permitiendo que tanto ‘gato’ como ‘lápiz’ refieran a los gatos y a los lápices, respectivamente, en los experimentos mentales de los gatos-robots y de los lápices-organismos.

En resumen, si la propuesta de Dodd-Evans es coherente, se puede responder al *qua problem* de manera diferente a como lo hace una teoría híbrida de la referencia como la de Thomasson. Quizá se puede argumentar que la teoría de Evans no es una teoría causal *pura* como la de Putnam, pues hay un elemento descriptivo, un contenido conceptual, que juega un papel en la determinación de la referencia de los términos. Pero, dado que el papel del contenido conceptual consiste meramente en ser el efecto causal producido por el referente y puede estar sustancialmente equivocado (siempre que haya un ajuste de nivel *mínimo* al referente), considero que la propuesta de Dodd-Evans, independientemente de las diferentes interpretaciones que pueda haber de la teoría de Evans, es de tipo externista: incluso para los introductores de un término, no hay inmunidad a la ignorancia o el error sobre la naturaleza de lo designado. Recapitulando la crítica a Thomasson, creo que su propuesta presenta problemas en sus planteamientos metafísico, epistemológico y semántico. En particular, creo que su propuesta semántica no está bien motivada, pues tanto su teoría metafísica como el aducido *qua problem*, además de las objeciones que se les pueden plantear, no conducen inexorablemente, desde mi punto de vista, a una teoría semántica híbrida como la que ella propone. Pero, sin duda, pienso que el mejor argumento contra cualquier teoría descriptivista en algún grado para los términos de clase artificial es mostrar los pertinentes argumentos de ignorancia y error. En el resto de la tesis discuto la viabilidad de tales argumentos externistas.

Capítulo 4. Argumentos externistas de ignorancia y error contra el descriptivismo

En este capítulo, considero los posibles argumentos decisivos a favor de una semántica externista para los términos de clase artificial y que sirven como réplica o alternativa a las propuestas descriptivistas consideradas en el capítulo anterior. En la sección 4.1, reviso el argumento putnamiano de los lápices-organismos y considero las dificultades que presenta al compararlo con el argumento de la Tierra Gemela para el término ‘agua’. Por estas dificultades, tanto los autores externistas como los descriptivistas están de acuerdo en que el experimento mental de los lápices-organismos no es el tipo de argumento que necesita el externista sobre los términos de clase artificial, pues no es un argumento de ignorancia o error acerca de la propiedad *relevante* de los lápices, su función. En la siguiente sección, 4.2, expongo dos argumentos externistas de ignorancia y error, ofrecidos por Kornblith y Nelson respectivamente, sobre la función de un tipo de artefacto. Sin embargo, estos argumentos han recibido objeciones por parte de autores descriptivistas como Schwartz y Thomasson, tal como expongo en la sección 4.3. La principal objeción descriptivista es que los términos de clase artificial implicados en los argumentos externistas son términos sin referente y, por tanto, los argumentos externistas, no muestran lo que pretenden.

A continuación, en 4.4, presento mi análisis de estos argumentos externistas. Por un lado, presento mis críticas a las objeciones descriptivistas a esos argumentos, y, sobre todo, me centro en responder la objeción de la falta de referencia de los términos en cuestión. Por otro lado, analizo el alcance que, desde mi punto de vista, tienen los argumentos externistas anteriores contra el descriptivismo. Finalmente, tras la consideración de los límites que presentan los argumentos externistas de Kornblith y Nelson, en 4.5, contemplo la posibilidad de construir argumentos de ignorancia o error a favor del externismo que salven el inconveniente de los argumentos externistas anteriormente mencionados y ofrezco algunos de estos argumentos. Además, intento defender tales argumentos de posibles réplicas y, en particular, de la posible réplica de la falta de referencia de los términos para artefactos implicados. Por último, intento dar algunas pistas sobre qué tipo de argumento se presenta como el más plausible para una propuesta externista sobre los términos de clase artificial.

4.1. ¿Un argumento externista de *Tierra Gemela* con un término de clase artificial?

4.1.1. El argumento de los lápices-organismos y las críticas descriptivistas

Como expuse en el **capítulo 1**, considero que los argumentos decisivos para defender una postura externista sobre algún tipo de término, sea de clase artificial o no, son los argumentos basados en la ignorancia o el error sobre la naturaleza *relevante* de aquello a lo que aplicamos el término en cuestión. La idea general tras esta consideración es que los otros tipos de argumentos que se han presentado contra el descriptivismo, el argumento epistémico y el argumento modal de Kripke, son más fácilmente rebatibles o acomodables por el descriptivismo si éste adopta la forma de un descriptivismo meramente de fijación de la referencia según el cual el término y la descripción asociada no son expresiones sinónimas o con (necesariamente) el mismo significado. En cambio, los argumentos de ignorancia o error se aplican a cualquier forma de descriptivismo, pues pretenden mostrar que la descripción asociada al término no conduce al referente y, por tanto, no puede fijar la referencia del término, sea o no la descripción una expresión sinónima al término en cuestión. En el **capítulo 1**, presenté este tipo de argumentos basados en la ignorancia y el error de los hablantes acerca del referente, tanto para el caso de los nombres propios como para el caso de los términos de clase natural⁷⁹. En este último caso, consideré decisivo a favor del externismo semántico el argumento putnamiano de la Tierra Gemela para el término ‘agua’, que puede considerarse como un argumento de ignorancia sobre el agua (pues, en el experimento mental, la descripción asociada por Óscar a ‘agua’ no es una descripción equivocada sobre el agua, pero es insuficiente para seleccionar únicamente la referencia del término, ya que también selecciona el agua gemela).

Ahora la cuestión, en el debate con el descriptivismo, es si, desde el externismo, puede construirse un argumento equivalente al de la Tierra Gemela para los términos para artefactos. Tal como expuse en **2.4.1**, Putnam tiene la intención de presentar un argumento externista de este tipo para los términos de clase artificial. Como vimos, el argumento consiste en imaginar que descubrimos que todos los lápices de la Tierra son

⁷⁹ Además, argumenté que el tipo de posibilidad implicada en ambos tipos de argumentos es la misma: la posibilidad epistémica (y no necesariamente la posibilidad metafísica).

en realidad organismos y no artefactos⁸⁰. En tal situación, se mostraría que la descripción que asociamos a ‘lápiz’ según Putnam, ‘artefacto hecho para escribir’, no determinaría la referencia del término, pues los lápices no serían artefactos, sino organismos. El argumento, tal como lo expone Putnam, no es un argumento exactamente análogo al experimento mental de la Tierra Gemela sobre el agua⁸¹, es, más bien, análogo al experimento mental de los gatos robots⁸², pero se podría intentar reformular como un argumento *de Tierra Gemela* del siguiente modo: supongamos que hay una Tierra Gemela que es una réplica exacta de la Tierra salvo en que los lápices de la Tierra Gemela son organismos provistos de células, nervios, etc., mientras que en la Tierra los lápices tienen un origen artificial. Podemos suponer también que, en un momento previo al descubrimiento del microscopio, Óscar y Óscar Gemelo asocian a ‘lápiz’ la misma especificación en términos de la forma y la función de esos objetos. Pero, desde la perspectiva externista putnamiana, ‘lápiz’ dicho por Óscar y dicho por Óscar Gemelo tienen intuitivamente diferentes significados, pues el primer término refiere a los *artefactos* hechos para escribir, mientras el segundo refiere a unos *organismos* que sirven para escribir. La diferente naturaleza de los referentes (o dominios de aplicación) de los usos de los términos en la Tierra y en la Tierra Gemela, respectivamente, determina que los significados de ‘lápiz’ en la Tierra y en la Tierra Gemela sean diferentes, pese a que la descripción asociada con el término sea la misma. En cambio, para el descriptivista, el término ‘lápiz’, tanto dicho por Óscar como dicho por Óscar Gemelo refiere indistintamente a los lápices de la Tierra y a los lápices de la Tierra Gemela, pese a que unos son artefactos y los otros son organismos. Y esto sería contraintuitivo para Putnam.

Sin embargo, como dije en 2.4.1 y en 3.1.1, tal argumento a favor del externismo semántico para ‘lápiz’ no funcionaría para un descriptivista que, como Schwartz,

⁸⁰ Realmente el experimento mental parece originario de Rogers Albritton, que es a quien Putnam lo atribuye (1975, 161).

⁸¹ Véase la sección 1.5.2 del capítulo 1 donde se expone el argumento putnamiano de la Tierra Gemela con el término ‘agua’.

⁸² Véase la sección 1.5.1 para la exposición de dicho argumento.

considerara que los lápices (así como cualquier otro tipo de los considerados artefactos) no tienen necesariamente un origen artificial. Así que, según Schwartz, suponer como hace Putnam que el origen no artificial de los lápices de la Tierra Gemela hace que allí el término ‘lápiz’ tenga una extensión diferente a la del término ‘lápiz’ en la Tierra es un error. Lo único que determina que algo sea un lápiz es su forma y función, y no su origen artificial⁸³, y tanto ‘lápiz’ dicho por Óscar, como ‘lápiz’ dicho por Óscar Gemelo referirían o se aplicarían indistintamente a las mismas cosas: a todo objeto con la función de escribir. Entonces, para Schwartz, como los lápices no son necesariamente artefactos, el experimento mental de la Tierra Gemela en la línea de lo propuesto por Putnam para los artefactos no mostraría que ‘lápiz’ dicho por Óscar y ‘lápiz’ dicho por Óscar Gemelo tienen diferentes significados. La intuición de Schwartz sería, precisamente, que tienen la misma extensión y que la referencia de ambos términos está determinada por la descripción en términos únicamente de la forma y la función que asocian los hablantes de la Tierra y de la Tierra Gemela.

Pese a que yo creo que esta formulación del argumento de los lápices-organismos en términos *de Tierra Gemela* muestra que nuestras intuiciones se corresponden más con las intuiciones de Putnam que con las de Schwartz, de todos modos, el argumento putnamiano de los lápices-organismos no está libre de algunos problemas, que considero a continuación y que hacen que no resulte tan persuasivo como el argumento de la Tierra Gemela con el término ‘agua’.

4.1.2. Ulteriores dificultades del argumento de los lápices-organismos

En primer lugar, está la dificultad que acabo de mencionar, destacada por Schwartz, de si los lápices son necesariamente artefactos o no. Para Putnam, *si* son artefactos, entonces son necesariamente artefactos. La posibilidad de descubrir que no lo son sería meramente epistémica. En cambio, para Schwartz, los lápices no son necesariamente artefactos, podrían crecer en los árboles, por ejemplo. En este punto, la cuestión parece ser un asunto de intuiciones y éstas están divididas. Igualmente, en la versión del argumento que presenta Putnam sobre los lápices-organismos, según la cual -sin apelar a ninguna Tierra Gemela- podríamos descubrir que los lápices son organismos vivos,

⁸³ Véase el **capítulo 2**, la sección **2.4.1**, para ver la razón que ofrece Schwartz a favor de esta afirmación.

está la discusión sobre qué descripción asocian los hablantes a ‘lápiz’. La cuestión es si tal descripción hace mención o no al hecho que los lápices son artefactos. Para Putnam, la descripción asociada sí hace referencia al hecho que los lápices son artefactos. Esto no parece implausible, pero, por otro lado, *ser un artefacto* no es una propiedad *superficial* o *directamente observable* en un objeto como sí lo es, por ejemplo, la propiedad de *ser un líquido* que plausiblemente explicitamos en la descripción que asociamos a ‘agua’. En este sentido, si comparamos el experimento mental de los lápices-organismos con el experimento mental de la Tierra Gemela con ‘agua’, parece que no están en igualdad de condiciones, pues en el caso de los términos para artefactos, a diferencia de lo que ocurre con ‘agua’, no hay acuerdo entre el externista y algunos descriptivistas como Schwartz sobre las características asociadas con los términos de clase artificial.

Otra dificultad que se le puede atribuir al experimento mental de los lápices-organismos es una dificultad que algunos ven basándose en un ulterior experimento mental⁸⁴. Este otro tipo de experimento mental, considerado por Marconi (1987) contra el argumento análogo de los gatos-robots, puede plantearse, adaptado al caso de los lápices, como sigue: supongamos que, después de haber descubierto que los objetos a los que llamábamos ‘lápiz’ son en realidad organismos, descubrimos, tras examinarlos todos cuidadosamente, que hay un único *lápiz* que no es un organismo, sino un artefacto (tal como pensábamos al principio). O supongamos que en vez de descubrir que *todos* los lápices son organismos, descubrimos que todos los lápices son organismos excepto uno (o unos pocos) que, efectivamente, es un artefacto. En tal situación, la intuición de algunos podría ser que no diríamos que hemos descubierto que los lápices son organismos (o que los gatos son robots), sino que realmente sólo hay un *auténtico* lápiz (o un *auténtico* gato). Este ulterior experimento mental mostraría, desde posturas descriptivistas, que, para determinar la referencia de nuestros términos para artefactos, la descripción asociada por los hablantes tiene más peso que la verdadera naturaleza de los objetos *paradigmáticos*⁸⁵ a los que aplicamos el término. Así, este ulterior desarrollo

⁸⁴ Volveré sobre una variante de este experimento mental más adelante, en la sección 4.3.1, al considerar las críticas de Schwartz a ciertos argumentos externistas.

⁸⁵ Aquí, los lápices-organismos serían los objetos paradigmáticos porque serían los objetos claramente mayoritarios en la Tierra.

del experimento mental mostraría que la idea putnamiana que la referencia se determina por la similitud con los objetos paradigmáticos estaría equivocada.

Yo creo que, en estas situaciones, que podríamos llamar ‘situaciones de *todos-menos-uno*’, lo más importante no es saber qué diríamos o haríamos a partir del descubrimiento, sino determinar a qué referían los usos pasados del término implicado. Y el motivo es que lo que diríamos puede depender de muchos factores, entre ellos, factores de tipo psicológico o sociológico. Tomemos, por ejemplo, el conocido caso que presenta Mill con el nombre ‘Dartmouth’ de la ciudad inglesa. Seguramente, la ciudad se llamó así por estar en la desembocadura del río Dart. Ahora bien, ¿qué pasaría si el río Dart cambiara su curso de manera que Dartmouth ya no estuviera en su desembocadura? Sin duda, podríamos seguir usando ese nombre para la ciudad, aunque ya no estuviera en la desembocadura del río Dart. Pero también podríamos convenir que, debido a la importancia que tiene para una ciudad estar en la desembocadura de un río (porque digamos que, por ejemplo, puede suponer que varias veces al año haya una plaga de mosquitos especialmente agresivos), es más apropiado poner el nombre de ‘Dartmouth’ a la nueva ciudad que se encuentra ahora en la desembocadura y poner otro nombre a la antigua Dartmouth. En definitiva, podríamos decidir cambiar el nombre a Dartmouth o decidir no hacerlo; los factores implicados en tal decisión podrían ser de lo más variopinto y desconocidos con antelación. Así, igualmente, creo que las intuiciones sobre lo que haríamos en las mencionadas situaciones de *todos-menos-uno* no son tan importantes porque, según las circunstancias o los intereses que tengamos, podríamos hacer diferentes cosas (podríamos decidir aplicar ‘lápiz’ a todos esos ítems que ahora sabemos que son organismos y también al único ítem que es un artefacto, o podríamos aplicar ‘lápiz’ sólo a los organismos o sólo a los artefactos). Y, así como en el caso de Dartmouth creo que lo importante es si ‘Dartmouth’ refería o no a Dartmouth antes de que el río cambiara su desembocadura, en el caso de las situaciones descritas de *todos-menos-uno* creo que lo importante es dar cuenta de los usos y prácticas lingüísticos anteriores al descubrimiento. ¿A qué refería ‘lápiz’ (o ‘gato’) antes de descubrir que todos esos ítems a los que aplicábamos el término excepto uno son organismos (o robots)? A la hora de responder a esta pregunta creo que el externista dispone de una buena respuesta al decir que la mayoría de las veces conseguíamos referir a esas cosas a las que aplicábamos el término aunque estuviéramos equivocados sobre la naturaleza de la inmensa mayoría de esos ítems.

