



FILO:UBA
Facultad de Filosofía y Letras
Universidad de Buenos Aires

G

La polémica entre el atomismo conceptual y el neo-empirismo de conceptos.

Autor:

Haimovici, Sabrina

Tutor:

Skidelsky, Liza

2010

Tesis presentada con el fin de cumplimentar con los requisitos finales para la obtención del título Licenciatura de la Facultad de Filosofía y Letras de la Universidad de Buenos Aires en Filosofía.

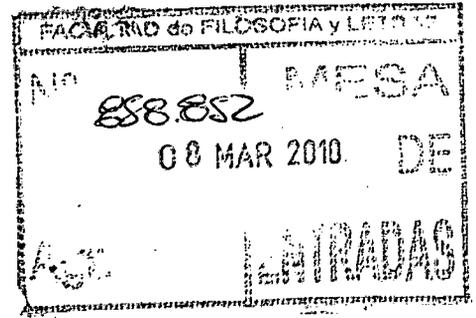
Grado



FILO:UBA
Facultad de Filosofía y Letras

FILODIGITAL
Repositorio Institucional de la Facultad
de Filosofía y Letras, UBA

Tesis
14.2.29



UNIVERSIDAD DE BUENOS AIRES

Facultad de Filosofía y Letras

Tesis de Licenciatura

**La polémica entre el atomismo conceptual
y el neo-empirismo de conceptos**

Estudiante: Sabrina Haimovici

Directora de Tesis: Dra. Liza Skidelsky

UNIVERSIDAD DE BUENOS AIRES
FACULTAD DE FILOSOFÍA Y LETRAS
Dirección de Bibliotecas

Marzo de 2010

ÍNDICE

Agradecimientos	3
Introducción	4
1. Las teorías de conceptos	4
2. El atomismo conceptual y el neo-empirismo de conceptos	13
3. Los requisitos explicativos fundamentales de una teoría de conceptos	15
4. La posibilidad de una teoría pluralista de conceptos	17
5. Plan de la tesis	19
Capítulo 1. El atomismo conceptual. La satisfacción de los requisitos de intencionalidad, alcance y composicionalidad	22
1. Características de los conceptos	22
2. Satisfacción de los requisitos	29
2.1 Alcance	29
2.2 Composicionalidad	30
2.3 Intencionalidad	32
3. Conclusiones	40
Capítulo 2. El atomismo conceptual. La satisfacción de los requisitos de adquisición y categorización	42
1. La Teoría Representacional/Computacional de la Mente	42
2. Satisfacción de los requisitos	48
2.1 Adquisición	48
2.2 Categorización	54
3. Conclusiones	60
Capítulo 3. El neo-empirismo de conceptos. La satisfacción de los requisitos de intencionalidad, alcance y composicionalidad	61
1. Características de los conceptos	61
2. Satisfacción de los requisitos	68
2.1 Intencionalidad	68
2.2 Alcance	72
2.3 Composicionalidad	77
3. Conclusiones	85
Capítulo 4. El neo-empirismo de conceptos. La satisfacción de los requisitos de adquisición y categorización	86
1. La Teoría de la Reactivación de Estados Perceptivos	86
2. Satisfacción de los requisitos	93

2.1 Adquisición	93
2.2 Categorización	96
3. Conclusiones	102
Capítulo 5. El pluralismo extendido	103
1. Los conceptos como una clase funcional	104
2. El abandono de un único tipo de estructura	106
3. Requisitos y representaciones	110
4. El poder explicativo del pluralismo extendido	116
5. Conclusiones	124
Conclusiones	126
Referencias	132

Agradecimientos

Quisiera expresar, en primer lugar, un profundo agradecimiento a Liza Skidelsky, por el apoyo y la paciencia con los que guió mi investigación. Fundamentalmente agradezco todo lo que me enseñó en este proceso, sus observaciones y sugerencias fueron un constante desafío a avanzar un poco más. Su dedicación y forma de trabajo son un ejemplo para mí.

Quiero agradecer a Sergio Barberis y Mariela Destéfano, con quienes compartí un grupo de lectura sobre los temas de esta tesis y quienes influyeron en muchas de las ideas que aquí desarrollo. Quiero agradecer también a todos los integrantes del grupo de investigación Cognición, Lenguaje y Conciencia (coordinado por Diana Perez y Liza Skidelsky). Cada reunión de los jueves es un incentivo para profundizar en la investigación filosófica. Discutir parte de este trabajo con ellos fue muy enriquecedor.

En los últimos años presenté trabajos relacionados con esta tesis en distintos congresos y jornadas. Quisiera agradecer a quienes escucharon y discutieron esas ponencias. Sus preguntas, observaciones y críticas contribuyeron a desarrollar mis ideas y darle forma a esta tesis.

Quiero agradecer con mucho cariño a mi familia y amigos, por el apoyo y en especial por la comprensión en toda esta última etapa de trabajo intensivo. Fueron siempre un sostén fundamental. Y, por último, a Patricio, por todo. No sé si hubiera podido hacer este trabajo sin su compañía, sé que seguro hubiese sido menos feliz.

Introducción

1. Las teorías de conceptos

En los últimos años, las teorías de conceptos han recibido gran atención en el ámbito de las ciencias cognitivas y la filosofía de la mente. Se han elaborado distintas teorías respecto de qué son los conceptos, qué estructura tienen, cómo se adquieren y cómo intervienen en diversos procesos mentales. En general, hay algo en lo que todas acuerdan: los conceptos son entidades encargadas de explicar gran parte de las características de nuestra vida mental y de la relación entre nuestras mentes y el mundo. Así, algunos autores, como Fodor (1987), señalan que dado que los conceptos son los constituyentes de los pensamientos, estos son las entidades que permiten explicar, por un lado, la intencionalidad de los pensamientos y, de esta manera, la relación entre los pensamientos y el mundo y, por otro lado, la realización de inferencias deductivas que dependen de la estructura de los pensamientos. En cambio, otros autores, como Prinz (2002) y Murphy (2002), enfatizan el rol de los conceptos como portadores de conocimiento y se concentran en la función que aquellos cumplen en las tareas de categorización. De manera que, independientemente de la postura teórica adoptada, tanto filósofos como psicólogos acuerdan en que los conceptos cargan con una importante responsabilidad en nuestra vida mental y, en este sentido, una teoría de conceptos formará parte de la explicación de cómo está estructurado nuestro pensamiento y de cómo funciona, en general, nuestra mente.

Se puede clasificar a las distintas teorías de conceptos que se han propuesto en función de la postura que adoptan respecto de una serie de debates centrales en torno a los mismos. Estos debates son, particularmente, acerca de la ontología de los conceptos, su estructura, y los requisitos explicativos de una teoría de conceptos. También hay otros debates que han tomado entidad propia, entre ellos, la polémica innatismo-empirismo en adquisición de conceptos, la relación entre los conceptos y el lenguaje natural, y la distinción conceptual-no conceptual. En lo que sigue describiré de manera general qué está en juego en los debates centrales puesto que esta tesis se ocupará primordialmente de ellos (aunque, por supuesto, se tocarán aspectos relacionados con algunos de los otros debates mencionados).

Siguiendo a Margolis y Laurence (2008), se pueden delinear tres grandes perspectivas en función de las características ontológicas asignadas a los conceptos: la primera sostiene que estos son representaciones mentales, la segunda los identifica con habilidades o capacidades y la última con entidades abstractas. La perspectiva según la cual los conceptos son representaciones mentales postula que estos son entidades psicológicas. Esta perspectiva en general acompaña a una teoría representacional de la mente, según la cual el pensamiento se produce en un sistema interno de representaciones. Así, se postula que los conceptos son las representaciones más básicas de este sistema. Algunos consideran que se trata de particulares mentales que presentan una semántica composicional y una sintaxis combinatoria semejante a la de los lenguajes naturales (Fodor 1984). Otros autores que adoptan el enfoque de las representaciones mentales postulan, sin embargo, que éstas están constituidas por copias de representaciones perceptivas (Prinz 2002).

En segundo lugar, quienes consideran que los conceptos son habilidades sostienen que estos consisten en ciertas capacidades cognitivas de los agentes, tales como extraer inferencias o discriminar entre los ítems que caen bajo un concepto y aquellos que no. Por último, la postura que entiende a los conceptos como sentidos fregeanos los identifica con objetos abstractos. Según esta perspectiva, los conceptos no están “en la cabeza”, sino que son entidades abstractas independientes de la mente, que ésta capta cuando tiene un pensamiento (Margolis & Laurence 2008). Se podría agregar como cuarta perspectiva una versión mixta, según la cual los conceptos son representaciones mentales, pero se individualizan en función de los sentidos que expresan. Quienes adhieren a esta perspectiva sostienen que los conceptos son representaciones mentales, siendo los sentidos los intermediarios entre estas representaciones y sus referentes (Beck ms.).¹

Por otra parte, se han suscitado debates en torno a la estructura asignada a los conceptos. En primer lugar, se puede establecer una gran división entre las teorías en función de si proponen o no que los conceptos tienen una estructura interna. Así, se puede agrupar a la teoría clásica de las definiciones, la teoría de los ejemplares, la de los prototipos y la teoría-teoría como aquellas que proponen una estructura interna, en oposición al atomismo.² El atomismo conceptual sostiene que la mayoría de los conceptos

¹ Aunque véase Laurence y Margolis (2007) para los problemas que acarrea esta postura.

² Fodor (1999, p.14) advierte esta distribución tan dispar cuando afirma que: “A diferencia de prácticamente todos los demás que trabajan o trabajaron en esta tradición, considero que la teoría de los conceptos debe

léxicos no tiene estructura interna, sino que son atómicos y su contenido queda determinado exhaustivamente por relaciones nomológicas mente-mundo -las cuales no involucran constitutivamente a otros conceptos (Fodor 1999, 2008). Por otro lado, las teorías que asignan una estructura constitutiva a los conceptos se diferencian entre sí por el tipo de información que consideran que los conceptos transportan.

La teoría clásica propone que los conceptos tienen una estructura definicional, es decir, están constituidos por un conjunto de conceptos que expresan las condiciones necesarias y suficientes para caer bajo el concepto en cuestión (Laurence & Margolis 1999, Margolis & Laurence 2008, Pitt 1999). La teoría de los ejemplares, por otra parte, sostiene que los conceptos son conjuntos de representaciones de instancias que caen bajo una misma categoría (Smith & Medin 1981, Machery 2009). En cambio, según la teoría de prototipos, los conceptos poseen información de tipo estadístico y están estructurados a partir de rasgos típicos o diagnósticos para pertenecer a una categoría (Hampton 2006, Hampton & Jönsson 2008, Osherson & Smith 1981). Estas dos últimas vertientes suelen considerarse emparentadas, puesto que ambas sostienen que los conceptos refieren en virtud de una relación de similitud entre las representaciones conceptuales y las instancias que representan. Finalmente, la teoría-teoría sostiene que los conceptos están estructurados como teorías -i.e como conjuntos de proposiciones- y contienen información sobre mecanismos causales y propósitos teleológicos, así como sobre rasgos ocultos y esencias (Gopnik & Meltzoff 1997, Gopnik & Schulz 2007, Rehder & Burnett 2005, Sloman 2005).

Además, se puede identificar un cuarto grupo de teorías que, si bien asignan estructura interna a los conceptos, ésta no es una estructura uniforme, de manera que no postulan un único tipo de estructura conceptual. Se han propuesto, por un lado, teorías híbridas -i.e. que afirman que los conceptos tienen partes con estructuras de distinto tipo, por ejemplo, prototipos, definiciones, ejemplares- y, por otro lado, teorías pluralistas -que sostienen que la clase de los conceptos incluye estructuras de distinto tipo, es decir, que hay distintos tipos de conceptos en función de su estructura interna. Entre las primeras se destaca la teoría de los proxitipos de Prinz (2002, 2005) según la cual un concepto puede tener distintas estructuras con distinto tipo de información -e.g. representaciones de ejemplares, una parte prototípica, otra definicional e información de relaciones causales,

ser atomista. Me siento un poco solitario estando aquí afuera, sin compañía...".

entre otras. Se la puede considerar híbrida, en tanto propone distintos tipos de estructura como *partes* de cada concepto. En cambio, las propuestas pluralistas sostienen, por un lado, que la clase de los conceptos incluye distintos tipos de estructuras conceptuales y, por otro lado, que cada categoría corresponde a una multiplicidad de conceptos con estructuras distintas (Weiskopf 2009a, 2009b).

Es habitual que las teorías de conceptos se organicen en torno a *desiderata* o requisitos explicativos, esto es, que partan de la enumeración de una serie de fenómenos de los cuales la teoría pretende dar cuenta. Pues bien, no hay un acuerdo total entre las distintas teorías respecto de estos requisitos. Algunos autores distinguen entre los *desiderata* o requisitos explicativos de una teoría psicológica de conceptos y los de una teoría filosófica (Fodor 2000, Machery 2009). En las teorías psicológicas predomina el objetivo de explicar procesos cognitivos, fundamentalmente, la categorización, el razonamiento causal y el aprendizaje de relaciones causales. Las teorías filosóficas, en cambio, se centran en la explicación de la capacidad productiva y sistemática del pensamiento -i.e. el hecho de que podemos pensar infinitos pensamientos y el hecho de que podemos pensar pensamientos relacionados entre sí-, en cuestiones acerca del contenido intencional y en principios como la composicionalidad -que establece que el contenido de los conceptos compuestos debe quedar exhaustivamente determinado por el contenido de sus constituyentes y las reglas de combinación sintáctica (Fodor y Pylyshyn 1988, Fodor 1999, por ejemplo). De esta manera se puede identificar como teorías filosóficas a aquellas que se concentran más en la metafísica del pensamiento, fundamentalmente, en la naturaleza de sus constituyentes y la individuación de los mismos, en oposición a las teorías psicológicas, que se concentran más bien en los procesos cognitivos, fundamentalmente, en los procesos de categorización e inducción.