Por último, quiero considerar el que es, para muchos, el principal inconveniente que presenta el argumento putnamiano de los lápices-organismos: tanto para Schwartz, desde el descriptivismo, como para autores que comparten las intuiciones putnamianas sobre los términos para artefactos, el argumento de Putnam no es un argumento de ignorancia o error sobre la naturaleza *relevante* de los lápices. Y la naturaleza *relevante* de los lápices no es su origen, sino, según Schwartz, su forma y su función. Esto es algo que, en general, y con las matizaciones necesarias, acepta casi todo el mundo (también autores externistas, como Kornblith y Nelson, de los que hablaré enseguida). Así que, según las partes en conflicto, el argumento que necesita el externista es un argumento de ignorancia o error, por parte de todos, sobre *la función* de algún artefacto, que es la propiedad que hace que un artefacto sea el tipo de artefacto que es, y no sobre su origen artificial o natural. Este tipo de argumento ha sido, efectivamente, propuesto por autores que comparten las intuiciones externistas putnamianas. En la siguiente sección del presente capítulo, expongo dos argumentos externistas para términos de clase artificial, uno basado en la ignorancia y otro basado en el error acerca de cuál es la función de un artefacto. Estos argumentos pretenden ir contra la idea descriptivista de fondo consistente en pensar que la comunidad lingüística que usa un término para artefacto posee una descripción que singulariza la clase artificial en cuestión⁸⁶.

4.2. Dos argumentos externistas de ignorancia y error sobre la función de un artefacto

4.2.1. El argumento externista de Kornblith basado en la ignorancia

Kornblith (1980) propone el siguiente experimento mental como un argumento externista de ignorancia sobre la naturaleza *importante*, la función, de un tipo de artefacto: consideremos un antropólogo marciano que llega a la Tierra y se pone a investigar los artefactos humanos (sabiendo únicamente sobre esos objetos que se trata de artefactos hechos por humanos). El antropólogo marciano, que ignora la función de los pomos⁸⁷, señala un pomo que encuentra y dice ‘llamemos ‘glug’ a la clase de la que

⁸⁶ Ya posean la descripción algunos usuarios privilegiados (los creadores u otros expertos), o todos los hablantes.

⁸⁷ Kornblith pone el ejemplo con el término ‘doorstop’ (‘tope de puerta’ en castellano).

esto es un miembro'. Pese a la ignorancia del marciano sobre la función de los pomos, ha tenido éxito al usar 'glug' para referirse a los pomos, dice Kornblith. Y esta habilidad marciana para usar 'glug' no es necesariamente parásita o dependiente de la existencia de expertos humanos en pomos, según Kornblith. Podemos suponer que 'glug' es un término introducido solamente en el lenguaje marciano y que ningún ser humano asocia una descripción al término 'glug'. O, incluso, podemos suponer que los marcianos visitan la Tierra después de la desaparición de los humanos y, así, no habría ningún experto en pomos en la situación imaginada.

Este sencillo experimento mental de Kornblith intenta mostrar que es posible usar y referir con un término de clase artificial disponiendo solamente de una descripción insuficiente por parte de los hablantes para seleccionar el dominio de aplicación del término. Por eso es un argumento de ignorancia. En la situación expuesta, el único material descriptivo asociado por los marcianos a los pomos es que son objetos fabricados por los humanos (algo del todo insuficiente para determinar la referencia del término, pues *ser un artefacto* no singulariza una determinada clase artificial).

Además, Kornblith también destaca que su argumento muestra que se puede introducir por primera vez un término nuevo para artefacto sin que nadie, tampoco ningún experto, conozca una descripción asociada al término. En el ejemplo, los humanos sí podían asociar (o haber asociado) una descripción al término 'pomo', pero no al término 'glug'. Por tanto, no es el caso que, para cualquier término de clase artificial, tenga que haber expertos que asocien a él alguna descripción definida que sea verdadera de todos y sólo los miembros de la clase. Y si no hay expertos en el uso de un término para artefacto como 'glug', obviamente no puede haber deferencia. Hay que notar, además, que el descriptivista tampoco podría apelar, en virtud de la traducción entre 'pomo' y 'glug', a la deferencia, por parte de los usos marcianos de 'glug' a los usos expertos humanos de 'pomo', ya que la deferencia en la determinación de la referencia es un caso de cooperación social entre los hablantes. Y para que haya cooperación social tiene que haber una misma comunidad lingüística unida por nexos causales (una cadena histórica). Y es obvio que, en el ejemplo, los marcianos y los humanos pertenecen a comunidades lingüísticas diferentes. Sus usos de 'glug' y los usos humanos de 'pomo' no tienen ninguna conexión; en particular, no están conectados por ninguna cadena histórica de transmisión de la referencia.

Más bien, la situación en que se encuentra el introductor marciano del término ‘glug’ es análoga a *nuestra* situación cuando introducimos un término de clase natural. Como nosotros ante los objetos naturales, los marcianos encuentran unos objetos que no han creado y de los que hay que descubrir sus propiedades *relevantes* como la función, pues la función no es una propiedad obvia de los artefactos. Argumentos análogos a éste de Kornblith pueden construirse apelando al descubrimiento por parte de arqueólogos o antropólogos de artefactos desconocidos pertenecientes a otras épocas o culturas. En estas situaciones también puede darse la referencia con el término para artefacto sin que haya disponible una descripción que seleccione la clase artificial y sin que haya la posibilidad de deferir a ningún experto en el uso del término.

4.2.2. El argumento externista de Nelson basado en el error

El argumento de Nelson (1982) a favor del externismo semántico para los términos de artefacto se basa en el error de todos los hablantes sobre la propiedad *relevante* de un tipo de artefacto, su función, y la especificación asociada con el término de clase artificial correspondiente. De nuevo, el argumento es un experimento mental. Supongamos, dice Nelson, que los objetos llamados ‘lápices’ son realmente instrumentos repartidos por la Tierra por alienígenas malvados y que, por medio de los lápices, estos alienígenas espían y manipulan las actividades humanas. De hecho, los lápices son mecanismos a través de los cuales los alienígenas nos manipulan psicológicamente y desencadenan nuestras guerras y nos controlan a su antojo. Para camuflar la verdadera función de los lápices, los alienígenas nos hacen creer que son instrumentos para dibujar, escribir o hacer marcas sobre determinadas superficies, pero los lápices jamás han hecho esto. Es decir, los alienígenas nos someten a una ilusión continua acerca de cuál es la función de los lápices.

El punto importante de la situación imaginada es que, si es coherente, muestra que *todos* los humanos podríamos haber estado equivocados sobre la función de los lápices, atribuyéndoles una función cuando realmente tienen otra. Los humanos podríamos haber usado el término ‘lápiz’ teniendo una descripción asociada equivocada y sin que nadie en la Tierra conociera la verdadera función de los lápices, pero, pese a todo, habríamos usado el término con éxito y habríamos conseguido, según la intuición de Nelson, referir a esas cosas que llamábamos ‘lápices’. Argumentos similares a éste pueden construirse apelando a la manipulación de la función de los artefactos que

conocemos por parte de otros seres inteligentes no humanos. Sin embargo, Nelson admite que la plausibilidad de su ejemplo (así como de otros similares que se pueden construir) depende de la plausibilidad que queramos conceder a este tipo de experimentos mentales que apelan a una *ilusión* colectiva. Pero siempre puede apelarse, dice, a este mecanismo que permite construir tales argumentos:

After all, if Putnam can hold that pencils are organisms, despite the fact that large numbers of people would give their occupations as pencil-makers, we can tell our story as well. (Nelson 1982, 362)

Otra cosa, dice Nelson, sería explicar por qué la gente en la Tierra ha pensado que los lápices tienen la función de servir para escribir, y cómo se explica el papel de los supuestos fabricantes terrestres de lápices, etc. Pero si el experimento mental es coherente, un hablante descriptivista *à la* Schwartz no podría identificar los instrumentos alienígenas como lápices, dice Nelson, puesto que no responden a la especificación asociada al término que hace referencia a la forma y la función supuesta de los lápices. Sin embargo, un hablante externista diría que tales instrumentos sí son lápices; que tengan la función que creemos que tienen es un asunto epistémicamente corregible.

Y, de nuevo, en esta situación, los usuarios humanos de ‘lápiz’ no podrían apelar a los usos expertos del término correspondiente alienígena, porque las prácticas lingüísticas humanas estarían desconectadas de las prácticas lingüísticas alienígenas. Ninguna cadena histórica conectaría los usos humanos de ‘lápiz’ con los usos alienígenas del correspondiente término (o del mismo término, suponiendo que los alienígenas también empleasen el mismo término ‘lápiz’). Así que, para explicar cómo los usos humanos de ‘lápiz’ consiguen referir, no podemos apelar al hecho que hay una descripción, la que conocen los alienígenas, que determina la referencia del término. De manera análoga, si Dios existiese, podría conocer la descripción que selecciona la extensión de cada término de clase natural, pero los usos humanos de los términos de clase natural estarían desconectados de los usos divinos, y habría que explicar igualmente cómo es posible la referencia de nuestros términos para objetos naturales aunque tengamos una descripción equivocada o insuficiente para seleccionar la referencia. El problema entonces para el

descriptivista es explicar cómo ha sido posible la referencia humana a los lápices en el experimento mental de Nelson.

En resumen, en los dos argumentos expuestos de Kornblith y Nelson se pretende mostrar que es posible introducir y/o usar un término para artefacto aunque todos los usuarios del término en cuestión desconozcan una descripción verdadera y singularizadora sobre el tipo de artefacto correspondiente. En este punto, parecería que el descriptivista tiene que conceder que hay una similitud más grande de la que quizás sería de esperar entre los términos de género natural y los términos para artefactos.

4.3. Objeciones descriptivistas a los argumentos externistas de ignorancia y error y réplicas a esas objeciones

4.3.1. Objeciones de Schwartz

En esta sección considero las objeciones que Schwartz (1983) plantea especialmente al experimento mental de Nelson de los lápices alienígenas. Schwartz empieza concediendo que las situaciones contrafácticas que presentan tanto Kornblith como Nelson son plausibles. Pero, centrándose en el argumento de Nelson, dice que aceptar como coherente la situación que plantea su experimento mental no obliga a aceptar que la semántica externista sea la apropiada para 'lápiz'. La razón para afirmar esto se basa en las consecuencias de la ulterior situación imaginable de *todos-menos-uno* o *todos-menos-algunos* a la que aludí en 4.1.2: si tras descubrir que los lápices son instrumentos de espionaje y control marcianos (nunca usados realmente para escribir), encontramos *un* lápiz que sí tiene la función de servir para escribir, entonces el término 'lápiz' referiría a ese ejemplar y no a todos los otros objetos espías alienígenas.

En concreto, la variante del experimento mental que propone Schwartz consiste en imaginar que, después de haber hecho el descubrimiento que los lápices son instrumentos de control alienígena, producimos un contraejemplo a dicha afirmación. Esto es, conseguimos fabricar un instrumento para escribir hecho de madera con un núcleo de grafito, etc. En fin, hacemos un lápiz conforme a la idea previa que teníamos acerca de los lápices. O manufacturamos un puñado de lápices e ideamos un test para discriminar entre los lápices *reales* y los impostores de Marte. Entonces, la intuición de

Schwartz, en esta nueva situación de *todos-los-lápices-son-espías-marcianos-excepto-uno* (o *todos-menos-algunos*), es que todos pensaríamos y diríamos que ahora tenemos un lápiz *real*, no uno de esos instrumentos impostores de Marte. Así que, para Schwartz, aun aceptando que el experimento de Nelson es plausible, no muestra que la semántica de ‘lápiz’ sea una semántica de tipo externista debido a que, en la situación imaginada, podemos construir un contraejemplo que mostraría que ‘lápiz’ responde más bien a una semántica descriptivista.

El punto importante de este nuevo escenario con un contraejemplo es, según Schwartz, que

[...] with nominal kind terms anything that falls under the original specification will be correctly called by the term even if the paradigms and all previous examples turn out not to. In the case of nominal kind terms the original specification “dominates”, as I would like to put it. (Schwartz 1983, 477)

De acuerdo con Schwartz, en el caso de los términos de clase nominal, como son para él los términos para artefactos, la especificación original asociada con el término *domina*, y este resultado, dice, es contrario a lo que deberíamos esperar si ‘lápiz’ respondiera a una semántica externista como la defendida por Putnam.

Para ver esto con más claridad, Schwartz contrasta el comportamiento de los términos de clase nominal (en los que están incluidos, para él, los términos para artefactos) con el de los términos de clase natural. Comparemos, propone Schwartz, el posible descubrimiento que todos los lápices son instrumentos espías alienígenas (y no objetos que sirven para escribir, como creíamos) con el descubrimiento que las ballenas son mamíferos (y no peces, como pensábamos). Como se ha dicho, la primera situación, según Schwartz, admite un contraejemplo a la afirmación que los lápices son objetos de espionaje alienígena de modo que un solo objeto que tuviera la función de servir para escribir sería un lápiz (de hecho, sería el *auténtico* lápiz). En cambio, si, tras descubrir que las ballenas son mamíferos y no peces, encontramos un pez que parece en todo una ballena, no diríamos que debemos volver a creer de nuevo que las ballenas son peces. Tras descubrir que todas las ballenas son mamíferos, ningún pez, subraya Schwartz, por

mucho que se ajuste al estereotipo de las ballenas, sería una ballena. Los nuevos candidatos a ballena sólo serían miembros de la extensión del término si tienen la misma naturaleza que los paradigmas, y ningún pez sería parte de la extensión de ‘ballena’. Este contraste entre los casos muestra que los términos de clase artificial y los términos de clase natural no funcionan de manera análoga. Una característica importante de los términos de clase natural es, según Schwartz, que la naturaleza subyacente *domina* sobre las descripciones asociadas al determinar si un nuevo candidato es un miembro de la extensión del término. Y el fenómeno expuesto de *crear un contraejemplo*, dice, muestra que con los términos para artefactos la situación es justo al revés: las descripciones asociadas *dominan* sobre la naturaleza empírica de los paradigmas al determinar si nuevos candidatos son miembros de la clase. Un nuevo candidato que se ajuste a las descripciones asociadas previamente al descubrimiento sería correctamente llamado por el término, dice Schwartz.

La hipótesis de Schwartz sobre por qué las descripciones asociadas *dominan* en el caso de los términos de clase nominal (que incluyen los términos para artefactos), y no en el caso de los términos de clase natural, es que las descripciones asociadas a los términos de clase nominal no son teorías empíricas sobre la naturaleza de las clases, sino especificaciones analíticas que sirven como definiciones de los términos⁸⁸. Y, como tales especificaciones analíticas, son metafísicamente necesarias y cognoscibles *a priori*. Así, ‘Los lápices son objetos que sirven para escribir’ es, según Schwartz, una afirmación necesariamente verdadera y, en tanto que cognoscible *a priori*, epistémicamente incorregible. Pero esto último lleva a Schwartz a rectificar la admisión que ha hecho, de entrada, al conceder que el ejemplo de Nelson es coherente. Según reconoce después, hay que matizar o retirar la afirmación que podría resultar (epistemológicamente) que los lápices no fueran instrumentos para la escritura, pues no es cierto que pudiéramos estar en la situación epistémica de equivocarnos sobre la función de los lápices. Cuando admitimos que es posible descubrir que los lápices no sirven para escribir, simplemente estamos revelando que nuestro uso de los términos tiene ciertas limitaciones, pero, estrictamente hablando, no hay la posibilidad de tal descubrimiento. Según Schwartz, lo que sucede es que, al descubrir que los objetos a los que aplicábamos ‘lápiz’ no son lápices, sino un fraude, podríamos expresar este hecho

⁸⁸ Véase 2.2.1 y 2.2.2.

diciendo que los *lápices no son lápices*. Lo podríamos expresar de esta manera contradictoria porque

[...] there is a certain persistence effect with referring terms so that we can continue to use them to refer to things that we have habitually used them to refer to even if we discover that those things do not fit the descriptions they must (analytically) fit if the term is to correctly apply to them. (Schwartz 1983, 478)

Así, podríamos decir por un tiempo, después del descubrimiento, que los lápices no son instrumentos para escribir, pero no estaríamos reflexivamente de acuerdo con que esto es una verdad tal como está expresada, dice Schwartz. Simplemente lo expresamos así porque temporalmente podemos carecer de otros términos más apropiados. Pero Schwartz cree que rápidamente describiríamos la situación como una en la que, sencillamente, no había lápices. Diríamos, según él, que los que pensábamos que eran lápices no eran lápices al fin y al cabo. En cambio, destaca de nuevo Schwartz, esto no sucede con los términos de clase natural: al descubrir que las ballenas no son peces, no estamos inclinados en absoluto a decir que *las ballenas no son ballenas*, sino que las ballenas no son como pensábamos (esto es, no son peces). Esto, de acuerdo con Schwartz, muestra que, en el caso de los términos de clase natural, el criterio que seguimos para la correcta aplicación del término tiene más en cuenta la naturaleza subyacente de la clase natural, sobre la que podemos hacer descubrimientos, que la descripción que asociamos con el término. Pero, con los términos de clase artificial, para una apropiada aplicación del término, lo decisivo es la descripción (analítica) asociada a él.

En la sección **4.4.1** presentaré mi crítica a estas consideraciones de Schwartz; antes, en la sección que sigue, expongo las objeciones de Thomasson a los argumentos externistas de Kornblith y Nelson.

4.3.2. Objeciones de Thomasson

Respecto al experimento mental de Nelson de los lápices espías alienígenas, el caso que más se ha tenido en cuenta desde el descriptivismo, la respuesta de Thomasson coincide con la opinión última por la que también se inclina Schwartz: la situación descrita no es,

en sentido estricto, plausible, pues es una situación en la que realmente no habría lápices (esto es, cosas que sirven para escribir), sino sólo instrumentos espías marcianos. Por tanto, en el experimento mental, el término ‘lápiz’ proferido por los hablantes humanos no tiene un referente, no refiere a nada. Pero, además, Thomasson cree que su propuesta particular, en la que la naturaleza de los artefactos está determinada por las intenciones del creador y sólo éste tiene un privilegio epistémico acerca de esa naturaleza, no está afectada por este tipo de experimento *à la Nelson*. Hablando de este argumento dice:

[...] this does not tell against the view that someone (namely makers) has privileged knowledge regarding the nature of artifactual kinds; it only shows that we can be wrong about who the makers are. (Thomasson 2009a, 68)

De acuerdo con Thomasson, lo que muestra el argumento de Nelson es que podríamos equivocarnos acerca de quiénes son los creadores de un tipo de artefacto (podríamos pensar que nosotros hemos creado los lápices, cuando realmente no es así porque *esos objetos* han sido creados por otros seres con otros propósitos). Y esto no impugna la idea thomassoniana de que los *auténticos* creadores de un tipo de artefacto sí son inmunes a la ignorancia o el error masivos acerca de la naturaleza de los artefactos que han creado. Thomasson admite que el asunto de si un individuo particular *es* realmente el creador de un artefacto es un asunto siempre abierto a la duda. Y es este aspecto el que explota Nelson en su argumento. Del mismo modo, Thomasson cree también que siempre está abierto a la duda el hecho si un objeto particular es realmente un artefacto. Y esto es lo que, según ella, explota Putnam en su experimento mental de los lápices-organismos⁸⁹. Pero lo importante es que la idea putnamiana que todos podríamos estar equivocados sobre la naturaleza de los lápices, ya que podrían resultar ser organismos, tampoco dice nada contra el punto de vista thomassoniano que, allí donde hay miembros de una clase artificial, tiene que haber algún conocimiento de la naturaleza de la clase artificial. Para Thomasson, el argumento de Putnam de los lápices-organismos

⁸⁹ Pero si, realmente, ciertos objetos como los lápices, considerados artefactos, resultaran no ser artefactos, el término correspondiente para esos objetos no conseguiría referir, de acuerdo con Thomasson, pues los hablantes no pueden equivocarse radicalmente al asociar un concepto de primer nivel como *artefacto* a un término como ‘lápiz’. El mismo razonamiento se aplica al argumento de los gatos-robots y el término ‘gato’.

sólo muestra que podemos equivocarnos sobre qué objetos son artefactos.

En definitiva, la idea principal de Thomasson es que es necesario que, *si* algún individuo *es* el fabricante de una cierta clase de artefacto, entonces tenga (*qua* fabricante) garantizada una concepción ampliamente correcta de las características esenciales de los artefactos de la clase que crea. Y esta idea no queda afectada, como se ha dicho, por el argumento de Nelson. Y tampoco por el argumento de Kornblith del antropólogo marciano. En este último caso, Thomasson admite que es posible que todo el mundo de la comunidad lingüística del antropólogo marciano sea ignorante o se equivoque sobre la naturaleza del artefacto en cuestión, pese a que ‘glug’ parece referir a los artefactos de una cierta clase. Pero no es verdad, dice Thomasson, que los términos para artefactos puedan referir independientemente de *todas* las creencias y conceptos humanos sobre la naturaleza de la clase artificial. Tomando un ejemplo equivalente al que propone Kornblith, un término introducido por un antropólogo (humano) para un tipo de artefacto que descubre en una cultura remota tiene un referente sólo si hay una clase a la que referir, dice Thomasson. Y hay una tal clase artificial a la que referir sólo si hay gente que intencionadamente creó aquellos artefactos (con algún concepto sustantivo en mente sobre cuáles son las propiedades relevantes para ser un artefacto de esa clase). Así, concluye Thomasson, los conceptos humanos sobre la naturaleza de una clase artificial juegan un papel crucial en la determinación de la referencia de los términos para artefactos. Volviendo a la situación que plantea Kornblith, mientras los antropólogos marcianos pueden equivocarse sobre cuáles son las características esenciales de los pomos, no pueden equivocarse en que (asumiendo que sus muestras sí pertenezcan a una clase artificial común en la Tierra), sean cuales sean las características que los creadores terrestres considerarían como relevantes para ser un miembro de la clase, esas características que los creadores humanos considerarían relevantes son las características que cuentan para ser un miembro de la clase artificial.

4.4. Análisis de los argumentos externistas de Kornblith y Nelson

4.4.1. Réplica a las críticas de Schwartz y Thomasson

Presento a continuación mis críticas a las objeciones descriptivistas que plantean Schwartz y Thomasson. A mi juicio, la crítica más importante que Schwartz y

Thomasson plantean al argumento externista de Nelson es la misma: en el experimento mental nelsoniano, el término ‘lápiz’ no consigue referir porque no hay lápices. De todos modos, me detengo primero en las consideraciones que hace solamente Schwartz contra el argumento de los lápices-espías-alienígenas. En primer lugar, creo que la situación de *todos-menos-uno* (o *todos-menos-algunos*) que plantea Schwartz no es concluyente a favor del descriptivismo. Schwartz afirma que, en la situación de *todos-los-lápices-son-instrumentos-para-el-espionaje-marciano-excepto-uno* (o algunos), aplicaríamos correctamente ‘lápiz’ sólo al objeto u objetos que se ajustasen a nuestra especificación previa asociada con el término, y no a la mayoría de los objetos a los que llamábamos ‘lápices’ antes del descubrimiento. Pero, de nuevo, como apunté en 4.1.2, creo que desconocemos qué haría nuestra comunidad lingüística ante el descubrimiento de que los lápices, excepto uno o unos pocos, son instrumentos alienígenas que no tienen la función de escribir. Podría suceder que aplicáramos ‘lápiz’ a los instrumentos alienígenas y que usáramos otro término para los pocos objetos que escriben, o bien que aplicáramos ‘lápiz’ tanto a los instrumentos alienígenas como a los instrumentos que escriben. Creo que se trataría de un asunto de estipulación de nombres por parte de la comunidad, pero parece que el hecho no mostraría nada respecto a la semántica de ‘lápiz’ y a la determinación de su referencia antes del descubrimiento. Además, si el motivo último por el que, según Schwartz, la especificación asociada a un término de clase artificial *domina* sobre la naturaleza de los ejemplares paradigmáticos es que la especificación está ligada analíticamente al término, ya expuse, en la sección 2.2.2, los motivos por los que pienso que esto no es así y que los términos de clase artificial no son términos de clase nominal.

Igualmente, el contraste que establece Schwartz entre cómo usaríamos ‘lápiz’ y cómo usaríamos ‘ballena’, en las situaciones de *todos-menos-uno*, también está guiado por sus intuiciones acerca de nuestro comportamiento lingüístico en esas situaciones hipotéticas. Pero podría pasar que decidiéramos usar el término ‘ballena’ también para los peces que tienen aspecto de ballena. A veces, incluso en el caso de términos de clase natural, *podría* pasar que la descripción asociada *domine* sobre la naturaleza relevante

de las muestras paradigmáticas⁹⁰. Así que, desconociendo lo que decidiríamos tras descubrir las situaciones de *todos-menos-uno*, no parece buena idea contrastar el caso de los términos de clase natural con el caso de los términos de clase artificial. Pienso que, en la situación en que descubrimos que todos los lápices menos uno (o algunos) son instrumentos para el espionaje alienígena, también es verosímil que siguiéramos llamando ‘lápiz’ a aquellas cosas que tienen la misma naturaleza que los lápices paradigmáticos (esto es, a los objetos para el espionaje alienígena y que no sirven para escribir). Y, además, más importantemente, tanto en la situación descrita por Nelson como en la situación de *todos-menos-uno*, el externista puede explicar la referencia de ‘lápiz’ en los usos previos a los descubrimientos en las situaciones respectivas: ‘lápiz’ refería a aquello que tenía la misma naturaleza relevante (la función) que los lápices paradigmáticos, sólo que, tras el descubrimiento, sabemos que estábamos equivocados sobre cuál era esa naturaleza (esa función).

Ahora bien, la objeción descriptivista que me parece más importante, y en la que coinciden Schwartz y Thomasson, consiste, en última instancia, en negar que el ejemplo de Nelson muestre que podríamos equivocarnos sobre la función que tienen los lápices porque, sencillamente, en la situación que propone Nelson, no habría lápices. Y no habría lápices porque, según la teoría de Schwartz, los lápices tienen, por definición, la función de servir para escribir. Es decir, el término ‘lápiz’, por estar analíticamente ligado a la descripción ‘instrumento (*con tal y tal forma*) que sirve para escribir’, referiría o se aplicaría en cualquier situación contrafáctica únicamente a todo instrumento que sirva para escribir. Thomasson también parece asumir este punto de vista al comentar el experimento mental de Nelson:

But it does not here “turn out” to be the case that there are pencils although makers are massively mistaken regarding their nature. The mistaken individuals are not makers, nor are the objects pencils. These people mistakenly believe they are the makers of certain artifacts, when in fact the aliens are the true creators of these artifactual kinds. Since these people have made nothing at all, the humans intending to make pencils have not made

⁹⁰ Creo que el externista no niega que en ocasiones pueda *dominar* la descripción, lo que niega el externista es que, para que haya referencia, el referente tenga que satisfacer *siempre*, en todos los casos, lo expresado por la descripción asociada.

pencils, nor any other kind of artifact. Since the aliens have an entirely different concept in mind (that of a monitoring device), [...] no pencils have been created at all (but only monitoring devices). (Thomasson 2003, 603)

Al reconocer que, en la situación imaginada, no habría lápices, sino sólo instrumentos para el espionaje alienígena, Thomasson reconoce implícitamente que los lápices necesariamente tienen la función de servir para escribir (porque así lo imponen los creadores de lápices). Así que, tanto para Schwartz como para Thomasson, el término ‘lápiz’ en el experimento mental de Nelson no tiene referente. Parece, entonces, que tendríamos que considerarlo como otros términos generales sin referente, como ‘éter’, ‘flogisto’, ‘unicornio’ o ‘bruja’, por mencionar algunos de los casos más célebres⁹¹.

Respecto a esta cuestión, quiero hacer dos comentarios. El primero es que, describir la situación que plantea Nelson como una en la que no habría lápices, es incurrir en petición de principio, pues, tanto Schwartz como Thomasson, ya presuponen que un lápiz sólo puede ser algo que sirva para escribir y, por tanto, es imposible descubrir que tiene otra función. Estos autores presuponen que los lápices sólo pueden ser artefactos que sirven para escribir y que no podemos hacer un descubrimiento *sustancial* sobre la función de los lápices. Pero, claro, esto es, precisamente, lo que está en cuestión.

Aunque se admita, como postulan Schwartz y Thomasson, que tener la función de servir para escribir es una característica necesaria de los lápices, me parece que podríamos, sin embargo, estar en una posición epistémica como la propuesta por Nelson, en la que, por todo lo que sabíamos, los lápices podrían no haber servido para escribir. Es decir, que la

⁹¹ Sin embargo, puestos a establecer el paralelismo, creo que el caso del término ‘bruja’ sería el más parecido al caso del término ‘lápiz’ en el argumento de Nelson porque ambos son términos del lenguaje coloquial (a diferencia de ‘flogisto’ o ‘éter’ que parecen ser términos introducidos desde un cierto marco teórico científico para una sustancia hipotética –aparte de que ‘flogisto’ y ‘éter’ son términos incontables para supuestas sustancias, no términos contables) y ambos son términos que se han aplicado muchas veces (bien o mal aplicados) a cosas existentes en el espacio y el tiempo (a diferencia, probablemente, de lo que ocurre con ‘unicornio’). En lo que sigue, para establecer el supuesto paralelismo, me centraré principalmente en el caso de ‘bruja’.

posibilidad implicada aquí no tiene por qué ser de tipo metafísico, sino que puede ser de tipo epistémico⁹².

En segundo lugar, aunque se pueda pensar que, en la situación que plantea Nelson, no está del todo claro cuál sería, *a partir del momento del descubrimiento*, el referente de ‘lápiz’ –pues, como expuse antes, podrían darse diferentes hechos respecto a esos usos futuros-, considero que lo importante es dar cuenta de *los usos pasados* de ‘lápiz’, antes de descubrir que eran instrumentos de espionaje alienígena. Respecto a esos usos pasados en la situación nelsoniana, ¿cómo explican Schwartz y Thomasson todos los intercambios lingüísticos pasados entre los hablantes humanos al decir cosas como ‘Pásame el lápiz’, puesto que niegan que ‘lápiz’ consiguiera referir? Yo creo, sin embargo, que el caso contemplado de ‘lápiz’ es diferente al caso de términos comúnmente admitidos como términos sin referente, tales como ‘bruja’. Intuitivamente, ‘bruja’ no tienen referente porque con esos términos no referimos a algo existente con lo que hayamos establecido una relación causal, no hay ninguna muestra que podamos señalar de esa supuesta clase⁹³. Recordemos que, para Putnam⁹⁴, una entidad x es una entidad de una cierta clase X si y sólo si es similar -en ciertos aspectos *relevantes* para nosotros- a las muestras paradigmáticas de esa clase (a las muestras a las que mayoritariamente aplicamos el término en el mundo actual). En el caso de las brujas, no tenemos muestras paradigmáticas de esa clase. En cambio, creo que, intuitivamente, en la situación nelsoniana, sí hay lápices: esas entidades que podemos señalar y que son similares a los lápices paradigmáticos (al grueso de los objetos a los que llamábamos ‘lápiz’ en la Tierra y con los que habíamos establecido una relación causal). Sin embargo, estas intuiciones deben justificarse de alguna manera, si no, parece que simplemente se presupone que las brujas no existen (algo que es generalmente

⁹² Esto es lo que pasaba también tanto con los argumentos de ignorancia y error que contenían términos de clase natural como con los que contenían nombres propios (véase **1.5.1**).

⁹³ Se supone que el término ‘bruja’, al menos en un sentido en que no es un mero término en el campo de la ficción o el mito, se ha aplicado muchas veces a entidades concretas, situadas en el espacio y el tiempo y con poderes causales.

⁹⁴ Véase sección **1.5.3**.

aceptado) y que los lápices en el caso nelsoniano, en cambio, sí existen (algo que podría no ser aceptable para muchos). Además, parece que la situación aludida de ‘lápiz’ es análoga a la situación de ‘bruja’: en ambas situaciones, se da la intención de referir a una cierta propiedad (tener la función de servir para escribir / tener ciertos poderes sobrenaturales) que supuestamente tienen *esas cosas* a las que aplicamos el término, y también se da, en ambos casos, la equivocación al atribuir esa cierta propiedad a *esas cosas* (pues, realmente, no tienen esa propiedad a la que pretendemos referir respectivamente). Entonces, ¿por qué deberíamos admitir que hay lápices y que el término ‘lápiz’ refiere a *esas cosas* cuando no admitimos lo propio con ‘bruja’? Respecto a las brujas, el motivo por el que parece que decimos que no existen (y, por consiguiente, el término ‘bruja’ no tiene referente) es porque la supuesta propiedad que se atribuía a las personas a las que se aplicaba el término realmente no era poseída por esas personas, con lo que, además, el descriptivista parecería tener razón en que el referente de un término debe tener la propiedad especificada en la descripción asociada al término en cuestión. ¿Qué diferencia hay, entonces, entre el término ‘bruja’ y el término ‘lápiz’ en el argumento nelsoniano?