No obstante, también hay propuestas teóricas que toman en cuenta tanto los requisitos psicológicos como los filosóficos (Prinz 2002, Weiskopf 2009b) y establecen una lista de *desiderata* más amplia. En el caso de Prinz (2002), se incluyen los requisitos de alcance -i.e. el requisito de dar cuenta de una amplia gama de conceptos tales como los de objetos, eventos, conceptos abstractos y teóricos-, contenido intencional -i.e. aquello acerca de lo cual son los pensamientos-, contenido cognitivo -que permite explicar los casos de diferencias entre conceptos coextensivos y semejanzas entre conceptos con distinta

referencia (los usualmente llamados casos Frege y casos de tierra gemela, respectivamente)-, adquisición, categorización -tanto el reconocimiento de instancias que caen bajo una categoría cuanto el anticipo de las características que tendrá una instancia por pertenecer a una categoría-, composicionalidad y publicidad –según este último requisito los conceptos deben ser compartibles por distintos individuos y por un mismo individuo en distintos momentos. En lo que sigue desarrollaré con más detalle los requisitos que más comúnmente se tienen en cuenta.

Intencionalidad

Una teoría de conceptos debe explicar en qué consiste que un concepto sea acerca de algo, i.e., tenga contenido intencional. Es necesario establecer qué tipo de información portan los conceptos -por ejemplo, si es conocimiento proposicional, estadístico, o de otro tipo y si la información está estructurada o es atómica-, y qué tipo de relación tienen con el mundo -por ejemplo, si se trata de una relación causal o de similitud-, así como qué relación tienen con los otros conceptos.

Las teorías semánticas tradicionales más desarrolladas respecto del contenido de los conceptos son la semántica informacional (Dretske 1981, Fodor 1990b) y la semántica de rol conceptual (Block 1986). Según la semántica informacional, el contenido consiste en información acerca del mundo (propiedades o eventos) y los conceptos portan esa información en virtud de relaciones de covariación nomológica con sus referentes. Según la semántica de rol conceptual, en cambio, el contenido está determinado por los roles inferenciales de los conceptos, es decir, por las inferencias en las que estos participan. Así, a diferencia de la semántica informacional, el contenido de un concepto se determina por las relaciones que mantiene con otros conceptos.

Por otra parte, distintos autores proponen que es necesario incorporar un requisito sobre el contenido cognitivo (Fodor 1987, Prinz 2002, Putnam 1975). Este tipo de contenido permitiría dar cuenta de los llamados casos Frege –casos de distintos conceptos con idéntico contenido intencional- y los casos de tierra gemela –en los cuales un mismo concepto tiene distintos contenidos intencionales en distintos mundos posibles. Este tipo de contenido se relaciona generalmente con las concepciones epistémicas de los individuos y se propone para permitir una individuación de conceptos de grano más fino que la que

permitiría únicamente el contenido intencional. Así se explicaría, por un lado, cómo en los casos Frege un individuo podría tener creencias contradictorias que involucren conceptos que refieren a la misma propiedad en el mundo (e.g. Luisa Lane cree que Superman vuela, y a la vez cree que Clark Kent no vuela) y, por otro lado, cómo en los casos de tierra gemela dos individuos pueden tener creencias y comportamientos idénticos en relación a propiedades distintas (e.g. en relación a H₂O en la tierra actual y a XYZ en la tierra gemela).

Algunos autores, como Fodor (1987), propusieron que los casos de tierra gemela se explican postulando un contenido estrecho (que, a diferencia del cognitivo, no es epistémico), el cual se define en función del contenido intencional o amplio. Mientras que el contenido amplio está constituido por propiedades en el mundo y determinado por relaciones nomológicas, el contenido estrecho es una función de contextos a condiciones de verdad o a contenido amplio (Skidelsky en prensa).³ Así, por ejemplo, AGUA en el contexto de esta tierra se proyecta en H₂O, mientras que en el contexto de la tierra gemela se proyecta en XYZ. Otros filósofos, como Block (1986), consideran que una semántica de rol conceptual puede dar cuenta del contenido cognitivo. En versiones más actuales, en cambio, el contenido cognitivo se explica apelando a la misma teoría semántica que explica el contenido intencional (Prinz 2002). Por eso, en §3, consideraré al requisito de contenido cognitivo como parte del requisito de contenido intencional.

Alcance

Somos capaces de pensar acerca de perros, repollos y licuadoras, pero también acerca de números, democracias o virtudes. Es decir, podemos representar propiedades de distinto tipo, tanto concretas como abstractas. Tenemos una variedad de conceptos que incluye, *inter alia*, conceptos de clases naturales, como AGUA, conceptos de artefactos, como TELEVISOR, conceptos derivados teóricamente, como PROTÓN y conceptos acerca de estados internos, como DOLOR (Prinz 2002). Al postular un requisito de alcance, se tiene en mente que una teoría de conceptos debe poder dar cuenta de nuestra capacidad representacional, es decir, del amplio espectro de conceptos con los que nuestra mente puede operar.

³ Fodor 1994 abandona la noción de contenido estrecho.

Composicionalidad

Nuestro pensamiento exhibe, al menos, dos fenómenos que una teoría de la mente debe explicar. Por un lado, es productivo, esto es, podemos, a partir de un sistema representacional finito, generar una infinita cantidad de pensamientos nuevos. Por otro lado, es sistemático, esto es, si somos capaces de generar ciertos pensamientos, también somos capaces de generar otros relacionados, recomblando sus constituyentes. A partir de un ejemplo de Miller, Margolis y Laurence (2007) ilustran muy claramente la productividad del pensamiento por analogía con la productividad del lenguaje. Si tomamos las oraciones de 20 palabras y consideramos (por estipulación) que al construir una oración, cada palabra es elegida entre un conjunto de 10 palabras, esto permite calcular que somos capaces de generar y entender por lo menos 10^{20} oraciones de 20 palabras. Para darle sentido a este número, Laurence y Margolis señalan que éste equivale aproximadamente a mil millones de veces la cantidad de neuronas del cerebro y a mil veces la cantidad de segundos en la historia del universo. Si suponemos entonces que, aproximadamente, cada oración se corresponde con un pensamiento diferente, la capacidad productiva del pensamiento resulta notable. Este es un fenómeno que requiere explicación en cualquier teoría de la mente y en particular del pensamiento. La sistematicidad también se puede ilustrar por analogía con los lenguajes naturales. Así como ningún hablante del castellano puede entender o producir la oración “El ratón persigue al gato” a no ser que también pueda entender o producir la oración “El gato persigue al ratón”, tampoco es plausible que una persona pueda pensar que el ratón persigue al gato, pero que no pueda pensar que el gato persigue al ratón.

Pues bien, ¿cómo se relacionan estos fenómenos con las teorías de conceptos? Si se asume que los conceptos son los constituyentes de los pensamientos, se puede apelar a la composicionalidad como una explicación de estos dos fenómenos. Fodor define informalmente la composicionalidad afirmando que “... la sintaxis y el contenido de un concepto complejo se determinan normalmente por la sintaxis y el contenido de sus constituyentes” (1999, p. 135). Respecto de la productividad, la representación de una cantidad infinita de conceptos por una mente con un sistema representacional finito se logra si cada concepto complejo resulta finitamente descomponible a la sintaxis y los contenidos de sus constituyentes. El argumento para explicar la sistematicidad es similar. El hecho de

que una mente que puede tener el pensamiento de que el ratón persigue al gato, también puede tener el pensamiento de que el gato persigue al ratón se explica por medio de la composicionalidad porque el contenido de los pensamientos complejos queda determinado de manera composicional por el contenido de sus constituyentes y las reglas de combinación sintáctica.

Publicidad

Según este requisito los conceptos deben ser compartibles, no sólo por distintas personas sino también por un individuo en distintos momentos. Este requisito se postula fundamentalmente por dos razones. En primer lugar, si se pretende que una teoría de conceptos contribuya a una psicología que formule explicaciones intencionales, i.e. que apele a estados mentales para explicar el comportamiento, estas explicaciones deben admitir generalizaciones sobre distintos individuos. Si los conceptos no fueran compartidos, esto no sería posible. Por ejemplo, se puede explicar que Julia agarró una llave y se acercó a la puerta de su casa a partir de su deseo de abrir la puerta y sus creencias de que la puerta estaba cerrada y esa era la llave para abrir su cerradura. El proyecto de una psicología de este tipo pretende extender estas explicaciones a otras personas, sus llaves y las puertas de sus casas. Para esto, es necesario suponer que las actitudes proposicionales de las distintas personas están constituidas por los mismos conceptos, en este caso, instancias del mismo tipo para LLAVE, PUERTA y CERRADURA. En segundo lugar, muchas teorías lingüísticas apelan a los conceptos para explicar la comunicación lingüística, afirmando que asociamos los mismos conceptos a las palabras del lenguaje natural. Si los conceptos no fueran públicos, entonces los pensamientos asociados a las emisiones lingüísticas serían diferentes para cada persona y esto tornaría difícil la comunicación. Puesto que la comunicación de hecho ocurre, esta es una razón para creer que los conceptos son compartidos (Fodor 1999, Prinz 2002).

Adquisición

Según este requisito, una teoría de conceptos debe incluir especificaciones acerca de cómo estos se adquieren. Es decir, debe explicar si los conceptos son innatos o aprendidos y cuáles son los mecanismos de adquisición o aprendizaje. El debate entre las teorías

empiristas e innatistas se da en torno de esta cuestión. Las primeras, por ejemplo la teoría de los proxitipos de Prinz (2002), proponen que la gran mayoría de los conceptos son aprendidos y postulan para esto mecanismos de dominio general. En cambio, las teorías innatistas, como la de Fodor (1984), postulan una base de conceptos innatos y recurren para explicar la adquisición a mecanismos de dominio específico (Margolis & Laurence 2008).

Las teorías empiristas proponen mecanismos de adquisición que, en general, son epistémicos, es decir, se basan en una relación racional con el mundo. Se supone que estos mecanismos son análogos a la formulación y testeo de hipótesis científicas, de modo que se establece una relación de evidencia entre la experiencia y los conceptos. La adquisición de conceptos podría basarse así en mecanismos inductivos de dominio general (Prinz 2002). Según las teorías innatistas, en cambio, la adquisición no es una forma de aprendizaje, sino que es un proceso causal bruto de disparo de conceptos ante los estímulos correspondientes en el mundo.

Categorización

Generalmente, en la literatura sobre el tema, se entiende que la categorización es la capacidad que tenemos de identificar a qué categoría pertenece un objeto. Se diferencia de la representación en sí, puesto que tiene que ver con la aplicación de un concepto a una cosa, y no meramente con nuestra capacidad de representar esa cosa. La categorización es entonces una relación que se establece entre una categoría o un concepto y un elemento del mundo cuando se *juzga* que dicho elemento cae bajo el concepto.⁴ Algunos autores, como Prinz (2002) incluyen también en este requisito la capacidad de anticipar las características que tendrá un objeto por pertenecer a una categoría.⁵

Los estudios psicológicos realizados en torno a tareas de categorización reportan diversos fenómenos que es preciso tener en cuenta al tratar este requisito. Por ejemplo, en algunos estudios se observó que generalmente las instancias más típicas de una categoría se

⁴ Tanto en la literatura filosófica como psicológica el uso del término "categoría" es ambiguo. Algunos autores consideran que "categoría" es sinónimo de "concepto" (Fodor 1999). Otros, en cambio, establecen una diferencia señalando que las categorías son aquello que los conceptos representan (Prinz 2002, Weiskopf 2009b). En esta tesis utilizaré el término de manera indistinta, puesto que considero que no es dañino, dado que en general el contexto aclara cuál es la interpretación adecuada.

⁵ Otros autores, como Murphy (2002) o Machery (2009), consideran que la capacidad de anticipar características debe explicarse como parte de los procesos relacionados con las inferencias inductivas y la incluyen, por lo tanto, dentro de este último requisito.

categorizan más rápido que las menos típicas -efecto de tipicidad (Rosch & Mervis 1975, Hampton 1997). Otros estudios, sin embargo, señalan un efecto inverso, i.e. que instancias atípicas se categorizan más rápido que otras típicas. Al postular un requisito de categorización, se pretende explicar este proceso de un modo que permita dar cuenta de los distintos fenómenos empíricos observados.

Inferencias inductivas y deductivas

Algunos autores, como Murphy (2002) y Machery (2009), consideran que la capacidad para anticipar características que tendrá un elemento si pertenece a determinada categoría forma parte de los procesos inductivos. Así, los procesos inductivos implican o bien la generalización de ciertas características de una categoría a elementos nuevos pertenecientes a la misma -un ejemplo de esto sería mi capacidad para anticipar que si me compro una heladera va a necesitar energía de algún tipo para funcionar-, o bien la generalización de propiedades de una categoría a otra categoría -por ejemplo, podemos inferir que los perros tienen pulmones a partir del hecho de que los gatos los tienen. La inclusión de un requisito como éste en una teoría de conceptos apunta a explicar cómo contribuyen los conceptos a este tipo de procesos. Respecto de las inferencias deductivas, podemos, por ejemplo, inferir a partir de que Ariel está de buen humor si gana Platense y que ayer ganó Platense, que entonces Ariel estará de buen humor. Fodor y Pylyshyn (1988) señalan que lo que permite realizar inferencias deductivas de este tipo es la sensibilidad de estos procesos a la sistematicidad de nuestros pensamientos, i.e. no es posible que una mente pueda inferir P de P&Q, pero no pueda inferir R de P&Q&R.⁶

2. El atomismo conceptual y el neo-empirismo de conceptos

El atomismo conceptual, cuyo principal representante es Jerry Fodor, y el neo-empirismo de conceptos, cuyo principal representante es Jesse Prinz, son las dos teorías actuales que compiten respecto de la explicación adecuada de aspectos de los debates centrales desarrollados en §1 y son las teorías de las cuales me ocuparé en esta tesis. Ambas

⁶ Fodor y Pylyshyn no incluyen al razonamiento deductivo como un requisito explicativo para las teorías de conceptos, pero analizan las condiciones que la arquitectura cognitiva debe satisfacer para poder dar cuenta de este tipo de razonamiento y apelan a la sistematicidad de las inferencias para sostener que los conceptos deben constituir un sistema simbólico como el lenguaje del pensamiento.

teorías proponen respuestas a la cuestión de la ontología de los conceptos, su estructura y los requisitos de una teoría de conceptos.

La teoría de conceptos de Fodor forma parte de su propuesta de una Teoría Representacional/Computacional de la Mente y tiene como meta principal describir cuáles son los “ladrillos” o elementos primitivos metafísicos a partir de los cuales se construyen los pensamientos. Con el objetivo último de contribuir a una psicología científica que explique la conducta en función de actitudes proposicionales tales como deseos y creencias, Fodor ofrece una propuesta de las características de los elementos constituyentes de las proposiciones que los agentes creen, desean, etc., y propone que se trata de símbolos de un lenguaje del pensamiento (Fodor 1984, 1987 y 1994). Prinz comparte con Fodor el objetivo de caracterizar los elementos a partir de los cuales están constituidos los pensamientos, pero se opone a la idea de que los conceptos son meramente *para pensar*, y propone en cambio enfatizar que los conceptos están orientados a la *acción*. Según Prinz (2005), los conceptos están estructurados de un modo que facilita nuestra interacción con el mundo. Así, propone que están estructurados a partir de rasgos perceptivos que nos permiten detectar instancias de categorías en el mundo.

Tanto Fodor como Prinz sostienen que los conceptos son representaciones mentales, de modo que comparten un enfoque general respecto de la ontología de los conceptos. Sin embargo, las características que asignan a estas representaciones son muy diferentes. Ambos realizan una distinción entre las propiedades semánticas de los conceptos -i.e. el contenido-, y las propiedades formales -el vehículo o formato que porta el contenido. Respecto a las propiedades semánticas, ambas teorías sostienen alguna variante de la semántica informacional que establece que el contenido de los conceptos se determina por relaciones nomológicas con el mundo. De modo que la polémica entre estas teorías gira, en particular, en torno a los vehículos de los conceptos. El atomismo conceptual identifica el formato de los conceptos con representaciones amodales, en el sentido de que no guardan ningún tipo de similitud estructural ni comparten recursos neuronales con los estados perceptivos. Se trata de símbolos de un lenguaje del pensamiento que poseen una sintaxis y una semántica combinatoria al estilo de los lenguajes naturales, pero sin ser ninguno de ellos (Fodor 1984, 1994, 1999). La vertiente neo-empirista, por su parte, considera que los conceptos están vehiculizados por un código modal, propio de cada uno de los sistemas

perceptivos, esto es, los conceptos son copias de representaciones perceptivas que pueden ser reactivadas en ausencia de los estímulos que las ocasionaron (Prinz 2002, 2005).

Por otro lado, mientras que desde el punto de vista atomista, los conceptos léxicos son atómicos, i.e., no poseen una estructura semántica interna (Fodor 1999), para el neo-empirismo, estos están internamente estructurados a partir de prototipos, teorías, etc.(Prinz 2002). Así, el atomismo informacional sostiene que las condiciones metafísicas necesarias para tener un concepto no involucran otros conceptos, i.e. no requieren que se satisfagan las condiciones de posesión de otros conceptos. Según el neo-empirismo, en cambio, dado que los conceptos tienen una estructura interna, para poseer un concepto será necesario poseer las partes que conforman esa estructura.

Finalmente, en relación al debate en torno a los requisitos de una teoría de conceptos, Prinz incorpora los ya mencionados en §1, i.e. intencionalidad, contenido cognitivo, alcance, adquisición, categorización, composicionalidad y publicidad, atendiendo así tanto a la metafísica de los conceptos como a su uso en los procesos psicológicos. Fodor, en cambio, en función de su interés por dar cuenta de los conceptos en tanto constituyentes de los pensamientos, se ocupa de los requisitos más relacionados con la metafísica de los conceptos. Incluye así los requisitos de intencionalidad, alcance, adquisición, composicionalidad y publicidad. De esta manera, la diferencia entre ambos está dada por los requisitos de contenido cognitivo y de categorización.

3. Los requisitos explicativos fundamentales de una teoría de conceptos

Dado que uno de los objetivos principales de esta tesis es comparar y evaluar las teorías de Fodor y Prinz en función de la satisfacción de los requisitos que considero fundamentales para toda teoría de conceptos, me ocuparé, por un lado, de los requisitos de intencionalidad, alcance y composicionalidad, que están ligados a las características metafísicas de los conceptos y, por otro lado, de los requisitos de adquisición y categorización que se relacionan con los procesos mentales. Distintos criterios contribuyeron a esta selección de requisitos fundamentales. En primer lugar, incluí tanto requisitos tradicionales de las teorías filosóficas de conceptos como de las teorías psicológicas, puesto que considero que una teoría de conceptos debe responder tanto a

cuestiones relacionadas con la metafísica de los conceptos como a cuestiones psicológicas relacionadas con lo que se hace con ellos.

En particular, el requisito de intencionalidad resulta fundamental si se pretende articular una teoría de conceptos con una explicación del contenido de los estados mentales. Es decir, si se considera a los estados mentales como estados intencionales que tienen un contenido y a los conceptos como los elementos que, en tanto constituyentes de los pensamientos, son los portadores primitivos del contenido intencional, entonces es necesario incluir una explicación de la semántica de los conceptos. De modo que no sólo es necesario explicar qué clase de entidades son los constituyentes de los pensamientos, sino también en qué consiste su contenido y cómo lo adquieren. Al evaluar la explicación del contenido intencional que ofrece cada teoría, consideraré también al contenido cognitivo, puesto que el contenido intencional no resulta suficiente para la individuación de conceptos.

También incluyo los requisitos de alcance y composicionalidad puesto que son los que demandan una explicación de nuestra capacidad representacional. El alcance de una teoría de conceptos equivale a su poder expresivo. Podrían formularse, por ejemplo, teorías que resulten muy adecuadas para explicar los conceptos de objetos perceptibles, pero éstas resultarían ser muy limitadas si fueran incapaces de abarcar otros tipos de conceptos, como los conceptos abstractos y teóricos, ya que gran parte de los conceptos con los que operamos corresponden a estos últimos. Por ejemplo, si hiciéramos una estadística de las palabras escritas en esta tesis, probablemente una buena cantidad esté asociada a conceptos abstractos como RAZONAMIENTO, CONCEPTO o EXPLICACIÓN. Una buena teoría de conceptos contribuirá a la explicación del pensamiento si puede dar cuenta de la amplia variedad de conceptos que podemos tener.

La composicionalidad, a su vez, se propone como la mejor explicación para la productividad y sistematicidad del pensamiento. Estos fenómenos empíricos reflejan la naturaleza del sistema representacional con el que contamos, que, no sólo exhibe relaciones estructurales entre los pensamientos, sino que, siendo finito, nos permite generar una cantidad infinita de pensamientos. La teoría de conceptos que se proponga para explicar la metafísica del pensamiento debe poder dar cuenta de estos fenómenos, de modo que debe satisfacer el requisito de composicionalidad. Por otra parte, dado que la explicación del

razonamiento es uno de los objetivos tanto de la filosofía de la mente como de la psicología cognitiva, el requisito de composicionalidad resulta inevitable puesto que contribuye a la explicación del razonamiento deductivo.

Respecto de los requisitos relacionados con los procesos cognitivos, el requisito de adquisición es necesario, ya que una teoría que postule características ontológicas de los conceptos y cómo estos se utilizan en diversos procesos, resultará insuficiente si no propone alguna hipótesis plausible acerca de cómo nuestra mente llega a equiparse con esas entidades. Por otro lado, el requisito de categorización es fundamental dado que es un proceso que explica gran parte de nuestro comportamiento hacia los objetos con los que interactuamos. Asimismo, hay una amplia gama de efectos observados en torno a este proceso (tales como el efecto de tipicidad), que resultan de central interés en teorías psicológicas sobre la mente. De modo que para articular una teoría de conceptos con las explicaciones psicológicas de estos fenómenos es necesario incluir el requisito de categorización. Además, algunos fenómenos relacionados con la inducción pueden ser abordados a partir de un tratamiento de la categorización, tal como hace Prinz (2002).

Esta selección constituye un conjunto de requisitos fundamentales, pero esto no significa que el resto de los requisitos mencionados en §1 no sean relevantes. Al igual que Weiskopf (2009b) -respecto de su propia lista de *desiderata*-, considero que son las dificultades para satisfacer este conjunto de requisitos fundamentales las que llevan a postular un pluralismo de conceptos, tal como sostendré hacia el final de la tesis. En este sentido, el agregado de otros requisitos no implicaría diferencias sustanciales ni afectaría a los argumentos a favor del pluralismo. Así, el segundo objetivo de esta tesis es ofrecer los lineamientos básicos y generales para una teoría pluralista de conceptos que permita superar las dificultades que enfrentan el atomismo y el neo-empirismo de conceptos.

4. La posibilidad de una teoría pluralista de conceptos

Con el surgimiento de teorías de conceptos pluralistas (Weiskopf 2009a, 2009b) y eliminativistas (Machery 2009), se ha suscitado un debate respecto de si es posible elaborar una teoría unificada de conceptos que dé cuenta de todos los *desiderata* en cuestión. Este debate involucra dos sentidos en los que se puede entender la noción de una “teoría unificada”. En primer lugar, puede referirse a una teoría que explica ciertos fenómenos

apelando a un único tipo de entidad. En este sentido, el debate se da entre las teorías que postulan un único tipo de entidad, tales como las de Fodor y Prinz, y aquellas que postulan un pluralismo de entidades, tales como las de Weiskopf o Machery. En segundo lugar, puede entenderse que “teoría unificada” se refiere a una teoría que explica todos los fenómenos relacionados con los conceptos, pero no necesariamente apelando a un único tipo de entidad. Machery señala que hay evidencia a favor de la existencia de distintas representaciones (ejemplares, prototipos y teorías) que cumplen el rol de los conceptos (i.e., según Machery, ser cuerpos de información almacenados en memoria a largo plazo, utilizados *por default* en la mayor parte de los procesos cognitivos) y hace especial énfasis en los fenómenos y efectos que cada tipo de entidad permite explicar. Muestra así que cada uno resulta adecuado para explicar ciertos fenómenos o procesos cognitivos, pero enfrenta limitaciones a la hora de dar cuenta de otros.

Considero que se debe trabajar en la elaboración de una teoría unificada en el segundo sentido mencionado, mientras que sería infructuoso seguir concentrándose en una teoría unificada en el primer sentido. Reconstruir y evaluar el atomismo conceptual y el neo-empirismo, me permitirá señalar sus ventajas y desventajas, para ofrecer luego una propuesta alternativa que supere sus límites explicativos. En la misma línea que adopta Machery (2009), la evaluación de estas teorías apuntará a mostrar que no pueden explicarse todos los fenómenos relacionados con los conceptos apelando a un único tipo de representación conceptual.

Ahora bien, es necesario aclarar qué entiendo por 'tipos de representaciones'. Dado que los conceptos entendidos como representaciones mentales tienen propiedades tanto semánticas como sintácticas, ¿cuál de estos dos aspectos determinará de qué tipo de entidad se trata? Se podría elegir el formato (perceptivo vs. simbólico) o el tipo de información que transportan (por ejemplo, información teórica vs. prototípica). Tanto la propuesta de Weiskopf (2009b) como la de Machery (2009) se concentran en la segunda opción, i.e. en el tipo de información que los conceptos contienen, y esto podría ser compatible con que hubiera un único formato para todas las representaciones. En la reconstrucción de las propuestas de Fodor y Prinz prestaré especial atención al formato de las representaciones mentales, dado que es el que está centralmente puesto en debate. La evaluación de los límites y ventajas de cada tipo de vehículo o formato permitirá esbozar una propuesta

alternativa que sostenga no sólo un pluralismo semántico -i.e. de tipos de información o contenido-, sino también de vehículos, al que denominaré pluralismo extendido.

5. Plan de la tesis

En los capítulos 1 y 2 me ocuparé de reconstruir las tesis principales del atomismo conceptual y evaluar la satisfacción de cada uno de los requisitos fundamentales propuestos. En el capítulo 1, me ocuparé de reconstruir las tesis respecto de la semántica y los vehículos de los conceptos, y evaluar la satisfacción de los requisitos de intencionalidad, alcance y composicionalidad; requisitos que dependen fundamentalmente de la caracterización metafísica de los conceptos. Señalaré que, si bien el atomismo permite representar un amplio espectro de conceptos y satisfacer el requisito de composicionalidad, tiene sin embargo dificultades con el requisito de intencionalidad a la hora de dar cuenta de la individuación de conceptos coextensivos.

En el capítulo 2 me ocuparé de elucidar cómo se articulan las tesis principales del atomismo con la teoría fodoriana de los procesos de pensamiento conceptual. Incluiré en este capítulo el análisis de los requisitos de adquisición y categorización, puesto que se trata de procesos que es preciso evaluar en función de la explicación que ofrece una teoría acerca de los procesos mentales. Ambos requisitos constituyen importantes desafíos para el atomismo. En primer lugar, Fodor se ve obligado a introducir en el proceso de adquisición de conceptos a los prototipos. Sostendré que la incorporación de otras estructuras conlleva dificultades, ya que la relación entre los prototipos y los conceptos no queda clara. En segundo lugar, tradicionalmente se recurre a la estructura de los conceptos para explicar la categorización, pero el atomismo no plantea ninguna estructura interna, por lo cual tiene que recurrir a elementos externos a los conceptos para explicar este proceso.