Antes de exponer la diferencia que creo que hay entre los dos casos, quiero aclarar que me parece que no es admisible el razonamiento tras la idea que ni ‘lápiz’ (en el caso de Nelson) ni ‘bruja’ tienen referente. Considero que decir que ‘lápiz’ y ‘bruja’ no refieren *solamente* porque las cosas a las que mayoritariamente se aplica el término no tienen la propiedad relevante que los hablantes les atribuían, presupone el descriptivismo y excluye que podamos referir al usar un término cuando estamos sustancialmente equivocados sobre la naturaleza de una clase. Pero, por todo lo expuesto hasta ahora, me parece razonable suponer que sí podríamos equivocarnos sobre la naturaleza de una clase (y especialmente cierto me parece en el caso de una clase natural). Así que, como tengo la intuición que ‘lápiz’ sí refiere en el argumento de Nelson y, en cambio, tengo la intuición que ‘bruja’ es un término sin referente, expongo a continuación lo que me parece que es una diferencia explicativa entre estos dos términos.

Consideremos el caso de ‘bruja’ y supongamos, como es plausible, que hubo una época en la que se creía que existían las brujas, mujeres con ciertos poderes sobrenaturales en virtud de haber hecho un pacto con el diablo. Si decimos que el término no refiere porque nunca hubo mujeres con esos poderes sobrenaturales, esto es, porque no hubo

(ni hay) nada que se ajustara a la especificación asociada con el término ‘bruja’, entonces parecería que el descriptivista tiene razón y que la referencia no es posible si cometemos un error *sustancial* sobre la propiedad que se considera relevante para la referencia del término. Sin embargo, centrándonos en el caso de las brujas, yo creo que la clave está en que, cuando la ciencia descartó, definitivamente, que pudiera existir alguna propiedad como la que se suponía a las mujeres llamadas ‘brujas’, ninguna otra propiedad *relevante para nosotros* era compartida por esas mujeres como para seguir aplicándoles el término⁹⁵. Si aceptamos la idea externista de Putnam según la cual la referencia de un término de clase está determinada por una propiedad *importante* de las cosas de esa clase, y también queremos admitir, como creo que es razonable, el postulado externista según el cual existe la posibilidad de que nos equivoquemos acerca de la propiedad *relevante* que atribuimos a ciertas cosas a las que aplicamos un término general, entonces se sigue que la referencia de un término general podría estar determinada por una propiedad desconocida por nosotros. Y esta propiedad desconocida de entrada podría pasar a ser *importante para nosotros* una vez descubierta. En el caso del término ‘bruja’, y puesto que la propiedad atribuida a los individuos a los que se aplicaba el término no existe, ¿qué otra propiedad *relevante* comparten esos individuos? Ninguna que sepamos. No sólo las personas a las que se llamaba ‘brujas’ no tienen la propiedad que se suponía, sino que no tienen en común ninguna otra propiedad relevante que nos haga pensar que nos equivocamos acerca de su naturaleza y que su naturaleza, realmente, consistía en otra propiedad, igualmente *importante*. En definitiva, nunca hubo brujas porque nunca hubo realmente un grupo de mujeres que compartieran *alguna* propiedad *relevante* para nuestros intereses clasificatorios (fuera la propiedad de tener ciertos poderes sobrenaturales u otra *equivalente*). Así que considero que el término ‘bruja’ no refiere porque se dan dos condiciones: no sólo porque no haya mujeres que tengan ciertos poderes sobrenaturales, sino también porque esas mujeres a las que aplicábamos el término no compartían ninguna propiedad *relevante* para nosotros (ni la supuesta erróneamente ni ninguna otra).

⁹⁵ La propiedad trivial que tenían en común de ser llamadas ‘brujas’ no es, típicamente, una propiedad *relevante* para nosotros como para seguir aplicando un término a los individuos que la comparten.

En cambio, en el experimento de Nelson, al descubrir que los lápices no tienen la función que pensábamos, se podría decir que otra propiedad *relevante equivalente*, la propiedad de tener la función de servir como mecanismos espías para los alienígenas, pasó a ser, razonablemente, la propiedad *relevante* (antes desconocida) decisiva para la pertenencia a la clase y que determinó la referencia del término desde un principio. Creo que esta diferencia entre los casos de ‘lápiz’ (en el argumento de Nelson) y los casos de términos generales sin referente, como ‘bruja’, podría dar cuenta de la intuición que ‘lápiz’ en el argumento externista de Nelson es un término que refiere (y refiere a esas cosas de las que creíamos que tenían la función de servir para escribir)⁹⁶, mientras que ‘bruja’ y otros términos por el estilo, no.

Igualmente, en el experimento mental de Kornblith, el término para artefacto ‘glug’ parece referir a los pomos. Lo que pasa en este caso es que la propiedad *relevante* que determina la pertenencia a la clase y la extensión del término está por descubrir por parte de los marcianos. Ellos simplemente postulan que hay una propiedad *relevante* de un cierto tipo (funcional), pero no saben cuál es. Así que creo que la referencia de ‘glug’ a los pomos es, de entrada, posible, a menos que ulteriores investigaciones marcianas determinaran que aquellos objetos a los que han aplicado el término ‘glug’ realmente no tenían ninguna propiedad relevante en común (ni la función ni ninguna otra propiedad *equivalente* que ocupara su lugar en importancia para los intereses clasificatorios marcianos). Sin embargo, el argumento de Kornblith ha recibido un tratamiento diferente al de Nelson por parte de los autores descriptivistas. Schwartz no se posiciona explícitamente respecto al argumento de Kornblith como sí lo hace frente al de Nelson, pero, sin embargo, creo que, desde una postura como la de Schwartz, el motivo por el que ‘glug’ no tendría referente es que los hablantes marcianos no asocian material descriptivo suficiente para llegar hasta el referente. Así, el motivo por el que ‘glug’ no referiría sería diferente al de por qué ‘lápiz’ en el argumento de Nelson no refiere. En el caso de ‘glug’, si la insuficiencia de la descripción asociada por los marcianos permite o no la referencia del término es más un asunto de intuiciones de un tipo o de otro, externistas o descriptivistas.

⁹⁶ Hay que notar que la nueva descripción que podríamos asociar a ‘lápiz’ en el experimento mental de Nelson, ‘instrumentos espías alienígenas’, no tendría por qué pasar a ser la descripción que determinara la referencia de ‘lápiz’ a partir del momento del descubrimiento, pues también hay la posibilidad de construir argumentos de ignorancia o error sobre esa nueva concepción de los lápices.

Thomasson, en cambio, no excluye que los marcianos logren referir con ‘glug’, pese a que nadie en la comunidad lingüística marciana conoce las características distintivas de los pomos. Pero, según parece sugerir, la referencia a los pomos con ‘glug’ sería posible en virtud de un conocimiento de tipo *general* que tienen los marcianos sobre los pomos: que son artefactos humanos y que, sean cuales sean sus características distintivas, éstas están delimitadas por las intenciones humanas de los creadores de los pomos. Sin embargo, como, en este punto, Thomasson no puede apelar a la deferencia lingüística a los expertos humanos, no está claro cómo puede explicar la referencia de ‘glug’ en los usos marcianos atribuyéndoles simplemente la intención de referir a aquello que sus creadores consideraron como las características relevantes para ser un artefacto de ese tipo. Sin embargo, como explico a continuación, en la sección siguiente, esto no tiene por qué suponer un problema mayor para Thomasson o para otras propuestas descriptivistas similares.

4.4.2. El alcance de los argumentos externistas de Kornblith y Nelson en la disputa con el descriptivismo

Desde mi punto de vista, los argumentos externistas presentados explotan la idea que es posible introducir y/o usar un término para artefacto en una comunidad lingüística diferente a la comunidad lingüística de invención de la función del artefacto en cuestión (comunidades entre las que, además, no puede establecerse una cooperación lingüística). Esta posibilidad, evidente incluso en el caso de un antropólogo humano que estudia un artefacto desconocido de otra cultura humana, obliga al descriptivista a dar una respuesta. Y creo que puede optar a dos líneas de razonamiento. En primer lugar, como expuse antes, el descriptivista puede apelar a la falta de referencia de los términos para artefacto implicados en los ejemplos. Puede decir, como Schwartz en los dos casos y Thomasson respecto al ejemplo de Nelson, que ni el antropólogo marciano consigue referir con ‘glug’ ni los humanos consiguen referir con ‘lápiz’. Ya he explicado, sin embargo, por qué creo que se puede defender que los términos implicados en los argumentos mencionados sí consiguen referir y son diferentes de los términos generales considerados tradicionalmente sin referente.

Otra posible línea de respuesta por parte del descriptivista es aceptar, como Thomasson parece admitir en el argumento de Kornblith, que sí hay referencia en los argumentos externistas. Sin embargo, esto no implicaría renunciar al descriptivismo (al menos a una

propuesta como la de Thomasson). Simplemente se trataría de conceder que ‘lápiz’ y ‘glug’ refieren en los experimentos mentales citados, pero matizando que esos casos son casos muy extraños, marginales, de introducción y uso de un término para artefacto. En la situación estándar de introducción y uso de un término de clase artificial, en la que la comunidad lingüística que introduce y usa el término es la misma que ha creado el tipo de artefacto, la tesis descriptivista se sostendría perfectamente. Y la tesis descriptivista dice que la referencia de un término para artefacto está necesariamente determinada (al menos en parte) por una descripción cognitivamente accesible a los usuarios del término o a alguien a quien los usuarios del término pueden deferir al usar el término. Como esta tesis se cumpliría, de acuerdo con el descriptivista, en los casos habituales de uso de los términos de clase artificial, lo único que tendría que matizar el descriptivista es que es una tesis relativa a la comunidad lingüística que crea un tipo de artefacto y usa el término correspondiente. En comunidades lingüísticas diferentes a la comunidad de invención de un tipo de artefacto (y sin posibilidad de establecer cooperación social con la comunidad inventora), siempre es posible usar y/o introducir el respectivo término para el artefacto con ignorancia o error por parte de todo el mundo en esa comunidad sobre la función del artefacto. Así que, aunque se conceda que la referencia con un término de clase artificial es posible en situaciones como las que plantean Kornblith y Nelson, el descriptivismo podría seguir presentándose como plausible y sin argumentos en contra respecto a las situaciones no marginales de uso de los términos de clase artificial. No creo que tener que restringir la tesis descriptivista de determinación de la referencia de los términos de clase artificial a situaciones en que la comunidad lingüística en cuestión es la creadora del tipo de artefacto menoscabe seriamente una teoría como la de Thomasson, incluso aun admitiendo que hay situaciones especiales en que la referencia se produce externísticamente.

De hecho, Thomasson, refiriéndose al argumento de Nelson pero igualmente ampliable al caso de Kornblith, dice, creo que con razón, que este tipo de argumento externista no afecta al núcleo de su teoría intencional según la cual los creadores de un tipo de artefacto son inmunes a la ignorancia y el error (masivos) sobre los artefactos que crean. Efectivamente, el argumento de Kornblith mostraría, a lo sumo, que los marcianos podrían ignorar las características distintivas de un tipo de artefacto que ha creado otra civilización. Igualmente, el argumento de Nelson mostraría que los humanos podríamos estar equivocados sobre un artefacto que *no* hemos creado (y que habíamos creído que

hemos creado). Lo importante, para salvaguardar la propuesta principal de Thomasson, es que los humanos, en el caso de Kornblith, y los alienígenas, en el caso de Nelson -los verdaderos creadores de los artefactos en cuestión- no pueden ignorar ni equivocarse sobre el tipo de instrumento que han creado. Los humanos (al menos los creadores, según Thomasson) sí conocen la descripción que selecciona la referencia de ‘pomo’, y, en el segundo caso, los alienígenas (al menos los creadores) sí conocen la descripción singularizadora asociada a su término para los lápices. Naturalmente, en las comunidades creadoras de las funciones de los pomos y de los lápices sí es posible la deferencia lingüística a los expertos en pomos y a los expertos en lápices (esto es, a los creadores, de acuerdo con la propuesta de Thomasson), y la referencia de estos términos se determinaría mediante la descripción singularizadora que poseen esos expertos. Así que la tesis thomassoniana que los creadores de un tipo de artefacto no pueden ignorar o equivocarse sustancialmente sobre las propiedades importantes de los artefactos que han creado quedaría intacta. Y, del mismo modo, cualquier teoría que defendiera que la comunidad que ha creado un tipo de artefacto no puede ignorar o estar sustancialmente equivocada sobre los artefactos que ha creado y usado (sea por el papel que juegan los creadores, o bien otros usuarios) tampoco se vería afectada en esta idea por los argumentos externistas expuestos.

Así, a fin de cuentas, concediendo que los argumentos externistas revisados muestran que es posible introducir y/o usar un término para artefacto con ignorancia o error, por parte de toda una comunidad lingüística, acerca de la naturaleza de un tipo de artefacto, estos argumentos sólo socavarían la tesis descriptivista que defiende que no es posible, *en ningún caso*, que la referencia del término se determine por algo no accesible cognitivamente. El alcance de los dos argumentos externistas se circunscribe únicamente a las comunidades lingüísticas desconectadas del uso del término de la comunidad introductora del artefacto. En el caso del argumento de Kornblith, la ignorancia se circunscribe a la comunidad marciana, y en el caso del argumento de Nelson, el error se circunscribe a la comunidad lingüística humana. Pero parecería que, al menos, la comunidad lingüística creadora del artefacto sí ha de tener acceso cognitivo a la descripción que singulariza la clase artificial.

4.5. La posibilidad de construir argumentos de ignorancia y error para toda comunidad lingüística

4.5.1. Argumentos universales de ignorancia y error

Como he expuesto, la idea descriptivista de fondo que permanece intacta por los argumentos externistas de Kornblith y Nelson es que siempre que hay un artefacto y su correspondiente término, parece que *tiene que haber* alguien (sea el creador o no del artefacto), plausiblemente perteneciente a la comunidad lingüística que ha creado el artefacto en cuestión, que conoce la descripción que selecciona ese tipo de artefactos. Así, todo término para artefacto tiene (o ha tenido) una descripción asociada singularizadora de la clase en alguna comunidad lingüística -presumiblemente la comunidad introductora de tal artefacto- y, al menos en esa comunidad, la referencia del término para artefacto queda determinada por esa descripción que alguien conoce. El descriptivista puede acogerse a que, independientemente de lo que pase en otras comunidades lingüísticas, al menos en la comunidad lingüística en la que se ha creado el artefacto no es posible la ignorancia y el error por parte de todo el mundo.

En este punto, hay que destacar la diferencia principal entre el argumento de Putnam de la Tierra Gemela sobre el agua y los argumentos externistas que hemos visto. En el argumento de Putnam se muestra que *todos*, tanto los hablantes de la comunidad lingüística de la Tierra como los hablantes de la comunidad lingüística de la Tierra Gemela, podríamos habernos equivocado al aplicar ‘agua’ a una determinada sustancia, pues el significado de ‘agua’ no viene dado por ninguna descripción que cualquier hablante pueda asociar, sino por la naturaleza de la propia sustancia agua. En cambio, el argumento de Nelson no excluye que al menos algunos hablantes de alguna comunidad lingüística sean inmunes a la ignorancia o el error respecto a la descripción que asocian al término de clase artificial. Esto es, el experimento mental de los lápices alienígenas no es un argumento *universal* de error, no muestra que los alienígenas (los creadores de la verdadera función de los lápices) podrían estar equivocados en la descripción que asocian a su término para los lápices. Igualmente, el argumento de Nelson no muestra que *cualquier* hablante de *cualquier* comunidad podría ser ignorante acerca de la naturaleza de los pomos, sino sólo que los marcianos poseen una descripción insuficiente para seleccionar la extensión del término correspondiente.