De manera análoga, en los capítulos 3 y 4 reconstruiré las tesis principales del neo-empirismo de conceptos. En el capítulo 3 reconstruiré las tesis respecto de la semántica y los vehículos de los conceptos y evaluaré la satisfacción de los requisitos de intencionalidad, alcance y composicionalidad. El neo-empirismo tiene dificultades con estos tres requisitos. Prinz intenta compatibilizar una semántica informacional para dar cuenta del contenido intencional con una estructura híbrida para los conceptos que constituiría el contenido cognitivo. Pero al hacer esto, por un lado, corre el riesgo de que el

contenido intencional quede reducido a contenido cognitivo, ya que este último resulta constitutivo en la determinación del contenido intencional y, por otro lado, tiene dificultades para dar cuenta de la individuación de conceptos. Respecto del requisito de alcance, mostraré que es un desafío dar cuenta de la representación de conceptos teóricos y abstractos recurriendo únicamente a representaciones perceptivas. Por último, Prinz también tiene dificultades con el requisito de composicionalidad, puesto que no sólo debe dar cuenta de la composicionalidad de los prototipos (muy criticada en la literatura), sino también de la composicionalidad de los proxitipos en general. Argumentaré que su defensa de este requisito es incompleta y que los aspectos que deja abiertos, en particular respecto de si las representaciones perceptivas admiten la composicionalidad sintáctica, son especialmente vulnerables a críticas.

En el capítulo 4 intentaré elucidar la articulación de las tesis principales del neo-empirismo con la teoría de los procesos de pensamiento que defiende Prinz. Al igual que para el atomismo, incluiré en este capítulo el análisis de los requisitos de adquisición y categorización. Mostraré que si bien la teoría de proxitipos ofrece explicaciones atractivas de la adquisición y la categorización, ésta tiene problemas. La explicación de la adquisición no resulta adecuada para todo tipo de conceptos, en particular, para los teóricos y abstractos. Por otro lado, no queda claro que las herramientas que se ofrecen para dar cuenta de la categorización sean compatibles con la postulación de representaciones perceptivas.

A lo largo de estos cuatro capítulos intentaré mostrar que ninguna de las dos teorías puede dar cuenta de los requisitos fundamentales para una teoría de conceptos, puesto que cada enfoque prioriza un fenómeno a explicar y sacrifica o debilita otros igualmente importantes. Esto mostraría, al parecer, que no hay una única estructura conceptual que pueda dar cuenta de los fenómenos conceptuales básicos. Sin embargo, cada estructura parece resultar adecuada para explicar algunos de ellos, permitiendo dar cuenta de distintos fenómenos que se observan en los estudios empíricos. A la luz de estas consideraciones, en el capítulo 5 desarrollaré una propuesta alternativa que postula una pluralidad de representaciones mentales de distinto tipo, a la que denominaré pluralismo extendido, puesto que abarca tanto un pluralismo de estructuras semánticas como un pluralismo de vehículos.

Este enfoque pluralista presenta ventajas frente a las posturas teóricas que adhieren a la tesis de la uniformidad -la tesis de que los conceptos conforman una clase uniforme de entidades psicológicas (Weiskopf 2009b)-, puesto que hereda las ventajas explicativas de las distintas teorías. Sin embargo, se debe mostrar que no hereda sus dificultades y que puede, entonces, dar cuenta de los requisitos fundamentales. En este sentido, uno de los mayores desafíos para este enfoque es mostrar que puede acomodar el requisito de intencionalidad, puesto que debe ofrecer una explicación de cómo se articulan los distintos modos de referir que tiene cada tipo de entidad. Por otra parte, el pluralismo presenta también nuevos desafíos. Obliga a revisar la concepción tradicional que trata a los conceptos como entidades singulares, según la cual, a grandes rasgos, hay *un* concepto asociado a cada ítem léxico, es decir, que se tiene, por ejemplo, *el* concepto GATO, *el* concepto BALCÓN -tesis de la singularidad (Weiskopf 2009b). Así, es necesario evaluar qué consecuencias acarrea el abandono de esta tesis, en especial, cómo se individualarán los conceptos y cómo se articularán las distintas estructuras en los procesos mentales conceptuales. Al ofrecer los lineamientos básicos de la propuesta del pluralismo extendido sugeriré posibles esbozos de respuestas a estos desafíos.

Para facilitar la lectura, a continuación presento las abreviaturas utilizadas en la tesis:

AP: actitud proposicional

LDP: lenguaje del pensamiento

MAS: mecanismos de acceso semántico

MDP: modo de presentación

RCA: modelo de recuperación, composición y análisis

RM: representación mental

RP: representación perceptiva

TRCM: teoría representacional/computacional de la mente

TREP: teoría de la reactivación de estados perceptivos

Capítulo 1

El atomismo conceptual.

La satisfacción de los requisitos de intencionalidad, alcance y composicionalidad

En este capítulo desarrollaré las tesis principales del atomismo informacional respecto de la metafísica de los conceptos y evaluaré la satisfacción de los requisitos de intencionalidad, alcance y composicionalidad. Mi objetivo será mostrar que, si bien el atomismo conceptual satisface los requisitos de alcance y composicionalidad, tiene inconvenientes respecto del contenido intencional en relación a la individuación de conceptos.

1. Características de los conceptos

La propuesta del atomismo conceptual puede sintetizarse en tres tesis principales, las cuales desarrollaré en orden luego de su presentación:

Lenguaje del pensamiento: los conceptos son representaciones mentales que tienen una semántica composicional y una sintaxis combinatoria.⁷

Semántica informacional: el contenido de los conceptos se determina por medio de relaciones nomológicas con propiedades en el mundo.

Atomismo: la mayoría de los conceptos léxicos son atómicos.

Lenguaje del pensamiento

La hipótesis del lenguaje del pensamiento (LDP) o mentalés forma parte de la Teoría Representacional/Computacional de la Mente (TRCM) a la que Fodor adscribe, según la cual el pensamiento está vehiculizado por un sistema interno de representaciones sobre las cuales operan causalmente los procesos computacionales. Me ocuparé de la TRCM en el capítulo 2. A los fines de comprender la hipótesis del LDP basta con decir que, según esta teoría, los estados mentales de actitud proposicional (AP) (tales como deseos y creencias, y en general, pensamientos) consisten en relaciones (computacionales)

⁷ La hipótesis del lenguaje del pensamiento está formulada para los constituyentes de los pensamientos, i.e. los conceptos. Sin embargo, cabe aclarar que si bien los conceptos son representaciones mentales, no toda representación mental es conceptual (Fodor 2008).

entre organismos y representaciones mentales (RM). Por ejemplo, mi creencia de que llueve implica que estoy en un tipo particular de relación o actitud (en este caso, la de creencia) con una RM que tiene por contenido que *llueve*.

Los contenidos de las AP están conformados así por un sistema interno de representaciones que es análogo a los lenguajes naturales. Se trata de un lenguaje compuesto por símbolos que se combinan para formar representaciones complejas según las reglas de combinación sintáctica y de composicionalidad semántica que rigen para los lenguajes naturales, aunque el mentalés no es ninguno de ellos (Fodor 1984, 1987). Este lenguaje, al igual que los lenguajes naturales admite, por un lado, una distinción entre sujeto y predicado y, por otro lado, ciertos operadores lógicos, como los cuantificadores y las variables. Pues bien, los conceptos son los símbolos de este LDP. Así, el atomismo informacional sostiene que los conceptos son RM, i.e., son particulares mentales que están físicamente instanciados –tienen una realización física a nivel cerebral- y poseen una estructura simbólica semejante a la de un lenguaje natural. De modo que estas representaciones poseen propiedades sintácticas y semánticas. El aspecto sintáctico de los conceptos consiste en la forma, vehículo o formato que transporta la información semántica, mientras que el aspecto semántico consiste en el contenido de los conceptos, i.e. aquello en el mundo (estados, eventos o propiedades) que los conceptos representan.

Las propiedades sintácticas son propiedades que tienen los conceptos en virtud de su estructura interna. Fodor (1987) explica las propiedades sintácticas por analogía a las formas de las llaves y las cerraduras: la forma geométrica de una llave determina qué cerradura podrá abrir.⁸ De manera similar, la estructura sintáctica de un concepto determina sus poderes causales, puesto que determina qué transiciones o transformaciones se podrán realizar sobre este símbolo. Por ejemplo, un símbolo con la estructura P&Q admite una transformación que resulte en Q puramente en virtud de su forma sintáctica. Por otra parte, las diferencias sintácticas pueden pensarse como diferencias en la *ortografía mental* (Margolis & Laurence 2007). En función de su estructura sintáctica, se pueden diferenciar representaciones atómicas y moleculares. Las representaciones moleculares son las que tienen una estructura de *constituyentes* sintácticos, que pueden ser a su vez atómicos o

⁸ Fodor (1987, p. 18) afirma que “la sintaxis de un símbolo es una de sus propiedades físicas de orden superior. Como aproximación metafórica, podemos pensar en la estructura sintáctica de un símbolo como un rasgo abstracto de su forma.”

moleculares (Fodor & Pylyshyn 1988). Así AUTO, por ejemplo, es una representación atómica, mientras que EL AUTO FANTÁSTICO es una representación molecular, que tiene como constituyentes atómicos a EL, AUTO y FANTÁSTICO.⁹

Las propiedades semánticas se relacionan con el contenido de los conceptos, es decir, consisten en *acerca de* qué son esas representaciones. Según la TRCM, los estados mentales de AP tienen propiedades semánticas o condiciones de satisfacción, en función de su relación con el mundo. Por ejemplo, si creo que la Tierra es plana, el contenido de mi creencia -i.e., la Tierra es plana- representa un hecho o estado del mundo y es éste el que permite evaluar si mi creencia es verdadera o falsa. Dado que el contenido de estos estados mentales se compone a partir de la combinación de los contenidos de sus constituyentes (en este caso, TIERRA, ES y PLANA) los portadores primitivos del contenido intencional son los conceptos.

Para dar cuenta de cómo los conceptos adquieren las propiedades semánticas, el atomismo informacional adopta la semántica informacional, la cual reconstruiré a continuación.

Semántica informacional

Respecto al contenido intencional, Fodor sostiene, siguiendo a Dretske (1981), que el contenido de los conceptos consiste en información acerca de propiedades en el mundo. Una representación R porta información acerca de P si las instancias de P son “nómicamente responsables” por las instancias de R (Fodor 1990a). Así, los conceptos adquieren su contenido en virtud de la existencia de leyes psicofísicas causales (sustentadas por contrafácticos) que relacionan instancias de nuestras representaciones mentales con instancias de propiedades en el mundo. Por ejemplo, las instancias del concepto LIEBRE caen bajo una ley según la cual las instancias de este concepto son causadas por (o serían efectos de) la propiedad de *ser liebre* instanciada en el mundo. Esta ley está sustentada por contrafácticos del tipo: Si se instanciara la propiedad de *ser liebre*, se instanciaría el concepto LIEBRE. Estas relaciones nómicas son las “hacedoras de significado”. De esta manera, Fodor adhiere a una teoría causal de la representación (aunque no apela a cadenas

⁹ En esta tesis, los nombres de los conceptos están en mayúsculas y los nombres de los valores semánticos de los conceptos están en itálicas.

causales actuales, sino contrafácticas), según la cual las condiciones para la posesión de conceptos no son epistémicas, sino puramente nomológicas: no se trata de “saber algo”, sino de estar enlazado nómicamente con la propiedad adecuada. En consecuencia, tener un concepto resulta semejante a tener una etiqueta que está relacionada con una propiedad determinada. Por medio de estas etiquetas es que pensamos acerca de cosas, relaciones, etc.

Por otra parte, Fodor (1999) introduce una distinción entre distintos tipos de propiedades con las cuales nuestras mentes se enlazan: las *propiedades de clase natural*, por un lado, y las *propiedades de apariencia*, por otro lado. Las propiedades de clase natural son propiedades esenciales, intrínsecas, mientras que las de apariencia son propiedades que las cosas tienen en virtud de los estados mentales que evocan en mentes como las nuestras. Por ejemplo, en el caso de AGUA, la posesión de este concepto implica estar enlazado con la propiedad esencial de la sustancia agua, i.e. H_2O , por lo cual *ser agua* es una propiedad de clase natural. En cambio, en el caso de PICAPORTE, por ejemplo, no hay propiamente leyes acerca de picaportes *qua* picaportes, sino que PICAPORTE cae bajo leyes acerca de nuestra mente, puesto que *ser picaporte* es una propiedad de apariencia. No hay ninguna esencia oculta, ni propiedad intrínseca de los picaportes (tal como una microestructura) en virtud de la cual puedan caer bajo una ley que conecte a los picaportes, *qua* picaportes, con el concepto PICAPORTE, sino que *ser picaporte* es una propiedad que se define a partir de cómo ciertas cosas en el mundo resuenan en nuestras mentes, es decir, a partir de los estados mentales que despiertan en nosotros. Tanto en el caso de los conceptos de clase natural como en el de los conceptos que refieren a propiedades de apariencia, su contenido se explica por relaciones nomológicas con propiedades. De esta manera, Fodor hace extensiva la semántica informacional a la mayoría de los conceptos léxicos (ya sea que representen propiedades de clase natural o de apariencia), excluyendo de esta explicación únicamente a los conceptos lógico-matemáticos (Fodor 1999).

Se ha objetado a la semántica informacional que las relaciones causales entre el mundo y los conceptos no son suficientes para determinar el contenido, puesto que puede ocurrir que un concepto se dispare frente a cosas que no pertenecen a su extensión. Es decir, existe la posibilidad de que, por ejemplo, se dispare el concepto MARIPOSA frente a una polilla, y según una teoría semántica puramente causal no se podría decir que este es un caso de representación errónea, puesto que habría que decir que lo que sea que causa la

instanciación de MARIPOSA, pertenece a la extensión de MARIPOSA y así no habría error posible. Esto, además, da lugar al llamado *problema de la disyunción* (Fodor 1990a, Prinz 2002, Laurence & Margolis 1999): si un concepto representa aquello que causa su instanciación y este concepto puede tener “causas fortuitas”, entonces siguiendo a la semántica informacional habría que decir que este concepto está nómicamente relacionado con más de una propiedad, es decir, que su contenido es una disyunción. En el caso del ejemplo en el cual una polilla causa MARIPOSA, habría que decir que el contenido de MARIPOSA es *polilla* o *mariposa*. La solución que ofrece Fodor (1990b, 1990c) a este problema consiste en una dependencia asimétrica.