Me parece, entonces, que el poder de persuasión del externismo semántico para los términos de clase artificial pasa por ofrecer argumentos de ignorancia o error de tipo *universal*; esto es, argumentos de ignorancia o error sobre la función de un tipo de artefacto que afecten a toda comunidad lingüística (incluida, claro, la comunidad creadora del tipo de artefacto en cuestión). ¿Es posible construir este tipo de argumento con un alcance *universal*⁹⁷ como el de la Tierra Gemela pero con los lápices o con cualquier otro artefacto en vez de con el agua? La dificultad en construir un argumento de *Tierra Gemela* estriba en que, si la función del artefacto ha de jugar un papel análogo al de la naturaleza subyacente de una sustancia natural, entonces la situación requerida sería la siguiente: en la Tierra Gemela hay un tipo de artefacto igual en todo a un tipo de artefacto de la Tierra salvo en la propiedad que consideramos *relevante* para los artefactos (la función) y, por tanto, tendrían que ser dos tipos diferentes de artefactos. Pero imaginar tal situación es problemático. Podríamos suponer que los dos artefactos son indistinguibles perceptiblemente e idénticos en todo; lo único que sucede es que en la Tierra el artefacto desempeña una función, y, en la Tierra Gemela, desempeña otra función. Por ejemplo, podría ser el caso que en la Tierra Gemela unos instrumentos indistinguibles de los destornilladores tuvieran la función de servir como objeto decorativo o como objeto para hacer palanca (sin nunca ser usados para atornillar o destornillar), mientras en la Tierra tuvieran la función de atornillar y destornillar. Pero, en este caso, la intuición que tenemos es que se trataría del mismo artefacto pero que realiza dos funciones diferentes⁹⁸. Si viajáramos a la Tierra Gemela, no sólo diríamos que usan los destornilladores para hacer palanca, sino que podríamos usar esos

⁹⁷ El adjetivo '*universal*' debe entenderse de una manera razonablemente limitada: afectaría a todos aquellos seres inteligentes pero con limitaciones cognitivas o con un sistema cognitivo similar al humano. Es obvio que no afectaría a seres omniscientes o con una facultad cognitiva especial para acceder de manera inmediata a la verdadera función de los artefactos. Esta es una limitación plausible, pues en la crítica al argumento de Nelson, el descriptivista presupone que los alienígenas tienen un sistema cognitivo similar al nuestro, en el sentido de que podrían conocer la función de los artefactos que han creado en virtud de ser la comunidad creadora del artefacto, no en virtud de ninguna otra facultad especial.

⁹⁸ Devitt (2005) no comparte esta intuición y, para él, desempeñar una función distinta implica ser un tipo de artefacto distinto. Así, un objeto indistinguible de una mesa pegado por sus patas a un techo, para adornar, deja de ser una mesa y pasa a ser un adorno. Yo creo que este punto de vista no es intuitivo: una mesa con ese uso nos seguiría pareciendo una mesa.

artefactos como destornilladores. Más bien, sería un caso en que un artefacto es susceptible de ser usado con dos funciones diferentes⁹⁹.

Hay que notar, además, que otra dificultad de plantear este tipo de experimento está en que, si los dos artefactos han de tener funciones diferentes en la Tierra y en la Tierra Gemela, respectivamente, no está claro que Óscar y Óscar Gemelo puedan asociar con sus respectivos términos descripciones iguales (salvo que la mención a la función sea excluida de esa descripción, cosa poco probable). Se violaría, de esta manera, una de las condiciones del experimento mental (Óscar y Óscar Gemelo deben asociar la misma información al término en cuestión).

Quizá otra vía para intentar presentar un argumento equivalente al de la Tierra Gemela con 'agua' es considerar que en la Tierra y en la Tierra Gemela podría haber dos artefactos indistinguibles superficialmente, en el nivel perceptible, pero que tienen alguna diferencia no obvia (por ejemplo, en alguna pieza muy pequeña y escondida) que hace que tengan funciones diferentes. Supongamos, por ejemplo, que en la Tierra Gemela hay un artefacto indistinguible superficialmente de los microondas terrestres, pero que, debido a un pequeño componente que tiene, en vez de calentar los alimentos que se introducen en él, los refrigera. En este caso, estaríamos ante dos artefactos intuitivamente diferentes (pues tienen funciones diferentes y uno no puede realizar la función del otro) e indistinguibles a simple vista. Pero, de nuevo, las descripciones asociadas en la Tierra y en la Tierra Gemela, respectivamente, a cada uno de los artefactos, serían descripciones diferentes (pues, plausiblemente, harían referencia a funciones diferentes). Así que, no estaríamos ante un argumento como el putnamiano, pues, aunque estemos ante artefactos diferentes, las descripciones en cada planeta serían, también, diferentes (confirmando que se trata de artefactos de diferente tipo).

La situación que exige un escenario *de Tierra Gemela* es una situación en que tengamos dos artefactos indistinguibles, excepto en la función, en la Tierra y en la Tierra Gemela respectivamente, a los que, tanto los terráqueos como los terráqueos gemelos, asocien las mismas propiedades *relevantes* (la misma función), pero que realmente tengan

⁹⁹ Véase la discusión acerca de la función posible en el **capítulo 3**, sección **3.4.1**.

funciones diferentes y sean, por ello, artefactos intuitivamente de un tipo diferente. Y creo que, por las razones consideradas, es muy difícil (si es que es posible) formularlo en esos términos. Pero el recurso putnamiano de la Tierra Gemela es sólo un tipo de argumento de ignorancia o error (en los casos conocidos del agua y de los lápices-organismos). Quizá pueda ser más fácil formular otro tipo de argumentos de ignorancia o error con términos de clase artificial tomando como punto de partida los argumentos de Kornblith y Nelson, pero intentando evitar el inconveniente que presentaban éstos; esto es, intentando evitar que tengan sólo un alcance *local*.

Empiezo por un argumento parecido al que plantea Nelson, basado en el error –debido a una ilusión colectiva- de una comunidad lingüística sobre la función de un artefacto, pero evitando el inconveniente de que haya otra comunidad lingüística que sí podría tener acceso cognitivo a la verdadera función del artefacto. Tomando el mismo ejemplo de los lápices, supongamos que descubrimos que estos artefactos nunca han tenido la función de servir para escribir ni marcar trazos sobre ninguna superficie. Lo que ocurre es que los lápices tienen la propiedad -en virtud, digamos, de una propiedad que hemos descubierto que tiene el grafito sobre nosotros- de producir en nuestro sistema óptico la ilusión que los movimientos que hacemos con los lápices se corresponden con trazos marcados en las superficies, pero, realmente, en las superficies contra las que apretamos los lápices, no aparece ni nunca ha aparecido ningún trazo. Todo ha sido fruto de una ilusión. En tal situación, parece que los usos del término ‘lápiz’ han referido a esas cosas a las que mayoritariamente¹⁰⁰ aplicábamos el término ‘lápiz’, y de las que ahora sabemos que nunca han tenido la función que creíamos que tenían. *Todos* habíamos estado equivocados sobre la naturaleza *relevante*, la función, de los lápices.

Este recurso de la ilusión óptica puede ser apelado para construir otros argumentos similares de error *universal*. Otro experimento mental de error, que me parece más convincente, sería el siguiente: supongamos que descubrimos que el detergente que usamos para lavar la ropa en realidad no consigue desprender las manchas de las superficies de los tejidos. Realmente lo que sucede es que los componentes que usamos para hacer los detergentes -digamos, ciertos componentes tensioactivos- tienen la propiedad de crear en nosotros la ilusión óptica de que las manchas desaparecen cuando

¹⁰⁰ Quizá no a todas porque podría haber habido aplicaciones incorrectas de ‘lápiz’.

no es así. Entonces, en tal situación, la descripción que asociábamos a ‘detergente’, algo así como ‘sustancia que sirve para lavar las manchas’, no sería una descripción verdadera de los artefactos a los que aplicamos ‘detergente’. Habríamos estado *todos* equivocados, también los fabricantes y los más expertos en detergentes, al pensar que esos ejemplares a los que llamamos ‘detergente’ sirven para limpiar. Pero, aún así, parece que con nuestros usos de ‘detergente’ hemos conseguido referir a esas muestras que realmente nos producen determinadas ilusiones ópticas. Aquí quiero notar que quizá la posibilidad implicada en este tipo de argumentos es meramente epistémica y no metafísica, pues quizá es imposible metafísicamente que algún componente del jabón detergente cause el efecto mencionado. Pero, de nuevo, si la situación es posible como situación epistémica, creo que es suficiente para su validez como argumento de error¹⁰¹.

Por otro lado, también se puede intentar reconstruir un argumento de ignorancia sobre la función de un tipo de artefacto, como el de Kornblith, pero evitando, de nuevo, que haya otra comunidad de hablantes que sí conozca su función. Supongamos, por ejemplo, que alguien crea por azar, probando diferentes combinaciones de piezas, un cierto artilugio complejo del que desconoce si puede tener o no alguna función, y acuña un término nuevo para referirse a él, ‘glig’, por ejemplo. Supongamos además que desconoce una descripción que individualice lo que ha creado; el autor sólo sabe describir rudimentariamente su forma y materiales, sin que ese material descriptivo sea suficiente para individualizar el artefacto. El resto de hablantes de la comunidad que tienen noticia de tal artefacto tampoco están en una situación epistémica mejor. Supongamos ahora que, algún tiempo después, se descubre que el glig puede servir para una determinada función (que tiene que ver, por ejemplo, con complementar a otro artefacto también nuevo). Creo que, en esta situación, la intuición que tenemos es que, al menos durante un tiempo, todo el mundo era ignorante sobre la propiedad *relevante*, la función, del artefacto en cuestión y, pese a todo, el correspondiente término para artefacto ya refería a esa cosa o cosas que tenían esa función. Si las situaciones de este tipo son plausibles, entonces también es posible la ignorancia *universal* respecto a la propiedad *relevante* de un cierto tipo de artefacto.

¹⁰¹ Como argumenté en 1.5.1, la posibilidad de ignorancia o error implica en los argumentos con nombres propios o términos de clase natural también es, en última instancia, de tipo epistémico.

En definitiva, si estos argumentos son coherentes y, efectivamente, evitan el inconveniente que presentaban los anteriores argumentos externistas analizados, serían argumentos de error e ignorancia *universales*, en los que todo el mundo estaría equivocado o sería ignorante sobre la función de un cierto tipo de artefacto.

4.5.2. La posible objeción a los argumentos externistas *universales* de la falta de referencia

Quiero ahora considerar la objeción contemplada anteriormente que también se podría hacer a los argumentos que acabo de presentar y mi réplica a esa objeción. Así, igual que sucedía en los primeros argumentos externistas con ‘lápiz’ y con ‘glug’, una posible objeción a las nuevas situaciones planteadas es que, tanto ‘lápiz’ o ‘detergente’ como ‘glig’, no refieren. Y, en las situaciones imaginadas, ‘lápiz’ no refiere porque no hay lápices (objetos que sirven para escribir); igualmente, ‘detergente’ no refiere porque no hay detergente (sustancia que desprende las manchas de la ropa sucia). Y ‘glig’ tampoco refiere porque, aunque quizás haya un glig en el ejemplo, el término no lo selecciona ya que la descripción asociada es insuficiente para determinar la extensión del término.

Contra estas razones que se podrían aducir, ya he argumentado en la sección 4.4.1. Repitiendo otra vez el mismo razonamiento, creo que ‘lápiz’ en el ejemplo de error universal consigue referir a esas cosas a las que llamábamos ‘lápices’¹⁰², aunque nos equivocáramos al pensar que sirven para escribir, y no es un término como ‘bruja’ porque los lápices, a diferencia de las brujas, comparten una propiedad (existente) *relevante* para nosotros, aunque sea una propiedad diferente de aquélla que supusimos en un principio. ¿Cuál es esa propiedad relevante que tienen los lápices en el experimento mental y que *sustituye* a la propiedad de servir para escribir que suponíamos que tenían? ¿Qué propiedad contribuye a determinar la referencia de ‘lápiz’ en el argumento? Creo que quizá se podría considerar que hay dos propiedades distintas candidatas a ser la propiedad *relevante* (seguramente junto a otras propiedades como la forma, etc.) de los ítems a los que aplicamos el término: (i) la propiedad *que* causa cierta ilusión óptica en nosotros; esto es, la propiedad de estar compuesto de grafito, o (ii), la propiedad *de* causar esa ilusión en nosotros (en virtud de la composición química del grafito). Mi intuición es que las propiedades *relevantes* de los lápices serían tanto la

¹⁰² Lo mismo se aplica al caso de ‘detergente’.

propiedad de contener grafito como la propiedad de causar un cierto efecto óptico en nosotros. Así, creo que, en el experimento mental, la propiedad que pasa a ser *relevante* de los lápices es que están hechos de un cierto material (el grafito) que causa en nosotros una cierto efecto óptico. Ahora bien, como esas dos propiedades distintas (estar compuesto de grafito y causar una cierta ilusión óptica) no son necesariamente coextensionales, siempre queda abierta la posibilidad de descubrir otros objetos que también nos causen ese mismo efecto óptico pero que en su composición no contengan grafito. En tal situación, ¿referiría ‘lápiz’ en el argumento expuesto a cualquier cosa que causara en nosotros la ilusión que deja trazos en las superficies o sólo a los objetos que causan esa ilusión pero que contienen grafito? Creo que sería un asunto de elección, por parte de la comunidad lingüística que hace el descubrimiento, sobre cómo se va a usar el término (si abarcando a todo aquello que causa el efecto o a aquello que causa el efecto en virtud de su composición química).

Creo que, en el experimento mental, la propiedad (o propiedades) de estar hechos de un cierto material (el grafito) que causa en nosotros una cierto efecto óptico es una propiedad (descubierta) suficientemente *importante para nosotros* como para ser la propiedad de los lápices que contribuye a determinar la referencia del correspondiente término. Creo, además, que el hecho que se trataría de una propiedad *relevante* para nosotros queda remarcado por el hecho que, tener esa propiedad por parte de los lápices, explica por qué habíamos supuesto erróneamente en ellos la propiedad de servir para escribir. Y el mismo poder explicativo tiene el descubrimiento que hacemos en el experimento mental sobre el detergente: la propiedad que hemos descubierto que tiene el detergente explica por qué estábamos equivocados y suponíamos otra propiedad *relevante* en el detergente. En cambio, como ya he destacado en 4.4.1, en el caso de las brujas, al aceptar (por algo que, aunque no sea un descubrimiento, tiene que ver con la adquisición de más conocimiento) que las personas a las que se aplicaba el término no compartían la propiedad supuesta en ellas, ninguna otra propiedad *relevante* común a esas personas ocupó su lugar. Creo que esta diferencia puede dar cuenta de la diferencia intuitiva que me parece que hay entre los términos sin referente y términos como ‘lápiz’ o ‘detergente’ en los ejemplos de error universal propuestos.

En el caso del término ‘glig’ implicado en el argumento de ignorancia universal, y como ya remarqué antes, simplemente creo que es más un asunto de intuiciones sobre si, en

una situación en que nadie asocia al término una descripción que individualice el dominio de aplicación del término, el término puede referir o no. Yo creo que sí logra referir del mismo modo que un nombre propio o un término de clase natural logra referir intuitivamente aunque nadie conozca una descripción singularizadora del individuo o de la clase en cuestión.

Otra variante de esta objeción de la falta de referencia, que quizás es más vívida y que podría plantearse al menos contra el argumento del error *universal* respecto a los lápices o el detergente, es que no hace falta hipotetizar situaciones como las propuestas, pues ya hay casos reales del todo análogos: por ejemplo, el caso de los amuletos o de las varitas mágicas, casos en los que también parece que se trata de artefactos sobre los que *todos* hemos estado equivocados un tiempo sobre su función o propiedad *relevante*. Y los términos ‘amuleto’ o ‘varita mágica’ -seguiría la objeción- o bien no tienen referente (porque intentaban referir a una propiedad que no existe) o bien su referencia está determinada por las creencias de la gente (así, ‘amuleto’, por ejemplo, referiría a todo aquello de lo que la gente –o alguna gente- cree que trae buena suerte). Y si ‘lápiz’ o ‘detergente’ en los argumentos planteados, son como ‘amuleto’ o ‘varita mágica’, las dos opciones expuestas, respecto a la referencia de estos términos, son indeseables para el externismo.

Bien, yo creo que, aun concediendo que es verdad que, en algún momento, *todos* los hablantes estuvieron equivocados acerca de la naturaleza de las cosas a las que aplicaban el término ‘amuleto’ o ‘varita mágica’, estos casos son diferentes a los casos de ‘lápiz’ o ‘detergente’ tratados en los argumentos. Me parece que, como en el caso de ‘bruja’, lo más sensato es decir que ‘amuleto’ y ‘varita mágica’ no refieren por dos motivos: porque las cosas a las que se aplican o aplicaban esos términos no tienen la propiedad que suponíamos y porque tampoco tienen ninguna otra propiedad *relevante* en común. Una vez asumido que las cosas a las que llamábamos ‘amuletos’ o ‘varitas mágicas’ realmente no tienen la propiedad atribuida, creo que la única propiedad que se podría defender que era común a esas cosas es la propiedad consistente en que *algunas personas pensarán de ellas que tenían ciertos poderes*. Pero creo que ésta no es una propiedad *relevante* a la hora de determinar la referencia del término porque no es una propiedad que esté *en* los objetos a los que aplicamos el término, ni hay nada en esos ítems que cause la propiedad de que creamos que tienen ciertos poderes sobrenaturales.