Esta propuesta especifica dos condiciones que se deben cumplir para que X tenga a *x* como contenido:

1. Es una ley que instancias de *x* causan instancias de X.
2. Para todo *y* (distinto de *x*), si *y* *qua* *y* causa de hecho instancias de X, esto depende asimétricamente de que instancias de *x* causen instancias de X.¹⁰

La primera afirmación prácticamente enuncia la tesis central de la semántica informacional. La idea es que hay una ley -e.g. 'Los topos causan TOPOS' o 'Las vizcachas causan VIZCACHAS'- que está sustentada por contrafácticos del tipo 'Si se estuviera en presencia de topos, estos causarían pensamientos acerca de topos'. Sin embargo, esto deja abierto el problema de la disyunción. De acuerdo con la condición 1, X porta información acerca de *x*, pero no se asegura que sólo acerca de *x*. La segunda condición especifica la asimetría. Esta condición establece que las instancias de *y* no podrían causar instancias de X si las instancias de *x* no causaran instancias de X.¹¹ En otras palabras, no podríamos pensar en instancias de *y* como X si no pudiéramos pensar en instancias de *x* como X (Fodor 2008). Si en determinadas circunstancias, por ejemplo en una noche oscura, alguien piensa TOPO frente a un castor, la relación *castor*-TOPO depende asimétricamente de la

¹⁰ Fodor (1990b) agrega una tercera condición según la cual debe haber, de hecho, una relación causal entre algunos *x* y X, pero luego abandona esta condición para adoptar una semántica informacional puramente contrafáctica (Fodor 1994).

¹¹ Esta condición debe entenderse en términos sincrónicos (Fodor 1987). La ley que conecta a *x* con X depende asimétricamente de la que conecta a *y* con X, de tal modo que si se interrumpe la conexión *x*-X en un tiempo *t*, también se interrumpen todas las conexiones *y*-X en *t* (Adams & Aizawa, en prensa).

ley que conecta las instancias de TOPO con las instancias de la propiedad *ser topo*. De esta manera puede hacerse frente al problema de la disyunción, puesto que, por la conjunción de las dos condiciones, todas las conexiones que dependen asimétricamente de otras para causar instancias de X quedan excluidas del contenido de X.

Atomismo

Contra lo que sostienen muchas teorías de conceptos (e.g. teorías de prototipos o definicionales), Fodor afirma que la gran mayoría de los conceptos léxicos son atómicos, i.e. no tienen estructura interna.¹² Para la mayoría de los conceptos léxicos, la relación nómica que determina su contenido no depende de manera constitutiva de ninguna otra relación que mantengan con otros conceptos y, por lo tanto, el contenido de estos conceptos es atómico (Fodor 1999, pp. 32-33). Es decir que para poseer, por ejemplo, el concepto GATO, no es necesario estar en relación con ninguna propiedad distinta de la de *ser gato*. Para entrar en relación (o “enlazarse”) con la propiedad de *ser gato*, las mentes como las nuestras necesitan estar en un entorno determinado, i.e un entorno en el que esté la propiedad de *ser gato*. Pero, en un entorno así, también se está en contacto con otras propiedades relacionadas con la de *ser gato*, de modo que puede ocurrir que junto con GATO se adquieran también otros conceptos relacionados. Sin embargo, resulta contingente que se pueda adquirir GATO al mismo tiempo que OREJA, PATA o BIGOTE, de manera que estos conceptos no son constitutivos de GATO. Así, una persona podría tener el concepto GATO sin tener, por ejemplo, el concepto BIGOTE.

Fodor (1999) argumenta extensamente en contra de las propuestas que asignan estructura interna a los conceptos. Tal como las reconstruye, tanto las teorías que proponen conceptos estructurados a partir de rasgos definicionales (condiciones necesarias y/o suficientes para caer bajo un concepto), cuanto las que proponen rasgos no definicionales (rasgos estadísticos, como en el caso de los prototipos) adoptan alguna forma de teoría semántica de rol inferencial. Estas teorías sostienen que la determinación e individuación del contenido de los conceptos primitivos está, al menos parcialmente, constituida por relaciones inferenciales, esto es, los conceptos se determinan e individualizan por las

¹² A grandes rasgos, los conceptos léxicos se corresponden con palabras (como opuestas a frases) en el lenguaje natural.

inferencias en las que participan. En el caso de los conceptos entendidos como definiciones, las inferencias constitutivas son modales, es decir, cada rasgo de la definición constituye una condición necesaria y el conjunto de todos los rasgos una condición suficiente para que algo caiga bajo el concepto. En el caso de los conceptos entendidos como prototipos o estereotipos (Fodor utiliza estos términos de manera indistinta), las inferencias constitutivas son estadísticas: tener un concepto es conocer qué propiedades o características exhiben típicamente las cosas a las que éste se aplica.

Fodor rechaza las teorías definicionales por la falta de realidad psicológica de las definiciones. A primera vista, si se piensa en ejemplos como SOLTERO, que puede definirse como hombre no casado, puede parecer intuitivo pensar que los conceptos tienen definiciones. Pero un análisis más profundo revela que esto no es posible para la mayoría de los conceptos. Si se quisiera definir, por ejemplo, un concepto bastante concreto como MESA, no resulta para nada evidente cómo definirlo de manera que, por un lado, sea distinto de ESCRITORIO o MESADA y que, por otra parte, no excluya mesas de una sola pata ni tampoco mesas en forma de cubo que no tienen patas. Es decir, no podemos especificar un conjunto de condiciones necesarias y suficientes para que algo caiga bajo el concepto MESA. Y si no es posible especificar ese conjunto para la mayoría de los conceptos, es decir, si la mayoría de los conceptos no tienen definiciones, no es posible sostener que los conceptos son definiciones.

Contra las estructuras de rasgos prototípicos, Fodor argumenta que éstas no se combinan de manera composicional, puesto que los prototipos compuestos no heredan los rasgos prototípicos de los prototipos que los componen. Fodor sostiene que un ejemplar prototípico de MANZANA PÚRPURA, por ejemplo, no hereda los rasgos prototípicos de MANZANA puesto que está entre sus rasgos prototípicos el ser roja. Dado que parte del contenido prototípico de MANZANA no forma parte del contenido de MANZANA PÚRPURA, dicho contenido no puede ser constitutivo de MANZANA. Por lo tanto, los rasgos prototípicos no pueden ser constitutivos del contenido de los conceptos. Otro ejemplo que ofrece Fodor en contra de la identificación de conceptos con prototipos es el de los conceptos complejos booleanos como NO GATO. Este concepto no tiene un prototipo: un perro es tan buen ejemplar de NO GATO como una chaucha, y un perro y una chaucha no tienen nada en común *qua* ejemplares de NO GATO más que no ser gatos. Por

lo tanto, concluye Fodor, los conceptos no pueden ser prototipos (Fodor 1999, Fodor & Lepore 1996, Connolly, Fodor, Gleitman, & Gleitman, 2007).

Fodor ofrece, a cambio, la propuesta informacional atomista, según la cual los conceptos son símbolos de un LDP que representan propiedades en el mundo en virtud de relaciones nómicas entre esas propiedades y los conceptos. A continuación analizaré si a partir de esta propuesta se puede dar cuenta de los requisitos de alcance, composicionalidad e intencionalidad.

2. Satisfacción de los requisitos

2.1 Alcance

Como mencioné en §1, Fodor no propone la misma teoría semántica para todos los tipos de conceptos, sino que establece una distinción entre conceptos, como AGUA, que refieren a propiedades de clase natural, conceptos como MESA que refieren a propiedades de apariencia, y conceptos como CONJUNCIÓN que son conceptos lógicos. Así, a partir de las consideraciones de Fodor (1990b, 1994, 1999) se puede elaborar una taxonomía, distinguiendo tres tipos de conceptos: conceptos que refieren a propiedades de clase natural, conceptos que refieren a propiedades de apariencia y conceptos lógicos y matemáticos.

Los conceptos de clase natural refieren a propiedades en el mundo en virtud de leyes psicofísicas que conectan a los conceptos con propiedades esenciales. Así, por ejemplo AGUA refiere a la propiedad H_2O , y el concepto teórico ELECTRÓN refiere a la propiedad de *ser partícula subatómica con carga negativa*. Los conceptos que refieren a propiedades de apariencia, por otra parte, lo hacen en virtud de leyes psicofísicas causales acerca de cómo esas propiedades resuenan en nuestras mentes. Por ejemplo, LICUADORA no es un concepto de clase natural puesto que no hay una microestructura o esencia oculta de las licuadoras, pero todas las licuadoras tienen en común la propiedad de ser causas de LICUADORA en mentes como las nuestras.

De esta manera, el atomismo informacional explica satisfactoriamente el contenido de la gran mayoría de los conceptos y no tiene mayores problemas para dar cuenta del contenido de conceptos abstractos, tales como VIRTUD (a diferencia de otras teorías como la de proxitipos, que se verá en el capítulo 3). El contenido de este concepto estará

determinado por una relación nómica entre el concepto y la propiedad que causaría su instanciación (Fodor 1990b). Es decir, hay una propiedad abstracta que, al instanciarse bajo ciertas condiciones físicas, causa la activación de VIRTUD en mentes como las nuestras. Y, puesto que *virtud* debe ser una propiedad de apariencia (aunque Fodor no lo señale explícitamente), ésta quedaría definida como aquella que tiene la característica de causar en nuestras mentes el concepto VIRTUD. Así, tampoco es necesario que Fodor especifique cómo es que se instancia en el mundo esta propiedad abstracta y difícil de caracterizar. No importa que distintos mecanismos puedan instanciar la virtud (podría ser una serie de acciones de cierto tipo, por ejemplo) y tampoco es necesario precisarlos. En esta explicación lo importantes es que, si se instanciara la propiedad de la *virtud*, se activaría VIRTUD en una mente que esté cerca -y cuán cerca es algo que la psicofísica debe especificar.

En relación a los conceptos lógicos y matemáticos, Fodor no se pronuncia explícitamente a favor de ninguna teoría semántica particular, pero sostiene que, por ejemplo, en el caso de CONJUNCIÓN, este concepto está definido por su tabla de verdad (Fodor 2008). Supongo que, a tono con el espíritu de este comentario, la manera adecuada de dar cuenta de este tipo de conceptos en la teoría fodoriana sería por medio de definiciones, puesto que para estos conceptos es posible especificar el conjunto de condiciones necesarias y suficientes para su posesión y aplicación.¹³

De esta manera, la semántica informacional “aumentada” con definiciones para un grupo reducido de conceptos (únicamente para los lógicos y matemáticos) puede dar cuenta de la amplia variedad de conceptos con los que pensamos.

2.2 Composicionalidad

El requisito de composicionalidad establece que el contenido de una representación compleja debe quedar determinado en función del contenido de las representaciones atómicas que la componen y las reglas de combinación sintáctica. Es decir que la composicionalidad requiere una sintaxis y semántica combinatoria (Fodor & Pylyshyn 1988). Esto es lo que permite explicar la productividad y sistematicidad del pensamiento.

¹³ Fodor (1990b) adhería a una semántica del rol conceptual para este tipo de conceptos, pero en *LOT 2* (p. 199) rechaza esto explícitamente.

El atomismo informacional da cuenta de este requisito apelando a la estructura en *constituyentes* de las representaciones mentales. Según Fodor (2008), las representaciones mentales simbólicas admiten una estructura en constituyentes sintácticos, que se diferencia de la mera estructura de *partes*. En una estructura compleja que admite sólo partes, cada parte realiza la misma contribución a la estructura total y no hay una forma canónica de recortar o delimitar esas partes. Las representaciones icónicas son, según Fodor, el ejemplo paradigmático de una estructura que sólo admite partes, ya que de cualquier manera que se divida una representación icónica, por ejemplo una foto de un payaso, se obtendrá una parte de una foto de un payaso.

En una estructura en constituyentes, en cambio, los constituyentes sintácticos se diferencian de las partes en el sentido de que no contribuyen todos de la misma manera a la representación compleja, sino que admiten, entre otras, relaciones de tipo predicativo. Los ejemplos paradigmáticos de este tipo de estructuras, a las que Fodor llama representaciones discursivas, son las proposiciones. Por ejemplo, en la proposición: 'El ratón asustó al elefante', 'el ratón' y 'asustó al elefante' son dos constituyentes moleculares, que se dividen a su vez en los constituyentes atómicos 'el' y 'ratón', y 'asustó' y 'al elefante', respectivamente. En cambio, 'ratón asustó' o 'ratón ... elefante' son partes de la proposición, no son constituyentes, puesto que no tienen un análisis sintáctico ni una interpretación semántica. De manera que no cualquier recorte que se haga de esta proposición delimita los constituyentes, sino que hay una estructura canónica que se delimita en relación con la contribución semántica y sintáctica que hace cada representación. Pues bien, el mentalés exhibe, por definición, este tipo de estructura, puesto que sus símbolos admiten, por un lado, relaciones sintácticas, y por otro, composicionalidad semántica. Es decir que admiten, al igual que los símbolos de los lenguajes naturales, formas lógicas que se reflejan en la sintaxis, así como una relación semántica de composicionalidad entre las representaciones atómicas que dan lugar a las representaciones moleculares.

Así, el atomismo informacional puede explicar cómo podemos generar (o tener) una cantidad infinita de pensamientos a partir de un conjunto finito de representaciones -i.e. la productividad del pensamiento-, por medio de la composición sintáctica y semántica de representaciones atómicas para formar representaciones complejas. Y permite también dar cuenta de la sistematicidad, i.e. el hecho de que haya una relación entre nuestra capacidad

para pensar, por ejemplo, que el ratón asustó al elefante y para pensar que el elefante asustó al ratón. No podemos tener alguno de estos pensamientos sin tener la capacidad para pensar el otro. La sistematicidad es así una relación estructural entre pensamientos y, en consecuencia, el hecho de que ambas representaciones complejas del ejemplo estén compuestas por los mismos constituyentes (junto con reglas de combinación sintáctica) permite explicar por qué no podría darse el caso de que una mente pudiera tener uno de los pensamientos, pero no ser capaz de tener el otro.

De esta manera, la satisfacción del requisito de composicionalidad es otra de las grandes ventajas del atomismo informacional.

2.3 Intencionalidad

El atomismo informacional resulta especialmente atractivo por su explicación de la intencionalidad de los conceptos. Tal como señalan Laurence y Margolis (1999), es una ventaja proponer que las únicas condiciones de posesión de un concepto consisten en relaciones nomológicas con el mundo y que puede haber distintos mecanismos que permitan que se realicen estas relaciones. Al no considerar a los factores epistémicos como constitutivos de los conceptos, se puede dar cuenta de la posesión de un concepto independientemente de que se tengan creencias erróneas respecto de la propiedad que este concepto expresa o se ignoren características fundamentales de esa propiedad. Laurence y Margolis toman a las enfermedades como ejemplo para ilustrar esta ventaja. En otra época se creía que la viruela era causada por espíritus malignos o “mala sangre”. Hoy decimos que en esa época estaban equivocados acerca de las causas de la enfermedad o que ignoraban características centrales tales como la forma de transmisión de la enfermedad, pero suponemos que, a pesar de las creencias erróneas, su concepto VIRUELA refería a la misma propiedad. La ventaja fundamental del atomismo informacional consiste en separar a las creencias de la representación, al postular que la individuación del contenido de las representaciones no depende de las creencias de los individuos, sino de relaciones puramente nómicas con el mundo.

A pesar de esta clara ventaja y de adoptar él mismo una forma de semántica informacional, Prinz (2002) señala una serie de objeciones a la versión fodoriana y propone ciertas modificaciones principalmente para resolver el problema de la disyunción. Estas

objecciones no apuntan a modificaciones sustanciales en la teoría semántica, que es, al parecer, la mejor teoría disponible para dar cuenta de la intencionalidad de los conceptos. Prinz critica que la condición de la dependencia asimétrica (que se vio en §1) no resulta suficiente para resolver los casos de “gemelos en esta tierra” -i.e. casos de propiedades indistinguibles, como *ser olmo* y *ser haya*, que darían lugar a un contenido disyuntivo. Sin embargo, considero que esta crítica pasa por alto el tratamiento que ofrece Fodor de estos casos en *El olmo y el experto*, donde sostiene que estos conceptos son deferentes, en el sentido en que las capacidades discriminativas de los expertos en las propiedades en cuestión (i.e. expertos en olmos y hayas) permiten individualarlos, sin necesidad de apelar a un contenido disyuntivo.¹⁴

Considero, sin embargo, que la semántica informacional atomista acarrea dificultades a la hora de individualar conceptos coextensivos, puesto que, en este caso, no se puede recurrir a relaciones nómicas mente-mundo para distinguirlos ya que la propiedad en cuestión es la misma, ni tampoco a factores epistémicos puesto que estos no son constitutivos de los conceptos. Desarrollaré esta dificultad a continuación.¹⁵

Muchos autores, por ejemplo Prinz (2002), señalan que el contenido intencional no es suficiente para la individuación de conceptos, puesto que deja sin explicar los llamados casos Frege -los casos en los cuales se tienen dos conceptos distintos con el mismo contenido intencional, es decir, dos conceptos que tienen distintos poderes causales a pesar de referir a la misma propiedad- y los casos de tierra gemela -los casos en los cuales un concepto tiene distintos contenidos intencionales en el mundo actual y en otro mundo posible.¹⁶ Fodor admite que puesto que se dan casos de conceptos atómicos coextensivos – i.e. que tienen el mismo contenido intencional-, la relación nómica con el mundo no es suficiente para su individuación. Los conceptos coextensivos están relacionados

¹⁴ Prinz también recurre a la noción de conceptos deferentes, como veremos en el capítulo 3, para explicar el contenido de los conceptos teóricos.

¹⁵ Formulé parte de esta crítica en Haimovici (2009a).

¹⁶ No me ocuparé de los casos de tierra gemela, aunque vale la pena señalar que el tratamiento de Fodor (1994) de estos casos resulta muy hábil y permite descartarlos velozmente. Fodor señala, en primer lugar, que ciertos casos de tierra gemela son nomológicamente imposibles. Por ejemplo, por los conocimientos de química disponibles, la única propiedad que tiene las características por medio de las cuales nos enlazamos con AGUA es H_2O y no XYZ (la sustancia de la que estaría compuesta el agua en la tierra gemela). Aunque esto no descarta todos los casos de tierra gemela sino sólo los del estilo H_2O/XYZ , Fodor propone tratar al resto como accidentes y considero que éste es un tratamiento adecuado.

nómicamente con la misma propiedad en el mundo y tienen, por lo tanto, el mismo contenido. ¿Qué es, entonces, lo que los diferencia?

Fodor (1999) afirma que la diferencia radica en una cuestión sintáctica. Dos RM con el mismo contenido que están instanciadas de formas sintácticamente distintas son dos conceptos distintos. De esta manera, la particularidad de la explicación fodoriana radica en que no apela a otro tipo de contenido –como el contenido cognitivo– para explicar estos casos, sino que sostiene que son los vehículos los responsables de una individuación de grano más fino que la que permite el puro contenido intencional. Es suficiente para que dos conceptos sean distintos que tengan distinto vehículo –esto, es, que sean símbolos distintos del LDP– o que tengan distinto contenido intencional. Así la individuación de conceptos se da tanto a partir de la semántica (o contenido) como de la sintaxis (o forma).

Este modo de individuación requiere elucidación. A continuación ofreceré una manera de entender, en la teoría fodoriana, en qué consiste que dos RM con el mismo contenido puedan instanciarse de formas sintácticamente distintas. Para esto abordaré dos cuestiones, cuya relación, a pesar de no ser explícita en los textos fodorianos, exhibe las dificultades que la presente teoría tiene para la individuación de conceptos coextensivos. En particular, me ocuparé de los Modos de Presentación (MDP) y de la determinación del contenido de los conceptos que covarían con propiedades de apariencia.

Frege (1892) apela, para la distinción de conceptos coextensivos, a sus MDP. Estos son sentidos –i.e. objetos abstractos– y corresponden a la manera en que los objetos se presentan al pensamiento o, en términos fodorianos, a cómo las cosas nos impactan. Cada MDP sirve como un modo de presentar un objeto al pensamiento. Pero, dado que los MDP son objetos abstractos que la mente capta, según Fodor (1999) debería ofrecerse un argumento independiente para mostrar que existe una única manera de captar cada MDP.¹⁷ Si hubiera una proliferación de maneras de captar cada MDP, esto resultaría o bien en una proliferación de conceptos (uno por cada manera de captar cada MDP) o bien en una asociación de un concepto con varias maneras distintas de captar un MDP. Si sucediera esto último, los MDP fallarían en la individuación de los conceptos.

Contra la estrategia fregeana, Fodor argumenta que si se tiene como objetivo que los MDP puedan ser causas o efectos de procesos mentales y así funcionar como aquello que

¹⁷ En la teoría fregeana este hecho se establece por mera estipulación.

permite distinguir (en función de sus consecuencias causales) entre las actitudes proposicionales de Edipo como Deseo casarme con Yocasta y Deseo casarme con mamá, entonces los MDP deben estar “en la cabeza” -i.e. deben ser particulares mentales físicamente instanciados. Esto es así porque lo único que puede jugar un rol causal en los procesos mentales son las entidades físicamente instanciadas. Fodor (1987, 1994) propone entonces analizar las AP como relaciones entre tres elementos: un individuo, una proposición -i.e. un determinado contenido- y un MDP -i.e. un vehículo que es una oración del mentalés. La propuesta de Fodor consiste, de este modo, en “materializar” los MDP. Así, los MDP están físicamente instanciados en la cabeza, esto es, son particulares mentales que son los vehículos sintácticos de los conceptos. De esta manera, lo que distingue a los conceptos coextensivos es su MDP o vehículo (Fodor 1994, 1999). Decir que dos conceptos distintos son coextensivos significa que el mismo contenido intencional (una propiedad en el mundo) es transportado por dos MDP o vehículos sintácticos distintos, es decir que la misma propiedad está asociada nómicamente con distintos símbolos del LDP.

Ahora bien, si los MDP son las maneras en que concebimos las cosas y estos están en la cabeza en tanto vehículos sintácticos de los conceptos, entonces, por ejemplo, la sintaxis de YOCASTA equivale a la manera en que la mente es impactada por la propiedad instanciada en el mundo que causa el contenido de YOCASTA. Y la sintaxis de MAMÁ equivale a la manera de la mente de ser impactada por aquello en el mundo que causa el contenido de MAMÁ.¹⁸ Pero, el problema es que la propiedad instanciada en el mundo que es causa de YOCASTA y MAMÁ es la misma.¹⁹ Resta explicar, entonces, cómo a partir de experiencias con una misma propiedad pueden adquirirse conceptos distintos y por qué distintos vehículos covarían con instancias de la misma propiedad. La clave para esta

¹⁸ Me referiré siempre a MAMÁ, en lugar de MI MAMÁ o LA MADRE DE EDIPO, dado que pretendo enfocar el problema desde el punto de vista de Edipo, para el cual MAMÁ es sin dudas un concepto léxico atómico que funciona como un nombre. Tomar MI MAMÁ o LA MADRE DE EDIPO daría lugar a lecturas distintas, si se los considerara conceptos complejos en lugar de atómicos. Por otra parte, Fodor (2008, p. 65) señala que si bien podría entenderse a los nombres como descripciones abreviadas (en cuyo caso no habría problemas para la individuación), por los menos algunos nombres propios no son descripciones, sino conceptos atómicos. De esta manera, el ejemplo de YOCASTA y MAMÁ, puede tratarse como un caso de conceptos atómicos coextensivos.

¹⁹ En este ejemplo, la propiedad instanciada que causa ambos conceptos es la propiedad de *ser determinado individuo*. Podría hablarse directamente de *tal individuo* como el contenido con el cual se enlazan los conceptos, pero dado que Fodor habla siempre de “propiedades” mantendré la terminología. De todas maneras, considero que las consecuencias señaladas se siguen tanto para propiedades como para individuos.

explicación podría estar en los mecanismos de adquisición de conceptos.²⁰ Dados los mecanismos que Fodor propone, debe explicarse cómo una misma propiedad causa dos conceptos atómicos distintos, i.e., dos MDP distintos. Considero, como intentaré mostrar a continuación, que esto conduce a un dilema que la teoría de Fodor no puede resolver sin renunciar a algún aspecto de su metafísica de los conceptos.

Si dos conceptos son coextensivos, esto significa que portan el mismo contenido. En el caso de Edipo, YOCASTA y MAMÁ portan la misma información acerca de un individuo. Si se los diferencia por su sintaxis, hace falta explicar cómo es que la información que portan se vincula con dos vehículos distintos. El caso de los conceptos coextensivos requiere o bien que una propiedad se relacione a través de una única ley psicofísica con dos conceptos distintos, o bien que haya más de una ley en relación a la misma propiedad. En el primer caso, se trataría de leyes como: 'Los encuentros con instancias de *Yocasta* disparan los conceptos YOCASTA y MAMÁ' o 'La propiedad *ser Yocasta* resuena en nuestras mentes YOCASTA y MAMÁ'. En el segundo caso, habría dos leyes distintas, una acerca de la instanciación de MAMÁ ante encuentros con *Yocasta* y otra acerca de la instanciación de YOCASTA ante encuentros con la misma propiedad. Ambas alternativas presentan problemas.

Si se opta por la primera opción, i.e., que una propiedad se relacione a través de una única ley con dos conceptos distintos, se estaría diciendo que tanto YOCASTA como MAMÁ caen bajo una ley que los vincula con la propiedad *ser Yocasta*.²¹ Esta sería una ley acerca de cómo nuestras mentes resuenan a cierta propiedad. Pero entonces, ante la instanciación de esta propiedad, tendría que darse la co-ocurrencia de ambos conceptos, es decir que en la mente de Edipo deberían resonar ambos conceptos ante encuentros con *Yocasta*. No sería posible tener un concepto sin el otro, y esto iría en contra del atomismo conceptual, en la medida en que esta teoría implica que la posesión de un concepto atómico no depende de tener otros conceptos, sino de relaciones puramente nómicas con el mundo.

Las dificultades surgen porque no está clara la vinculación entre que una propiedad resuene de una determinada manera y que sea transportada por un vehículo determinado.

²⁰ Me ocuparé de estos mecanismos en el capítulo 2. Por ahora es suficiente con decir que Fodor propone que hay diversos mecanismos de acceso semántico, entre los cuales se cuenta la percepción, que sirven como mediadores en la adquisición de conceptos, pero que no son constitutivos de los mismos.

²¹ O *ser la madre de Edipo* (ver nota 18). De todas maneras se trataría de propiedades idénticas.

Tomados de forma independiente, tanto YOCASTA como MAMÁ parecen estar funcionando como la teoría dice que deberían. Una propiedad impacta la mente de Edipo y, en virtud de una ley, se dispara un concepto. De esta manera, Edipo se convierte en poseedor de una etiqueta para su enlace con dicha propiedad, por ejemplo, la etiqueta YOCASTA. Pero, si el enlace se produce con dos vehículos distintos, ¿cuál de esos vehículos es el que covaría con instancias de la propiedad? “Ambos”, debería responder Fodor. “Depende”, respondería probablemente Edipo, y su historia parece darle la razón. Pero ¿de qué depende? Si las condiciones de posesión de ambos conceptos son exactamente las mismas esto resulta inexplicable en la teoría fodoriana.