Además, este tipo de propiedad podría justificar que tienen referente muchos términos que intuitivamente diríamos que no lo tienen. Si se insiste en que la referencia de los términos ‘amuleto’ o ‘varita mágica’ está determinada por las creencias de la gente, ¿de qué creencias se trata?, ¿qué descripción –accesible cognitivamente a algunos hablantes- determinaría la extensión del término? En el caso de ‘varita mágica’, no parece haber ninguna descripción de este tipo capaz de seleccionar a unas varitas, entre todas las varitas, como las que tienen ciertos poderes, como no sea cierto criterio arbitrario difícil de especificar. En definitiva, creo que ‘amuleto’ y ‘varita mágica’ no refieren porque no hay en las cosas a las que se aplican estos términos ninguna propiedad que se relacione causalmente con nosotros del modo apropiado. Y creo que la intuición que hay tras la idea de la relación de referencia es que los términos generales referenciales refieren a alguna propiedad de las cosas con la que tenemos contacto causal. Me parece, entonces, que lo más sensato es defender que ‘amuleto’ es un término general sin referente¹⁰³. En cambio, por todo lo expuesto, creo que ‘lápiz’ o ‘detergente’ sí consiguen referir en los argumentos, y no son como los términos ‘amuleto’ o ‘varita mágica’.

Quiero hacer notar que, desde una teoría externista como la defendida por Dodd¹⁰⁴, inspirada en la de Evans y que admite que la ignorancia o el error por parte de todos los hablantes son compatibles con la referencia de los términos, también se puede argumentar que términos como ‘amuleto’ o ‘varita mágica’ no refieren y, en cambio, ‘lápiz’ o ‘detergente’ en las situaciones comentadas, sí. Según esta teoría, en el caso de ‘lápiz’, la descripción o cuerpo de información (errónea) que la comunidad lingüística asocia con el término, *objeto que sirve para escribir y hacer trazos*, tiene su fuente causal (dominante) en las cosas a las que llamábamos ‘lápiz’, en virtud de presentar esas cosas cierta propiedad, a saber, la propiedad de causarnos cierta alucinación al interactuar nuestra visión con el grafito. Es decir, que, según esta teoría, ‘lápiz’ o ‘detergente’ tienen referencia, pues, en los ejemplos, podemos identificar la fuente

¹⁰³ Pese a que considero que ‘amuleto’ no tiene referente, admito que tiene un uso en nuestro lenguaje que quizás se podría considerar bien delimitado. Mi intuición es que no refiere por los motivos aducidos, pero, de todos modos, los términos sin referente de nuestro lenguaje suponen un problema también para el descriptivista a la hora de explicar su uso bien establecido en nuestras prácticas lingüísticas.

¹⁰⁴ Véase la sección 3.4.2 del capítulo 2.

causal de las creencias (por lo demás, equivocadas) que asociamos con los respectivos términos. Hay una propiedad en los lápices, o en el detergente, que se relaciona del modo apropiado con las creencias que teníamos sobre los lápices o el detergente. La propiedad del detergente que desconocíamos, consistente en causarnos la impresión errónea que quita las manchas de la ropa, es lo que explica que tengamos esas creencias erróneas sobre el detergente. En cambio, en el caso de ‘amuleto’ o ‘varita mágica’, no hay ninguna propiedad obvia en las cosas mismas a las que aplicamos el término que sea la fuente causal de nuestras creencias sobre esas cosas a las que aplicamos estos términos. Creo que sería muy difícil admitir que las varitas a las que se aplicó el término eran la fuente causal de la información asociada con ‘varita mágica’, mientras que otras varitas, prácticamente indistinguibles de las primeras, no eran consideradas varitas mágicas. Y lo mismo para el término ‘amuleto’. Así que, según la propuesta externista de Dodd, ‘amuleto’ y ‘varita mágica’ no refieren a las cosas a las que se aplicaba el término. Y, de manera más general, parece que lo más sensato es decir que tales términos no refieren a nada, pues no parece haber una fuente, al menos una fuente fácilmente identificable, que explique el origen de las creencias que tenemos sobre las cosas de las que decimos (o decíamos) que son amuletos o varitas mágicas¹⁰⁵.

4.5.3. La idiosincrasia de los argumentos universales propuestos

Por último, quiero considerar otra posible objeción del todo diferente contra los argumentos *universales* expuestos que conduce a una reflexión ulterior sobre las características del tipo de argumentos *universales* que he propuesto. La objeción consistiría en decir que los argumentos no son válidos porque los términos implicados en ellos, ‘lápiz’ (o ‘detergente’) y ‘glig’, no son términos de clase artificial. Y no serían términos de clase artificial precisamente porque los supuestos artefactos implicados no son tales, ya que no presentan en los argumentos ninguna función. Así, por ejemplo, aun admitiendo que, en el primer argumento, ‘lápiz’ refiere a esas cosas a las que llamábamos lápices y de las que ahora sabemos que no sirven para escribir, el término ‘lápiz’ no sería un término de clase artificial pues no designaría una clase artificial. En el argumento -continuaría la objeción- los lápices no son artefactos porque ni sirven para escribir ni tienen ninguna otra función; habíamos supuesto que desempeñaban una

¹⁰⁵ Lo mismo puede aplicarse al término ‘bruja’ considerado antes. También, desde esta teoría, se podría defender que es un término sin referente.

función, pero no es así. Entonces, no sólo nos habíamos equivocado al atribuirles una función en particular, sino también al suponer que tenían alguna función. Y si, realmente, no tienen *ninguna* función, ¿por qué deberíamos considerarlos artefactos y, a los términos correspondientes, términos de clase artificial? El mismo razonamiento sirve para ‘detergente’. Incluso se podría decir que ‘glig’, aunque desde el principio refiera al glig, no es un término de clase artificial hasta que se descubra alguna función que pueda desempeñar.

Respecto a esta posible crítica, quiero decir que el descriptivista había concedido que los argumentos decisivos a favor del externismo eran argumentos de ignorancia y error respecto a la característica *relevante* de una clase artificial, su función. Pero, por el problema de alcance limitado que presentaban los argumentos de Kornblith y Nelson, el tipo de argumento que el externista debía de presentar es un argumento de ignorancia o error de alcance *universal*. Esto es, un argumento que muestre que la ignorancia o el error acerca de la función de un tipo de artefacto afectan a todas las comunidades lingüísticas (no sólo a algunas que no han creado el tipo de artefacto en cuestión). Así, este requisito excluye que el argumento externista de alcance *universal* sea uno en que una comunidad lingüística ignora o se equivoca al atribuir una función a un cierto artefacto porque éste tiene realmente otra función conocida por otra comunidad. Entonces, parece que la situación deseada es una en que una comunidad lingüística usa un término para artefacto, pese a que ignora o se equivoca sobre la función que le atribuye, pero sin que haya otra comunidad que conozca (o haya conocido) la verdadera función. Y ésta es la situación que intentan plantear los argumentos *universales* presentados. ¿Son los términos ‘lápiz’ y ‘glig’, en los citados argumentos, términos de clase artificial pese a que refieren a objetos que no tienen una función? Creo que intuitivamente sí. O quizá se podría decir que los lápices, en el ejemplo, tienen la *verdadera* función de causarnos unas ciertas ilusiones ópticas. Pero creo que esto no es una manera apropiada de describir la situación, sino que, más bien, se trataría de que los lápices tienen ese *efecto* (antes que esa función) sobre nosotros.

De todos modos, quien encuentre sumamente problemático que consideremos artefactos a cosas que, como los lápices o el detergente en los experimentos mentales, no tienen una función, puede considerar argumentos análogos a los propuestos pero con una pequeña variación: supongamos que descubrimos que los lápices sólo pintan y escriben

sobre determinadas superficies (algunas que no contienen celulosa) y no sobre otras (las que contienen celulosa), de modo que, debido a una ilusión óptica, nos parece que dejan trazos sobre algunas superficies cuando no es así. O supongamos que descubrimos que el detergente realmente sólo logra desprender ciertas manchas (las que no son de grasa, digamos), mientras que nos parece que también desprende otras (las de grasa) cuando realmente esto último no es el caso. En estos casos, tendríamos artefactos que sí tienen *parte* de la función que les hemos atribuido (y serían artefactos con una función) pero sobre los que estaríamos equivocados en la especificación de su función. Sin embargo, no veo razón por la que ‘lápiz’ en el ejemplo propuesto en primer lugar, aunque no se aplique a objetos con una función propiamente, no pueda ser considerado un término de clase artificial.

Aquí hay que notar que el error que cometemos acerca de la propiedad *relevante* de los lápices no es un error por presuponer una cierta función cuando, realmente, los lápices tienen *otra* función. No se trata de un error por considerar una propiedad concreta de un tipo cuando se trata de otra propiedad concreta del mismo tipo; aquí el error es sobre el tipo mismo de propiedad. Presuponemos que, como los lápices o el detergente son artefactos, lo *importante* para su pertenencia a la clase es su función, pero podría ser que lo *importante* fuera otro tipo de característica, como, por ejemplo, el efecto que pueden tener sobre nosotros o su composición material. En este caso, nos habríamos equivocado sobre el tipo de propiedad relevante. Este tipo de error también es posible con las clases naturales. Así, respecto al agua, por ejemplo, no sólo podríamos equivocarnos en que su composición química es H₂O en vez de otra composición química que podríamos descubrir mañana¹⁰⁶, sino que podríamos equivocarnos al pensar que las muestras de agua comparten una estructura química (sea la que sea), pues podríamos descubrir que sólo comparten características superficiales. Aquí el error sería sobre el tipo de propiedad común que atribuimos a las muestras de agua; no sería una propiedad subyacente sino una propiedad o propiedades de tipo superficial. En tal situación, las características superficiales pasarían a ser las propiedades *relevantes*, *importantes*, para la referencia del término. Sin embargo, este descubrimiento, como

¹⁰⁶ El descubrimiento que plantea Nelson consistente en que los lápices son instrumentos de espionaje alienígena es equivalente a este hipotético descubrimiento que las muestras de agua no tienen la estructura química H₂O, sino otra diferente.

vimos en 2.3.2, no implicaría que la semántica de agua es de tipo descriptivista, pues podemos desconocer o equivocarnos sobre cuál es el tipo de propiedad importante para la pertenencia a una clase natural. No sabemos *a priori* qué tipo de propiedad es. Y el mismo tipo de error, que también considero admisible para las clases artificiales, es el que está en juego en el argumento de error *universal* sobre los lápices o el detergente. En resumidas cuentas, creo que es posible ignorar o equivocarse sobre el tipo de propiedad *relevante* de un tipo de artefacto y, si descubrimos que la propiedad relevante no es de tipo funcional, sino que tiene que ver con su estructura y/o los efectos de ésta, no es admisible decir que no se trata, entonces, de un artefacto. Creo que es contraintuitivo decir que los lápices o el detergente en los ejemplos propuestos no son artefactos.

Ahora bien, aunque creo que los lápices y el detergente en los argumentos son, ciertamente, artefactos, sí podría apreciarse que, en los argumentos expuestos, no se comportan como otros artefactos. Es decir, se podría sospechar que quizá el tipo de argumento de error *universal* propuesto no es ampliable coherentemente a *cualquier* término de clase artificial y, en cambio, el argumento putnamiano de la Tierra Gemela con ‘agua’ sí podría exponerse con cualquier otro término de clase natural, como ‘oro’ o ‘tigre’, digamos, llegando a las mismas conclusiones. Así, ¿podríamos equivocarnos, debido a una ilusión colectiva, o al motivo que se prefiera, sobre la función que tienen, por ejemplo, las sillas o las mesas? Yo creo que parece realmente implausible¹⁰⁷, o, al menos, no está claro qué supondría creer que las sillas sirven para sentarse pero que realmente esto no fuera el caso. Y lo mismo con las mesas, las camas o los destornilladores. ¿Qué diferencia hay, entonces, entre las sillas y los lápices o el detergente?

Me parece que hay al menos dos diferencias que hacen que, en el caso de los lápices y el detergente, un experimento mental de error *universal* sobre su función resulte verosímil, mientras que un argumento de error *universal* sobre la función de las sillas o las camas, no. Mi intuición es que una primera diferencia tiene que ver con poder apelar a una explicación plausible del mecanismo causal implicado en el error colectivo. En mi

¹⁰⁷ Al menos en un escenario que no sea un escenario escéptico del estilo de los escenarios de *a brain in a vat*.

opinión, para que este tipo de argumentos sea verosímil, hay que ofrecer una explicación de por qué –*en virtud de qué*– es posible que cometamos el error que cometemos. Y me parece que una explicación así puede proporcionarse apelando a una cierta estructura interna (tipo de material o composición química) del artefacto como la responsable de nuestra creencia errónea. Así, en los argumentos de error *universal* considerados, con los lápices y el detergente, la ilusión colectiva que hace que estemos equivocados sobre la función que atribuimos a estos artefactos puede explicarse, creo que verosímilmente, apelando a que el material o composición de los artefactos implicados produce en nosotros el efecto de la ilusión óptica y el consiguiente engaño.

En cambio, si intentamos situarnos en una situación epistémica en que las sillas, por ejemplo, no tienen realmente la función que creemos que tienen, no está claro cómo podría explicarse (fuera de suposiciones escépticas radicales) el engaño al que habríamos estado sometidos. Aquí, apelar a alguna propiedad material que tienen las sillas, como el material del que están hechas, resulta difícilmente convincente, pues las distintas sillas están hechas de diferentes materiales y seguramente ni siquiera la mayoría de las sillas comparten un mismo material en su composición. Y parece difícil encontrar otra propiedad compartida por las sillas que pudiera desempeñar ese rol explicativo. En definitiva, mi intuición es que, descartando los escenarios escépticos más radicales, podría ser más fácil construir argumentos de error universal respecto a la función de un tipo de artefacto cuando los artefactos de ese tipo en cuestión incumplen, en alguna medida, el principio de realizabilidad múltiple¹⁰⁸, pues la composición material de un tipo de artefactos puede resultar un factor causal verosímil que actúa sobre nosotros induciéndonos el engaño¹⁰⁹.

¹⁰⁸ Véase la sección 2.2.1.

¹⁰⁹ Aunque quizá no es cierto que todos los lápices ni todas las muestras de detergente comparten entre los miembros de sus respectivas clases algún material de su composición, me parece aceptable considerar que, en alguna medida, son tipos de artefactos que incumplen el principio de la realizabilidad múltiple, pues parte de su composición química (compartida por la mayoría de los miembros de la clase) tiene que ver con que sean el tipo de artefacto que son en un modo en que la composición química de las sillas no tiene que ver con que sean el tipo de artefacto que son.

Esto no quiere decir que la no satisfacción (en alguna medida) del principio de realizabilidad múltiple por parte de una clase de artefactos sea suficiente para construir el argumento requerido con ese tipo de artefacto. Así, por ejemplo, se podría considerar que (la mayoría de) los cuchillos no satisfacen de hecho el principio de realizabilidad múltiple, en alguna medida digna de considerarse, ya que habitualmente llevan acero (de un tipo u otro) en la composición de su hoja, y, sin embargo, no parece verosímil un argumento de error *universal* sobre la función de los cuchillos apelando a que el acero que contienen provoca en nosotros la apariencia de que cortan cuando realmente no es así. Y lo mismo puede decirse de otros artefactos para los que su composición material es una característica importante para que sean el tipo de artefactos que son como, por ejemplo, el pegamento. Me parece, entonces, que en la plausibilidad de los argumentos de error universal hay implicado, al menos, otro factor aparte de una composición material común a los miembros de un tipo de artefactos.

Una hipótesis sobre este segundo factor diferenciador entre las sillas y el detergente es que este factor podría tener que ver con el tipo de función que les suponemos respectivamente, a qué afecta esa función y qué supondría que realmente no se diera esa función. Así, suponemos que el detergente sirve para quitar las manchas, y el efecto de esa función (en caso de tenerla realmente) parece que, en condiciones normales, afectaría a nuestro sistema perceptivo (veríamos que ya no está la mancha que estaba en una prenda antes de someterla a la acción del detergente) sin que ese efecto tenga ulteriores implicaciones. En cambio, pensemos en el caso del error universal respecto a la función de las sillas: creemos que sirven para sentarse en ellas, pero realmente no tienen esta función. Me parece que la falta de verosimilitud de este caso tiene que ver con que la función que atribuimos a las sillas no sólo afecta a nuestro sistema perceptivo (vemos que la gente se sienta en las sillas, por ejemplo) sino que también afecta a muchas relaciones causales del mundo, de modo que la posibilidad de un error así nos situaría en un mundo posible *demasiado lejano* al nuestro. Pensemos qué supondría que estuviésemos engañados sobre la función de las sillas y que éstas nunca hubiesen servido para sentarse encima. Este descubrimiento afectaría a otras muchas cosas, de manera que parece que se enmarca en un escenario escéptico.

De todos modos, delimitar y comprender las condiciones implicadas en la verosimilitud de los argumentos de error *universal* en lo que respecta a la función de un artefacto es

un asunto que requiere una investigación ulterior. En el presente trabajo, me he limitado a explorar cómo podría determinarse la referencia de los términos para artefactos y si podría determinarse de una manera externista, tal como sugiere Putnam al enfocar su atención en el caso de los términos para artefactos e iniciar el debate con el internismo semántico respecto a la determinación de la referencia de estos términos. Mi conclusión es que, si, como creo, la plausibilidad del externismo depende de la plausibilidad de los argumentos de ignorancia y error que puede ofrecer, el externismo para los términos de clase artificial no puede descartarse, pues creo que hay argumentos plausibles de ignorancia y error con algunos tipos de artefactos, quizás no con todos. Si aquellos términos de clase artificial con los que no es verosímil construir los argumentos requeridos por el externista constituyen o no un *tipo* diferente de términos dentro de los considerados términos de clase artificial es una cuestión que también demanda un estudio ulterior.

Conclusiones

Las principales conclusiones de la tesis, siguiendo el orden de los capítulos, son las siguientes:

(1) Los argumentos decisivos para defender una postura semántica de tipo externista para los términos de clase artificial son los argumentos de ignorancia y error.

Tanto para el caso de los nombres propios como para el caso de los términos de clase natural, en el **capítulo 1** vimos que los argumentos decisivos a favor del externismo en su disputa con el internismo, en lo relativo a cómo se determina la referencia de los términos, son los argumentos en que se muestra que los hablantes pueden estar equivocados o pueden ser ignorantes sobre las características distintivas del dominio de aplicación de un término. El motivo por el que estos argumentos son los más importantes para el externismo es que los otros argumentos que se han presentado contra el internismo semántico, los llamados ‘argumento epistémico’ y ‘argumento modal’ que presenta Kripke, no alcanzan a todas las formas de descriptivismo. En particular, el descriptivismo de fijación de la referencia no es socavado por estos argumentos. En cambio, los argumentos de ignorancia o error, los llamados a veces ‘argumentos semánticos’ de Kripke, socavan todas las formas de descriptivismo – excepto una versión *sui generis* del descriptivismo, el descriptivismo causal, que es refutable con argumentos independientes. Naturalmente, si los argumentos de ignorancia o error son los decisivos para defender el externismo semántico aplicado a los nombres propios y a unos términos generales como son los términos de clase natural, no hay ninguna razón para pensar que no va a suceder lo mismo con los términos de clase artificial.

Por otro lado, dado que la tesis internista respecto a la determinación de la referencia es que ésta se determina necesariamente por alguna entidad accesible cognitivamente, la tarea del externista que quiera negar esto ha de consistir en mostrar que la referencia de un término *podría* determinarse sin la concurrencia de ninguna entidad accesible cognitivamente. Esto es, el externista ha de mostrar que la referencia de un término puede determinarse sin que el hablante tenga acceso cognitivo a las propiedades relevantes del referente del término y esto quiere decir que el hablante puede

desconocer o puede equivocarse sobre las propiedades que seleccionan la extensión del término. En definitiva, la mejor y más directa manera de argumentar contra la tesis internista es presentar argumentos de ignorancia y error de los hablantes sobre las propiedades que seleccionan el referente del término.

(2) Una segunda conclusión de la tesis, conclusión del **capítulo 2**, es que los términos para artefactos no son términos de clase nominal.

Algunos autores descriptivistas, como Schwartz, defienden que los considerados términos de clase artificial son términos de clase nominal como ‘soltero’ o ‘cazador’. Sin embargo, si los términos de clase nominal son aquellos términos cuya descripción asociada funciona como una definición ligada analíticamente al término y que proporciona condiciones necesarias y suficientes de pertenencia a la clase, entonces los términos para artefactos no son términos de clase nominal. En primer lugar, no está claro que la definición asociada al término sea una descripción o definición analítica, pues no está claro que sea conocida *a priori* por los hablantes competentes o, incluso, por los expertos. Y, en segundo lugar, es muy difícil, si es que posible en absoluto, dar las condiciones necesarias y suficientes de pertenencia a una clase artificial de manera que esa clase artificial incluya a todos los objetos que consideramos, intuitivamente, miembros de la clase (esto incluye a los miembros rotos, a miembros con formas muy diferentes entre sí, etc.). En cambio, la situación con los términos considerados de clase nominal es otra muy diferente: estos términos sí tienen una definición asociada que está ligada de manera analítica al término y especifica condiciones necesarias y suficientes de pertenencia a la clase. Hay, por tanto, una diferencia crucial entre términos como ‘soltero’ y términos como ‘lápiz’.

Si los términos para artefactos fueran términos de clase nominal, el descriptivismo sería la tesis adecuada para explicar cómo se determina su referencia. Pero como no lo son, el externismo, a falta de otros argumentos en contra, no queda excluido para estos términos.

(3) Las propuestas descriptivistas para los términos para artefactos presentadas hasta ahora, básicamente las teorías de Schwartz y Thomasson, tienen problemas y no son satisfactorias al explicar cómo se determina la referencia de los términos de clase artificial.

Naturalmente, esta conclusión no niega que el descriptivismo pueda proporcionar una explicación satisfactoria de cómo se determina la referencia de los términos de clase artificial, pero, hasta el momento, no la ha proporcionado. En primer lugar, como expuse en el **capítulo 3**, las propuestas descriptivistas presentadas por Schwartz y Thomasson parten de la metafísica de los artefactos para llegar a conclusiones semánticas sobre los términos de clase artificial. Sin embargo, sus respectivas propuestas metafísicas sobre los artefactos no están libres de objeciones. Además, una dificultad importante respecto a la metafísica de los artefactos reside en especificar cuál es la esencia de las clases artificiales: si la función (y qué sentido de *función*), la función más otras propiedades, lo que determina la intención del creador, etc. Dentro del descriptivismo, los autores no se ponen de acuerdo en cuál es la esencia de los artefactos, una esencia que afirman que conocemos.

En segundo lugar, y más importante, las propuestas semánticas descriptivistas presentadas hasta ahora para los términos para artefactos tampoco son satisfactorias. Un descriptivismo *puro e individualista* como el de Schwartz no puede proporcionar la definición analítica, conocida *a priori*, asociada al correspondiente término de clase artificial y que da las condiciones necesarias y suficientes de pertenencia a la clase artificial. Además, al ser un descriptivismo *individualista*, la teoría de Schwartz también presenta el problema obvio de la ignorancia o error por parte de muchos hablantes respecto a la naturaleza de un tipo de artefacto cuando usan el correspondiente término. Un descriptivismo *comunitario* como el de Thomasson parece más prometedor, permitiendo la deferencia a los expertos cuando los hablantes son ignorantes o están equivocados sobre la naturaleza de un artefacto, pero su teoría híbrida de la referencia no está libre tampoco de críticas. Desde mi punto de vista, quizás la principal crítica a una teoría híbrida es que es difícil entender cómo funciona a la hora de determinar la referencia de los términos. Si algo *externo*, no necesariamente accesible cognitivamente determina *parcialmente* la extensión del término, ¿por qué no *enteramente*? Por otro lado, incluso si aceptamos la motivación principal para postular una teoría híbrida de la

referencia –el *qua problem*- hay, al menos, una teoría semántica alternativa que no es híbrida y también da una respuesta al *qua problem*. Creo que la propuesta de Dodd-Evans muestra que una teoría causal también puede ser una solución para los que consideran el *qua problem* un problema genuino y quieren darle una respuesta.

(4) Es posible formular un argumento externista de ignorancia o error de alcance *universal* respecto a la función que suponemos a un cierto tipo de artefacto.

Putnam ya había ofrecido un argumento *universal* de error, similar a un argumento de *Tierra Gemela*, pero que no resultó convincente ni para los filósofos descriptivistas ni para los externistas porque, aunque era un argumento en que el error afectaba a *todo el mundo*, no versaba sobre la característica *relevante* de la clase artificial implicada: su función. Tras el argumento putnamiano, Kornblith y Nelson han presentado argumentos externistas de ignorancia y error respecto a la función de un artefacto, pero el problema, en esta ocasión, es que estos argumentos externistas no tienen un alcance *universal*; esto es, la ignorancia o el error respecto a la función de la clase artificial no alcanzan a *toda* comunidad lingüística. Y, así, el descriptivismo, puede defenderse al menos para *alguna* comunidad lingüística en los contraejemplos que proporcionan Kornblith y Nelson.

Pero creo que el argumento requerido de error *universal*, respecto a la función de un tipo de artefacto, puede construirse apelando a una ilusión colectiva sobre la función que atribuimos a un artefacto y sin que ninguna otra comunidad lingüística quede a salvo de tal error. En el **capítulo 4**, presento algunos argumentos de esta clase: uno con los lápices, y el otro con el detergente. Además, creo que este tipo de argumentos puede defenderse de la principal objeción que puede plantearse contra ellos: la ausencia de referencia de los términos para artefactos implicados. Considero que términos como ‘amuleto’ o ‘varita mágica’ no refieren y, en cambio, ‘lápiz’ o ‘detergente’, en los citados argumentos, sí. La diferencia entre los dos casos es que, en el caso de ‘amuleto’ o ‘varita mágica’, las cosas a las que aplicamos esos términos no tienen en común ni la propiedad que les suponíamos ni ninguna otra propiedad *relevante* en su lugar. Así que estos términos no consiguen referir a nada. Pero en los casos de ‘lápiz’ y ‘detergente’, las cosas a las que aplicamos estos términos en los respectivos argumentos presentados sí tienen referencia porque, aunque esas cosas no tienen la propiedad que les habíamos

atribuido, hemos descubierto que tienen otra propiedad *relevante* en su lugar. Y creo que puede apelarse a esta otra propiedad *relevante* para explicar la relación de referencia entre el término para artefacto en cuestión y el mundo. Además, creo que la teoría causal de Dodd-Evans también puede explicar por qué algunos términos como ‘amuleto’ no refieren mientras que ‘lápiz’ o ‘detergente’, en los experimentos mentales, sí refieren.

(5) La conclusión final y más general de la presente tesis es que una teoría externista semántica, en la línea de las teorías causales de Putnam o Dodd, para los términos para artefactos, parece plausible al menos para algunos términos de clase artificial (quizás no para todos).

Creo que los argumentos cruciales para el externismo, argumentos basados en el error de todo el mundo respecto a la naturaleza *relevante* de los artefactos, son verosímiles en algunos casos (como, por ejemplo, los casos que hemos visto con ‘lápiz’ y ‘detergente’). Pero no está tan claro, intuitivamente, que sean verosímiles en otros casos -casos con sillas, mesas o puentes, por ejemplo- sin apelar a situaciones escépticas radicales. Entonces, ¿cuál es la diferencia entre el detergente y las mesas? Mi hipótesis es que una mayor plausibilidad para el externismo aplicado a los términos para artefactos (o sea, una mayor plausibilidad de los argumentos de error *universal* sobre la función de un artefacto) podría depender de varios factores, pero uno de ellos podría ser la importancia de la estructura interna o composición material del artefacto en cuestión. El motivo es que, para que este tipo de argumentos sea verosímil, mi opinión es que hay que ofrecer una explicación de por qué –*en virtud de qué*– es posible el error que cometemos, y creo que esta explicación puede proporcionarse apelando a una cierta estructura interna del artefacto como la responsable de nuestra creencia errónea (como vimos en los ejemplos de los lápices y el detergente). En cambio, en el caso de las mesas, ¿*en virtud de qué* podríamos equivocarnos sobre la función que suponíamos que tenían las mesas? Ninguna propiedad de las mesas podría explicar nuestro (hipotético) error sobre las mesas porque las mesas no tienen en común ninguna propiedad excepto su función, y no veo ninguna otro posible factor explicativo aparte de un escenario escéptico (y, precisamente por eso, no muy convincente). Así que creo que la relevancia de la composición material para una clase artificial, la no satisfacción en alguna medida

del principio de realizabilidad múltiple, podría ser un factor a favor de una mayor plausibilidad para los argumentos de error *universal* respecto a la función de una clase artificial.

De todas formas, creo que otros factores podrían estar implicados en la plausibilidad de estos argumentos, ya que éstos no parecen verosímiles con algunos artefactos de los que razonablemente se puede pensar que no respetan, en alguna medida, el principio de la realizabilidad múltiple, como, por ejemplo, el pegamento. Mi intuición es que la posibilidad de encontrarnos en un caso de error *universal* respecto a la función del pegamento (sin postular un escenario escéptico radical) no es muy convincente, a pesar de que el pegamento es un artefacto cuya función depende de una cierta composición material. Así que el estudio y la profundización en las condiciones implicadas en la verosimilitud de los argumentos de error *universal* en lo que respecta a la función de un artefacto requieren más investigación. Este trabajo acaba con la conclusión que el argumento requerido de error *universal* parece plausible al menos con algunos términos para artefactos: no sólo los casos que he considerado con ‘lápiz’ y ‘detergente’, sino también con los casos de ‘microondas’, ‘microscopio’, etc. Creo que, mientras estos argumentos sean coherentes, el externismo semántico no puede excluirse al explicar cómo se determina la referencia de los términos de clase artificial. Igualmente, la cuestión de si los términos para artefactos componen un grupo variopinto, con diferentes tipos de términos para artefactos, es una cuestión que necesita estudiarse.

Conclusions

The main conclusions, and further consequences, of this dissertation are the following:

(1) The decisive arguments in order to defend an externalist semantics for artifactual kind terms are the arguments of ignorance and error.

Both in the case of proper names and in the case of natural kind terms, we saw in **chapter 1** that the required arguments in favour of externalism in the dispute with internalism, regarding how reference is determined, are the arguments that show that speakers could be mistaken or ignorant about the distinctive characteristics of the referent of a term. The reason for that is that the other Kripkean arguments presented against semantic internalism, the so called ‘epistemic argument’ and ‘modal argument’, do not undermine all versions of descriptivism. In particular, the reference-fixing version of descriptivism is not undermined by those externalist arguments. In contrast, the Kripkean arguments of ignorance and error, the so called ‘semantic arguments’, undermine all forms of descriptivism –except a peculiar version of descriptivism, causal descriptivism, that must be refuted with independent arguments. Of course, if arguments of ignorance and error are the crucial arguments in order to defend the externalist position for proper names and some general terms such as natural kind terms, there seem to be no reason to think otherwise in the case of artifactual kind terms.

On the other hand, given that the internalist thesis with respect to the determination of reference is that reference is necessarily determined by some cognitively accessible entity, the externalist must show that the reference of a term *could* be determined without any cognitively accessible entity. That is, the externalist must show that the reference of a term can be determined without a speaker having cognitive access to the relevant properties of the referent of the term. Thus, the best way of arguing against internalist thesis is to present arguments where a speaker can be ignorant or mistaken about the properties that pick out the extension of the term in question. That is, to present arguments of ignorance and error.

(2) A second conclusion of the dissertation, a conclusion from **chapter 2**, is that terms for artifacts are not nominal kind terms.