Es necesario, entonces, recurrir a la segunda alternativa y postular que en estos casos hay una propiedad que se relaciona por medio de leyes distintas con conceptos distintos. Así, YOCASTA cae bajo una ley y MAMÁ bajo otra. Sin embargo, esta alternativa también acarrea dificultades. Fundamentalmente, al igual que con la opción anterior, sería imposible tener un concepto sin el otro. Para evitar esta consecuencia, podría sugerirse que hay *algo* en virtud de lo cual una propiedad se enlaza con uno u otro vehículo, es decir que hay mecanismos distintos que median a cada una de las leyes. Los mejores candidatos en la teoría para cumplir esta función son los prototipos (o más precisamente, los mecanismos inductivos de generalización que dan lugar a los prototipos). Fodor señala que hay una instancia inicial en la adquisición de conceptos en la cual (a partir de mecanismos inductivos de generalización) se forma un prototipo, que se relaciona con el concepto (o etiqueta), la que a su vez se enlazaría con la propiedad correspondiente. Es decir, la adquisición o formación de prototipos es un mecanismo más de acceso semántico que funciona como mediador en la adquisición de conceptos (Fodor 2001, 2008). Podría ocurrir que distintas experiencias con la misma propiedad dieran lugar a prototipos distintos y, de esta manera, a conceptos distintos. Pero si el hecho de adquirir distintos prototipos fuera lo que causa la adquisición de conceptos distintos, los conceptos estarían en una relación de covariación con los prototipos, y no con las propiedades. Es decir, los prototipos serían así los que determinan la instanciación de uno u otro concepto. De este modo, se estaría estableciendo una relación más fuerte entre los prototipos y los conceptos que la de ser los primeros meros disparadores de los segundos.

Ahora bien, esta explicación resulta difícil de compatibilizar con la teoría fodoriana, fundamentalmente, por dos motivos. En primer lugar, esta propuesta supone una relación constitutiva entre un prototipo y un concepto, lo cual Fodor (1999, 2008) rechaza. En segundo lugar, si la diferencia en los prototipos para una misma propiedad diera lugar a conceptos distintos, se pondría en riesgo el requisito de publicidad de los conceptos -requisito que, en términos de Fodor (1999), no es negociable. Habría una misma propiedad vinculada, vía distintos prototipos, con distintos conceptos. Pero si se pueden tener dos prototipos de la misma propiedad, ¿por qué no tres, o cuatro, o cien? Esto implicaría una proliferación de conceptos, difícilmente compartibles entre distintos individuos.²² En consecuencia, asignar un rol constitutivo a los prototipos atenta también contra la publicidad de los conceptos.

De esta manera, la individuación de los conceptos coextensivos presenta un problema que es a la vez semántico y sintáctico. Si se insiste en una diferencia puramente sintáctica, ésta resulta difícil de explicar a partir de relaciones causales nomológicas como únicas asignadoras de contenido. Pero, por otra parte, la teoría también genera dudas acerca de la identidad entre la propiedad de *ser Yocasta* y *ser la madre de Edipo*. A partir de la definición de las propiedades de apariencia, podría parecer hasta necesario que *ser Yocasta* y *ser la madre de Edipo* fueran propiedades de apariencia distintas. Si ser una propiedad de apariencia consiste en resonar en la mente de una determinada manera, entonces propiedades que generen distintos impactos deberían ser distintas propiedades de apariencia. Así, el disparo de conceptos sintácticamente distintos quedaría explicado por tratarse de efectos de propiedades distintas.

Sin embargo, si esta fuera la explicación –i.e., si el disparo de vehículos distintos es una consecuencia de distintas maneras de resonar en la mente-, entonces YOCASTA y MAMÁ ya no serían coextensivos. Ambos conceptos portarían información acerca de propiedades de apariencia distintas, con lo cual la diferencia entre ellos no sólo sería sintáctica sino también semántica.²³ El tratamiento de Fodor (1994, 2008) del caso de Edipo

²² Además, los prototipos se sustentan en los mecanismos de acceso semántico, los cuales no necesariamente son compartidos por distintos individuos.

²³ En *El Olmo y el Experto*, Fodor aclara que no sostendrá una semántica informacional “full-blown” para los nombres propios (Fodor 1994, p. 106). Sin embargo, en *Conceptos*, la hace extensiva a la mayoría de los conceptos léxicos: señala que la metafísica de PICAPORTE es la misma para “todos los conceptos que no son lógico-matemáticos y no expresan clases naturales” (Fodor 1999, p. 195). Como consecuencia, puede

y de los nombres propios coextensivos en general muestra que esta multiplicación de las propiedades de apariencia no es lo que tiene en mente. Insiste en que en estos casos los conceptos tienen el mismo contenido y la diferencia entre ellos es puramente sintáctica. Es decir, si se los entiende como conceptos del mismo tipo que PICAPORTE, la propiedad de apariencia a la que refieren tiene que ser la misma.²⁴

El problema al intentar explicar estos casos es el de la relación entre la semántica y la sintaxis de los conceptos. Una propiedad de apariencia consiste en resonar en la mente de una determinada manera. Esto parece ser una cuestión semántica/epistémica; es decir, el impacto de las propiedades en nuestras mentes parece ser algo del orden de cómo concebimos las cosas. Ahora bien, puesto que Fodor no considera los factores epistémicos como constitutivos de los conceptos, la manera en que las propiedades resuenan en la mente debe reflejarse en la sintaxis. Sin embargo, no queda claro que la sintaxis pueda recoger las diferencias epistémicas que permiten individuar este tipo de conceptos sin apelar a los mecanismos de acceso semántico (MAS) como los que determinan cómo nos impactan las propiedades y con qué propiedad covaría cada concepto.

Fodor da buenos argumentos por los cuales los factores epistémicos no deben ser parte constitutiva de los conceptos. Desplazar este tipo de factores permite, entre otras cosas, preservar la publicidad y asegurar la composicionalidad semántica de los conceptos, así como fijar la referencia de los conceptos independientemente de creencias que pueden resultar falsas. Como consecuencia, la semántica informacional atomista, junto con un mecanismo de adquisición causal-bruto sustentado por conexiones nomológicas, debería hacer todo el trabajo en las explicaciones tanto de la metafísica de los conceptos como de su adquisición. Sin embargo, los casos Frege aparecen como objeciones a esta teoría de los conceptos y parecen introducir nuevamente los factores epistémicos que se intentaba desplazar, puesto que gran parte del peso explicativo recae sobre los MAS. Considero que la teoría fodoriana se torna aún más vulnerable a este tipo de crítica a partir de ciertas

tratarse a los conceptos YOCASTA y MAMÁ como conceptos que refieren a propiedades de apariencia.

²⁴ No es mi intención ocuparme aquí de las consecuencias que puede tener la propuesta de propiedades de apariencia para el realismo metafísico al que Fodor suscribe. Este es un problema independiente de la individuación de conceptos. Lo importante en relación con la individuación en el caso de YOCASTA/MAMÁ es que Fodor trata a las propiedades de apariencia como instanciadas en el mundo. En particular, trata a *ser Yocasta* como una propiedad instanciada en el mundo e idéntica a *ser la madre de Edipo*.

aclaraciones que se agregan en *LOT2*, puesto que allí se sugiere que dos conceptos atómicos coextensivos podrían tener historias ontogenéticas distintas (Fodor 2008, p. 75). Si dos historias ontogenéticas distintas que involucran a la misma propiedad en el mundo dan lugar a distintos conceptos, esto apunta hacia los MAS como los que explican, en última instancia, qué conceptos tenemos.²⁵

3. Conclusiones

Como vimos, el atomismo informacional presenta claras ventajas explicativas al poder dar cuenta de los requisitos de alcance y composicionalidad. Sin embargo, esta teoría tiene dificultades en cuanto al contenido intencional relacionadas con la individuación de conceptos coextensivos. ¿Cuán problemático resulta esto para una teoría de conceptos? Bastante, si esta teoría intenta contribuir a las explicaciones psicológicas del comportamiento, tal como pretende Fodor. Lo que está en juego a la hora de individuar conceptos, en última instancia, es la posibilidad de formular leyes intencionales para dar cuenta de la racionalidad. En el caso de Edipo, tiene sentido decir que desea casarse con Yocasta y no desea casarse con su mamá. Y su comportamiento sólo resulta racional si hay una manera de explicar por qué estos dos deseos no son de la forma 'deseo que *p*' y 'no deseo que *p*'. Si se individuara a los conceptos sólo en función de su contenido intencional, esto no sería posible, por eso Fodor apela a la sintaxis. Pero, como intenté mostrar, la sintaxis tampoco recoge las diferencias epistémicas que explicarían este caso.

Por otra parte, al intentar resolver el problema de la individuación de conceptos coextensivos, Fodor introduce nuevos inconvenientes en su teoría. O bien no debería ser posible que estos casos ocurrieran, dada la metafísica del pensamiento que propone el atomismo, o bien el rol que juegan los prototipos y los MAS se torna más relevante que el de ser meros mediadores puesto que, en los casos de conceptos coextensivos, distintos

²⁵ También hay problemas relacionados con la individuación de los símbolos o vehículos del LDP que traerían aparejados, a su vez, problemas para formular generalizaciones acerca de la conducta para casos en los que intervienen conceptos coextensivos. Por razones de extensión no podré desarrollar esto. Basta con decir que el hecho de que dos conceptos coextensivos tengan roles causales distintos en el comportamiento no puede explicarse por generalizaciones intencionales (porque los contenidos en estos casos son idénticos), pero tampoco por generalizaciones formales o sintácticas porque no se pueden individuar tipos de vehículos para el caso interpersonal. Aydede (1998) sostiene que la individuación de los vehículos sólo se consigue apelando a su rol funcional/causal y, por ende, es implausible que los individuos compartan sistemas completos de representaciones. Para el problema de la individuación de los símbolos del LDP, véase también Schneider (2009a, 2009b).

mecanismos de acceso en contacto con las mismas propiedades dan lugar a distintos conceptos. Retomaré la relación entre los conceptos, los prototipos y los MAS en el próximo capítulo, en el cual me ocuparé de los requisitos de adquisición y categorización.

Capítulo 2

El Atomismo conceptual.

La satisfacción de los requisitos de adquisición y categorización.

En el capítulo 1 expuse las tesis centrales del atomismo informacional en relación a la metafísica de los conceptos y analicé la satisfacción de los requisitos de alcance, composicionalidad e intencionalidad. En este capítulo me ocuparé de evaluar la satisfacción de los requisitos de adquisición y categorización. Dado que, como mencioné en la introducción, estos requisitos se relacionan con los mecanismos y procesos mentales, los trataré en relación con la teoría computacional de los procesos mentales que postula Fodor. Mi objetivo será mostrar que no se puede dar cuenta de estos requisitos apelando a procesos computacionales, ni solamente a las características metafísicas de los conceptos. De aquí que el atomismo informacional debería articularse con otro tipo de representaciones y procesos para poder explicar tanto la adquisición como la categorización.

1. Teoría Representacional Computacional de la Mente

Fodor (1999) expone su TRCM como una conjunción de 5 tesis:

1. La explicación psicológica es típicamente nómica e intencional.
2. Las representaciones mentales son las portadoras primitivas del contenido intencional.
3. Pensar es computar.
4. El significado es información.
5. Lo que sea que distingue conceptos coextensivos está en la cabeza.

Las tesis 2, 4 y 5 tomadas en conjunto describen la metafísica de las representaciones mentales y conforman el aspecto representacional de la TRCM. Como vimos en el capítulo 1, las RM son símbolos que tienen propiedades semánticas y sintácticas. La cuarta tesis alude a las propiedades semánticas, i.e. al contenido de las RM. Según esta tesis, el contenido consiste en la información que las RM transportan en virtud de relaciones nomológicas con las propiedades a las que refieren. A su vez, la segunda tesis

establece que las RM atómicas, i.e. los conceptos atómicos, son las portadoras primitivas del contenido intencional, de modo que los pensamientos adquieren su contenido de manera derivada por composición de las RM atómicas que los constituyen. Así, el contenido intencional de las AP y de los pensamientos en general se explica a partir de los contenidos de las RM atómicas y reglas de combinación sintáctica, quedando establecida una relación de constitución entre los conceptos y los pensamientos: los primeros son los constituyentes de los segundos.

La quinta tesis sostiene que los conceptos coextensivos se individualizan por sus propiedades formales o sintácticas. Hemos visto, en el capítulo anterior, que la relación semántica entre los contenidos y el mundo no es suficiente para la individuación de conceptos coextensivos y que por eso Fodor apela a las propiedades sintácticas de las RM. Estas son propiedades intrínsecas de las representaciones que se suponen equivalentes, en última instancia, a la realización física neurológica de cada concepto y están, por lo tanto, en la cabeza. Esta es la única manera en que las propiedades que individualizan conceptos coextensivos pueden tener consecuencias causales en los procesos mentales (Fodor 1999). Los problemas del requisito de intencionalidad de los que me ocupé en el capítulo anterior se relacionaban en particular con esta tesis.

Es preciso aclarar que las tesis cuatro y cinco son específicas del atomismo informacional de Fodor. Esta semántica es compatible con la TRCM, pero no es la única que se puede adoptar (cualquier teoría semántica sería en principio compatible con una TRCM).²⁶ De todas maneras, me parece adecuado exponer la TRCM tal como Fodor (1999) lo hace, esto es, como la conjunción de estas cinco tesis, porque de este modo queda claro cómo se relaciona su teoría de los procesos mentales con su teoría de conceptos (de la cual forma parte la semántica informacional atomista).