Some descriptivist authors, like Schwartz, defend that artifactual kind terms are nominal kind terms, just as ‘bachelor’ or ‘hunter’. However, if a nominal kind term is a term whose associated description works as a definition analytically tied to the term and that provides necessary and sufficient conditions for membership of the kind, then artifactual kind terms are not nominal kind terms. First, it is not clear that the definition associated with the term is an analytic definition or description, since it is not clear that it is a definition known *a priori* by the competent speakers or, even, by the experts. Second, it seems very difficult, if possible, to give the necessary and sufficient conditions for belonging to an artifactual kind in such a way that this artifactual kind contains all the objects that we intuitively consider members of that kind (and this includes broken members, members with very different forms from the standard, etc.). In contrast, the situation with nominal kind terms is a very different situation: those terms do have an associated definition analytically tied to the term and it specifies necessary and sufficient conditions for belonging to the kind. So, I think there is a crucial difference between terms such as ‘bachelor’ and terms such as ‘pencil’.

If terms for artifacts were nominal kind terms, descriptivism would be the appropriate thesis in order to explain how their reference is determined. But they are not nominal kind terms; so externalism, without any further argument against it, is not ruled out as a plausible point of view for those terms.

(3) To the present moment, the descriptivist theories proposed for artifactual kind terms, basically theories in the line of Schwartz and Thomasson’s theories, have some problems and are not satisfactory in order to account for the determination of reference of terms for artifacts.

Of course, this is not to say that descriptivism cannot provide a satisfactory explanation for the determination of reference of artifactual kind terms, but up to the moment, I think it has not been the case. Firstly, as I explained in **chapter 3**, the descriptivist proposals by Schwartz and Thomasson start from the metaphysics of artifacts to arrive to semantic conclusions about terms for artifacts. However, their respective metaphysical theories about artifacts are not free from objections. Furthermore, there is an important difficulty concerning the metaphysics of artifacts in specifying what

constitutes the essence of artifactual kinds: whether the function (and what notion of function), the function plus other properties, whatever determines the creator's intention, etc. Descriptivists do not agree on what is the essence of artifacts, an essence that they claim we know.

Secondly, and more importantly, the descriptivist semantic proposals proposed up to now for artifactual kind terms are not satisfactory either. A *pure* and *individualistic* descriptivism as that of Schwartz's cannot provide the analytic definition, known *a priori*, associated with the correspondent artifactual kind term which gives the necessary and sufficient conditions for being a member of the artifactual kind. Moreover, Schwartz's *individualistic* theory also faces the obvious problem of ignorance or error by many speakers about the nature of the corresponding type of artifact. The Thomassonian *communitarian* descriptivism seems more promising, allowing for deference to experts when speakers are ignorant or mistaken about the nature of an artifact, but her hybrid theory of reference is not free from criticisms. From my point of view, perhaps the main criticism to such a semantic theory is that it is hard to understand how a hybrid theory is supposed to work. If something *external*, not necessarily cognitively accessible, is to *partly* determine the extension of a term, why not *wholly*? On the other hand, even if we accept the main motivation for a hybrid theory of reference –the *qua problem*- there is, at least, one alternative theory which is not hybrid and gives an answer to the mentioned problem. I think that Dodd's proposal shows that a causal theory can be also a solution for those who want to respond to the *qua problem*.

(4) It is possible to formulate externalist arguments of ignorance or error of *universal* scope concerning the *relevant* nature of an artifact, its function.

Putnam had presented an argument of *universal* error, the pencils-organisms thought experiment, but it was not sound because, although it was an argument where the error affected *everybody*, according to both externalist and internalist philosophers it was not an argument of error about the *relevant* property of the artifactual kind involved: its function. After the Putnamian argument, Kornblith and Nelson have presented externalist arguments of ignorance and error concerning the function of an artifactual

kind, but, this time, the problem is that those arguments do not have a *universal* scope; that is, the ignorance or error with respect to the function of the artifactual kind do not reach *every* linguistic community. And, thus, descriptivism can be preserved at least for *some* linguistic community.

But I think that the required argument of *universal* error, concerning function, can be construed appealing to a collective illusion, within a community, about the function we attribute to a certain kind of artifact, and without any other linguistic community safe from this mistake. In **chapter 4**, I present some arguments of that kind: one involving pencils, and other involving detergent. Moreover, I think that this type of argument can be defended from the main objection that can be made against it: the lack of reference of the artifactual terms involved. I defend that terms like ‘charm’ or ‘magic wand’ do not refer, and, in contrast, ‘pencil’ or ‘detergent’, in the mentioned arguments, do have reference. The difference between these two cases is that, in the former cases, the things called ‘charm’ (or ‘magic wand’) have in common neither the property we supposed them to have nor any other *relevant* property instead; so, ‘charm’ (and the same for ‘magic wand’) do not refer to anything. But, in the latter cases, the things we called ‘pencils’ (or ‘detergent’) in the proposed arguments do have reference because, although those things do not have the property we attributed to them, we have discovered that those things have another *relevant* property instead. And I think this another *relevant* property can be appealed to in order to explain the relation of reference between the term for artifact and the world. Furthermore, I think that the causal theory of Dodd-Evans can also explain why some terms like ‘charm’ do not refer whereas ‘pencil’ or ‘detergent’, in the thought experiments, do refer.

(5) The final and general thesis of this dissertation is that a semantic externalist proposal, along the lines of the causal theories of Putnam or Dodd, for terms for artifacts seems plausible at least for some terms of artifactual kind (maybe not for all).

I think that crucial arguments for externalism, arguments based on error by everybody concerning the *relevant* nature of artifacts, are available in some cases (like the cases we have seen with ‘pencil’ and ‘detergent’) but it is not so clear, intuitively, in some other cases (cases such as tables, chairs, bridges...) without appealing to radical sceptical

situations. Thus, what is the difference between detergent and tables? My hypothesis is that the greater plausibility of externalism for artifactual kind terms -the greater plausibility of arguments of *universal* error about the function of an artifact-, could depend on several factors, but one of them could be the relevance of the internal structure or material composition of the artifact in question. This is so because in order to make plausible this type of argument, an explanation of why *-in virtue of what-* the error is possible must be given, and I think that this explanation can be given appealing to a certain internal structure of the artifact as responsible of our mistake (as we saw in the examples of pencils and detergent). In contrast, in the case of tables, *in virtue of what* could we be mistaken about the function we supposed them to have? No property of tables could explain our (hypothetical) mistaken belief about tables, because tables do not have in common any property except their function, and I do not see any other possible explanatory factor besides a sceptical (and, because of that, not very convincing) scenario. So I think that the importance of the material composition for an artifactual kind, the no satisfaction –to some extent- of the principle of multiple realizability, could be a factor in the greater plausibility of arguments of *universal* error concerning the function of an artifactual kind.

Other factors could be involved in the plausibility of those arguments since they do not seem plausible with other artifactual kinds that can reasonably be said not to respect, to some extent, the principle of multiple realizability, such as glue, for instance. My intuition is that the possibility of being in a case of *universal* mistake with respect to the function of glue is not very plausible (without postulating a radical sceptical situation), even though glue is an artifact whose function depends on a certain material composition. So, the study and deepening of the conditions involved in the plausibility of the arguments of *universal* error concerning the function of an artifact require further investigation. This work ends with the conclusion that the required argument of *universal* error seems plausible at least with some terms for artifacts: not just the considered cases of ‘pencil’ and ‘detergent’, but also the cases of ‘microwave’, ‘microscope’, etc. So, I think that semantic externalism cannot be ruled out in explaining how the reference of artifactual terms is determined. Likewise, whether artifactual kind terms constitute a colourful *kind*, with different types of terms for artifacts, is an open question that must be studied.

Bibliografía

- Baker, Lynne R. (2004): “The Ontology of Artifacts”, *Philosophical Explorations* 7, 99-112.
- Carrara, Massimiliano y Vermaas, Pieter E. (2009): “The Fine-grained Metaphysics of Artifactual and Biological Functional Kinds”, *Synthese* 169, 125-43.
- Davidson, Donald (1992): *Mente, mundo y acción*, Paidós (Trad. de Carlos Moya), Barcelona.
- Dennet, Daniel C. (1987): *The Intentional Stance*, Basil Blackwell, Nueva York.
- Dennet, Daniel C. (1990): “The Interpretation of Texts, People and Other Artifacts”, *Philosophy and Phenomenological Research* 50, 177-94.
- Devitt, Michael y Sterelny, Kim (1987): *Language and Reality. An Introduction to the Philosophy of Language*, Basil Blackwell, Oxford.
- Devitt, Michael (2005): “Rigid Application”, *Philosophical Studies* 125, 139-65.
- Dodd, Julian (2012): “Defending the Discovery Model in the Ontology of Art: A Reply to Amie Thomasson on the *Qua* Problem”, *British Journal of Aesthetics* 52, 75-95.
- Donnellan, Keith (1983): “Kripke and Putnam on Natural Kind Terms”, en Carl Ginet y Sydney Shoemaker (eds.), *Knowledge and Mind*, Oxford University Press, Oxford, 84-104.
- Elder, Crawford L. (2009): “On the Place of Artifacts in Ontology”, en Eric Margolis y Stephen Laurence (eds.), *Creations of the Mind. Theories of Artifacts and their Representation*, Oxford University Press, Oxford, 33-51.

- Fitch, Greg y Nelson, Michael (2012): "Singular Propositions", *The Stanford Encyclopedia of Philosophy*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/win2012/entries/propositions-singular/>.
- García-Carpintero, Manuel (1996): *Las palabras, las ideas y las cosas*, Ariel, Barcelona.
- Gelman, Susan A. (1988): "The development of induction within natural kind and artifactual categories", *Cognitive Psychology* 20, 65-95.
- Gómez-Torrente, Mario (2006): "Rigidity and Essentiality", *Mind* 115, 227-60.
- Gutheil, Grant et al. (2004): "The role of historical intuitions in children's and adults' naming of artifacts", *Cognition* 9, 23-42.
- Kornblith, Hilary (1980): "Referring to Artifacts", *Philosophical Review* 89, 109-14.
- Kornblith, Hilary (2009): "How to Refer to Artifacts", en Eric Margolis y Stephen Laurence (eds.), *Creations of the Mind. Theories of Artifacts and their Representation*, Oxford University Press, Oxford, 138-49.
- Kripke, Saul (1986): "A Problem in the Theory of Reference: the Linguistic Division of Labor and the Social Character of Naming", en V. Cauchy (ed.), *Philosophy and Culture. Proceedings of the XVIIth World Congress of Philosophy*, Editions Montmorency, Montreal, 241-47.
- Kripke, Saul A. (1988): *Naming and Necessity*, Basil Blackwell, Oxford.
- Lambon Ralph, Matthew A. et al. (1998): "Are Living and Non-living Category-specific Deficits Causally Linked to Impaired Perceptual or Associative Knowledge? Evidence from a Category-specific Double Dissociation", *Neurocase* 4, 311-38.
- Lawler, Diego (2003): "Las funciones técnicas de los artefactos y su encuentro con el constructivismo social en tecnología", *Revista CTS* 1, 27-71.

- Lawler, Diego y Vega, Jesús (2010): “Clases artificiales”, *Azafea: Revista de Filosofía* 12, 119-47.
- Lawler, Diego y Vega, Jesús (2011): “Realizabilidad múltiple y clases de artefactos” *Revista CTS* 7, 167-78.
- Lawler, Diego y Vega, Jesús (próxima publicación en 2013): “Creating Artifactual Kinds” en M. Franssen, P. Kroes, T. A. C. Reydon y P.E. Vermaas (eds.), *Artifact Kinds: Ontology and the Human-Made World*, Synthese Library (Borrador).
- Locke, John (1975): *An Essay concerning Human Understanding*, Oxford University Press, Oxford.
- Malt, Barbara C. & Sloman, Steven A. (2006): “Category Essence or Essentially Pragmatic? Creator’s intention in naming and what’s really what”, *Cognition* 105, 615-48.
- Marconi, Diego (1987): “Two Aspects of Lexical Competence”, *Lingua e Stile* 22, 385-95.
- Marconi, Diego (2009): “Being and Being Called: Paradigm Case Arguments and Natural Kind Words”, *Journal of Philosophy* 106, 113-36.
- Martí, Genoveva (2004): “Rigidity and General Terms”, *Proceedings of the Aristotelian Society* 104, 131-48.
- Mill, John S. (1973): *A System of Logic Ratiocinative and Inductive*, en J.M. Robson (ed.), *Collected works of J. S. Mill*, volume VII, University of Toronto Press, Toronto.
- Millikan, Ruth G. (1989): “In Defense of Proper Function”, *Philosophy of Science* 56, 288-302.
- Millikan, Ruth G. (1999): “Wings, Spoons, Pills and Quills: a Pluralist Theory of Function”, *The Journal of Philosophy* 96, 191-206.

- Mingardo, Daria (2011): “Making (Another) Sense of a Debate on Artifacts Categorization” (Borrador).

- Nelson, James A. (1982): “Schwartz on Reference”, *Southern Journal of Philosophy* 20, 359-65.

- Pérez Otero, Manuel (1998): “Las distinciones modales de Kripke y el concepto de proposición”, *Teorema* 17, 61-74.

- Putman, Daniel A. (1982): “Natural Kinds and Human Artifacts”, *Mind* 91, 418-19.

- Putnam, Hilary (1962): “It Ain’t Necessarily So”, *The Journal of Philosophy* 59, 658–71.

- Putnam, Hilary (1973): “Meaning and Reference”, *The Journal of Philosophy* 70, 699-711.

- Putnam, Hilary (1975): “The Meaning of ‘Meaning’”, en Keith Gunderson (ed.), *Language, Mind and Knowledge*, University of Minnesota Press, Minneapolis.

- Putnam, Hilary (1979): “The Analytic and the Synthetic” en Hilary Putnam, *Mind, Language and Reality. Philosophical Papers Vol. 2*, Cambridge University Press, Cambridge.

- Russell, Bertrand (1905): “On Denoting”, *Mind* 14, 479-93.

- Russell, Bertrand (1986): *Los problemas de la filosofía*, Editorial Labor, Barcelona.

- Salmon, Nathan U. (1979): “How Not to Derive Essentialism from the Theory of Reference”, *The Journal of Philosophy* 76, 703-25.

- Schwartz, Stephen P. (1978): “Putnam on Artifacts”, *Philosophical Review* 87, 566-74.

- Schwartz, Stephen P. (1980): “Natural Kinds and Nominal Kinds”, *Mind* 89, 182-95.

- Schwartz, Stephen P. (1983): “Reply to Kornblith and Nelson”, *Southern Journal of Philosophy* 21, 475-79.

- Searle, John R. (1983): *Intentionality: An Essay in the Philosophy of Mind*, Cambridge University Press, Cambridge.

- Searle, John R. (1999): “Nombres propios y descripciones”, en Luis M. Valdés (comp.), *La búsqueda del significado*, Tecnos, Madrid, 106-15.

- Soames, Scott (2005): *Reference and Description: The Case against Two-Dimensionalism*, Princeton University Press, Princeton.

- Soames, Scott (2011): “Kripke on Epistemic and Metaphysical Possibility: Two Routes to the Necessary *A posteriori*”, en Alan Berger (ed.), *Saul Kripke*, Cambridge University Press, Cambridge.

- Sullivan, Arthur (2005): “Rigid Designation, Direct Reference, and Modal Metaphysics”, *Pacific Philosophical Quarterly* 86, 577-99.

- Thomasson, Amie L. (2003): “Realism and Human Kinds”, *Philosophy and Phenomenological Research* 67, 580-609.

- Thomasson, Amie L. (2007): *Ordinary Objects*, Oxford University Press, Oxford.

- Thomasson, Amie L. (2009a): “Artifacts and Human Concepts”, en Eric Margolis y Stephen Laurence (eds.), *Creations of the Mind. Theories of Artifacts and their Representation*, Oxford University Press, Oxford, 52-73.

- Thomasson, Amie L. (2009b): “The Significance of Artifacts for Metaphysics”, en Antonie Meijers (ed.), *Handbook of Philosophy of the Technological Sciences*, Elsevier Science, 191-212.

- Warmke, Brandon (2010): “Artifact and Essence”, *Philosophia* 38, 595-614.
- Wiggins, David (1980): *Sameness and substance*, Basil Blackwell, Oxford.
- Wikforss, Åsa (2005): “Naming Natural Kinds”, *Synthese* 145, 65-87.
- Wikforss, Åsa (2010): “Are Natural Kind Terms Special?”, en Helen Beebe y Nigel Sabbarton-Leary (eds.), *The Semantics and Metaphysics of Natural Kinds*, Routledge, Londres, 64-83.
- Wittgenstein, Ludwig (2004): *Investigaciones filosóficas*, Crítica (Trad. de Alfonso García Suárez y Ulises Moulines), Barcelona.