Respecto de las tesis restantes, la primera especifica el vínculo entre una teoría de conceptos y las explicaciones psicológicas, y la tercera define los procesos de pensamiento. En relación con la primera tesis, ésta sostiene que la psicología explica el comportamiento apelando a los contenidos de los pensamientos, en particular recurre a los contenidos de los deseos y las creencias -es decir de los estados mentales de AP-, y establece

²⁶ Otros filósofos adoptan una TRCM, sin adherir a una semántica informacional. Este es el caso de Block, por ejemplo, que defiende una semántica de rol conceptual.

generalizaciones legaliformes sobre esos contenidos. Fodor señala que las explicaciones científicas son típicamente legaliformes, es decir, consisten en subsumir eventos bajo leyes. La psicología no es una excepción: uno de sus objetivos es explicar eventos de tipo psicológico (estados mentales, comportamientos) a partir de leyes. Para explicar el comportamiento, la psicología formula generalizaciones legaliformes del tipo: si una persona *desea* P y *cree* que P sólo si Q, entonces, *ceteris paribus*, esa persona hará Q (o intentará llevar a cabo Q) (Fodor 1994). Por ejemplo, si una persona cree que trae mala suerte pasar por debajo de una escalera y desea evitar la mala suerte, entonces, al caminar por la calle, evitará a toda costa pasar por debajo de una escalera.

Este tipo de explicación se basa en el contenido del deseo y la creencia para predecir una cierta conducta, y supone, para esto, que los conceptos que aparecen en el contenido de los deseos y las creencias son los mismos. De lo contrario, ésta no sería una explicación racional del comportamiento. Si en el ejemplo anterior, MALA SUERTE tuviera un valor semántico diferente en la creencia y en el deseo, entonces ésta no sería una explicación racional de por qué se evita pasar por debajo de una escalera, puesto que, en este caso, aquello que se desea evitar no sería lo acarreado por pasar debajo de una escalera, sino alguna otra cosa. De aquí la primera tesis: las explicaciones psicológicas hacen uso de leyes que generalizan sobre los contenidos de los pensamientos.

Ahora bien, para poder adoptar explicaciones de este tipo, es preciso establecer también en qué consisten las AP. En términos de Fodor (1987, p. 17): “Para cualquier organismo *O*, y para cualquier actitud *A* hacia la proposición *P*, hay una relación ('computacional'/funcional) *R* y una representación mental *RM* tal que *RM* significa que *P*, y *O* tiene *A* si y sólo si mantiene *R* con *RM*.” Así, una AP se define como una relación psicológica entre un individuo y una RM, que es el contenido u objeto de la AP (el significado de esa RM es la proposición que el agente cree, desea, etc.).²⁷ De modo que los estados mentales de AP son estados relacionales, y la relación entre el organismo y la RM será diferente si se trata de una creencia, un deseo, etc.²⁸ Esto equivale a decir que hay un

²⁷ En el capítulo 1 utilicé la definición de AP de Fodor (1994) en la que se agregan los MDP como tercer elemento, quedando establecida una relación entre un individuo, un MDP -que es una oración en mentalés- y una proposición que es su contenido.

²⁸ Así, Fodor se opone a las concepciones de los estados mentales como monádicos (los cuales no admiten descomposición). Una ventaja del enfoque relacional es que permite explicitar relaciones entre distintas AP con el mismo contenido o entre dos AP del mismo tipo con contenidos distintos.

tipo de relación psicológica diferente correspondiente a cada tipo de AP. Estos distintos tipos de relaciones -las distintas AP- son relaciones computacionales, en el sentido en que se diferencian por sus roles funcionales, es decir, se individualizan por los procesos mentales en los que participan. Esto significa que se individualizan en función de los mecanismos de los cuales son *outputs* y los procesos a los que sirven de *inputs*. Fodor (1987) ilustra esto tomando una metáfora de Schiffer: creer que llueve es tener en la “caja de las creencias” una RM que significa que llueve. En cambio, desear que llueva es tener un caso de la misma RM, pero en la caja de los deseos. Por otra parte, también podemos tener dos AP del mismo tipo con distinto contenido, por ejemplo, podemos creer que llueve y creer que es de día. En este caso se está en un mismo tipo de relación, pero con RM que codifican distintas propiedades o estados del mundo. La individuación de los contenidos de las AP, es decir, de las RM, se explica (a diferencia de la individuación funcional de las AP mismas) a partir de la semántica informacional, que vimos en el capítulo 1.

La tercera tesis sintetiza el aspecto computacional de la TRCM, estableciendo que pensar es computar. Esto significa que los procesos de pensamiento consisten en transiciones causales entre símbolos: pensar es pasar de una RM a otra de manera ordenada -siguiendo determinados algoritmos- en virtud de las propiedades sintácticas o formales de esas RM. Por ejemplo, sobre una representación compleja como EL PASTO ES VERDE Y LA NIEVE ES BLANCA puede operarse una transformación que dé lugar a LA NIEVE ES BLANCA, porque esa representación tiene la forma P&Q y sobre representaciones de esta forma puede operarse una transformación que da lugar a Q (Fodor 2008). Los procesos mentales computacionales operan así sobre la estructura sintáctica de las RM y son sensibles *únicamente* a esta estructura. Las transformaciones de este tipo dependen, entonces, sólo de las propiedades sintácticas de las representaciones, pero tienen la virtud de preservar de manera confiable los valores semánticos que portan los vehículos representacionales sintácticos. Esto es así porque, como vimos en el capítulo 1, la sintaxis sirve de vehículo a los contenidos, es decir, es lo que codifica los contenidos intencionales y “refleja” así relaciones semánticas.

Según la primera tesis de la TRCM, las explicaciones psicológicas apelan a los contenidos de los estados mentales para explicar la conducta. Sin embargo, el contenido por sí mismo no tiene poder causal. Pues bien, la manera en que los contenidos pueden causar

comportamientos u otros estados mentales es por el hecho de estar “reflejados” en la sintaxis. Así, los procesos mentales operan en virtud de las propiedades sintácticas, pero éstas, a su vez, transportan el contenido, de modo que sirven de vínculo entre el contenido y las relaciones causales que permiten explicar la conducta. De esta manera, las transformaciones sintácticas “respetan” el contenido (Fodor 1987). Además, “en tanto que la sintaxis codifica aspectos semánticos que determinan los roles inferenciales de los símbolos, los procesos computacionales respetan la coherencia semántica del pensamiento. Así, son posibles mecánicamente razonamientos semánticamente coherentes (y que preservan la verdad)” (Skidelsky en prensa, p. 85).

Ahora bien, las propiedades sintácticas que gobiernan los procesos computacionales son propiedades locales de las RM. Esto significa que son propiedades que dependen sólo de las características intrínsecas o esenciales de las RM en cuestión, específicamente, dependen de las partes que constituyen una RM y de cómo esas partes están relacionadas entre sí. Como señala Fodor (2003, p. 26): “Para ver cuál es la estructura sintáctica de una oración no es necesario mirar “fuera” de ella...”. En este sentido, la sintaxis es una propiedad de las RM que no depende de ninguna relación con otras RM por fuera de la que está en cuestión (así como no depende de que se tengan otras RM en absoluto). De modo que las propiedades locales son esenciales para las RM, si cambiara alguna de ellas estaríamos en presencia de una RM diferente. Así, la identidad de las RM está definida sobre propiedades locales. Las propiedades globales, en cambio, se relacionan con elementos por fuera de la estructura constituyente de una RM, y así dependen de sistemas completos de RM o creencias -de allí su denominación. Por ejemplo, la relevancia es una propiedad global. No se puede determinar la relevancia de una RM atendiendo únicamente a sus características intrínsecas o estructura constituyente, sino que esta propiedad se determina de alguna manera por su relación con el resto de las RM. Si se desea evaluar la relevancia de una determinada RM para la confirmación o no de una hipótesis, ésta sólo se podrá determinar en relación con un conjunto de creencias de trasfondo. Otros ejemplos de propiedades globales son la simplicidad y la plausibilidad, las cuales también se determinan en función de conjuntos de creencias (potencialmente conjuntos completos de creencias) y serían imposibles de determinar atendiendo sólo a las propiedades intrínsecas de las RM.

Pues bien, los procesos computacionales pueden operar únicamente sobre las propiedades locales, dado que están gobernados exclusivamente por la sintaxis de las RM -son sensibles únicamente a la identidad de las RM y su estructura constituyente. Por este motivo, una TRCM sólo podrá explicar los procesos que dependen de este tipo de propiedades (Fodor 2003, 2008). Así, la TRCM resulta ideal para explicar los procesos de razonamiento deductivo, puesto que estos dependen únicamente de la forma lógica de los pensamientos, la cual está plasmada en la estructura sintáctica de las RM. Por ejemplo, deducir que tengo que darle de comer al gato a partir de las creencias de que si son las nueve de la noche, entonces tengo que darle de comer al gato y que son las nueve de la noche, no requiere considerar ninguna propiedad por fuera de las propiedades sintácticas de las RM que sirven de vehículo a cada una de las creencias.

Ahora bien, no todos los procesos mentales se pueden explicar atendiendo a propiedades locales, ya que parte de nuestra vida mental es sensible a propiedades globales.²⁹ Este es el caso, en particular, de los procesos de fijación de creencias empíricas -que son análogos a los de confirmación de hipótesis empíricas. Se trata de procesos inductivos o abductivos de inferencia (típicamente inferencias a la mejor explicación), que dependen de propiedades como la relevancia y la simplicidad para determinar qué se considerará como confirmatorio o no de una hipótesis. Según Fodor (1983, 2008) los procesos de confirmación de hipótesis científicas son *isotrópicos* -i.e., en principio, cualquier creencia previamente establecida puede influir en la determinación de los hechos relevantes para su confirmación- y son *quineanos* -el grado de confirmación de una hipótesis depende de propiedades globales, tales como la simplicidad y la plausibilidad. De

²⁹ La distinción entre procesos locales y globales (Fodor 2003) está relacionada con (aunque no es idéntica a) la distinción entre sistemas de *input* y sistemas centrales que se plantea en la teoría modular de la arquitectura cognitiva (Fodor 1983). No es mi intención desarrollar aquí esta teoría, pero cabe aclarar brevemente las características de cada uno de estos sistemas. Los sistemas de *input* son modulares, i.e. son mecanismos computacionales especializados que están informacionalmente encapsulados y son específicos de dominio, entre otras características, y tienen la función de generar representaciones del carácter (y distribución en algunos casos) de estímulos distales. Entre ellos se incluyen el lenguaje y los sistemas perceptivos, que no son equivalentes a los sentidos, sino más específicos -incluyen, entre otros, mecanismos de percepción de color y de análisis de formas y mecanismos que detectan el ritmo de los estímulos acústicos. Los sistemas centrales, en cambio, no están informacionalmente encapsulados y no son modulares. Son los que reciben la información producida por los sistemas de *input* y tienen como función principal la fijación de creencias mediante procesos inferenciales no demostrativos. De manera que se podría pensar que mientras que los procesos modulares son completamente computacionales, i.e. operan en virtud de las propiedades sintácticas locales de las RM, la mayor parte de los procesos centrales no parecen ser computacionales dado que operan en virtud de propiedades globales.

modo que si los procesos centrales de fijación de creencia son análogos a los procesos científicos de confirmación de hipótesis, estos también serán isotrópicos y quineanos. Así, este tipo de proceso resulta ser holista, puesto que depende de propiedades globales, las cuales se basan en una relación epistémica entre *todas* o gran parte de las creencias. Este tipo de procesos no resulta explicable en el marco de una TRCM, ya que los procesos computacionales que esta teoría postula, como vimos, dependen sólo de propiedades locales.

La moraleja es que hay procesos mentales que una TRCM no puede explicar y entre ellos se incluyen procesos fundamentales para nuestra vida mental -y nuestra vida en general- como la fijación de creencias y el aprendizaje. Este es el caso, también, de los procesos de categorización. Dado que la categorización es, según Fodor, un proceso abductivo de fijación de creencias, este proceso depende de propiedades globales. Ya sea que se postule que la categorización se lleva a cabo por un proceso de comparación de rasgos o recurriendo a información de tipo causal -como proponen, respectivamente, los defensores de las teorías de prototipos y ejemplares, y los defensores de la teoría-teoría-, será necesario delimitar qué conjuntos de rasgos o propiedades se tendrán en cuenta, i.e. cuáles serán *relevantes* para la tarea, de modo que será necesario recurrir a propiedades globales que no son parte de la estructura constituyente de las RM. Respecto de la adquisición, si se considera que consiste en un aprendizaje que implica la formulación y testeo de hipótesis, este proceso tampoco podrá explicarse computacionalmente, puesto que sería un caso de proceso inductivo o abductivo de fijación de creencias.

A continuación analizaré cómo trata el atomismo informacional los *desiderata* de adquisición y categorización, teniendo en cuenta el trasfondo (y estos problemas) de la TRCM.

2. Satisfacción de los requisitos

2.1 Adquisición

La adquisición presenta ciertas dificultades para el atomismo. Según la teoría fodoriana, los símbolos son innatos. Esto quiere decir que hay vehículos representacionales simbólicos que están disponibles de manera innata para adquirir su contenido intencional ante los estímulos correspondientes. Así, la adquisición de contenido no es innata, sino que

se desencadena o dispara a partir de encuentros con propiedades. Como vimos en el capítulo anterior, Fodor sostiene que los conceptos atómicos adquieren su contenido en virtud de la existencia de leyes psicofísicas que conectan instancias de propiedades en el mundo con instancias del concepto correspondiente (Fodor 1999). Por ejemplo, las instancias del concepto CARACOL caen bajo una ley según la cual las instancias de este concepto son causadas por (o serían efectos de) la propiedad de *ser caracol* instanciada en el mundo. Estas relaciones nómicas son las relaciones hacedoras de significado.

Una explicación usual de la adquisición de conceptos sostiene que se trata de un aprendizaje, en el sentido de estar mediado por relaciones racionales de confirmación de hipótesis. Según esta postura, la relación entre la experiencia y los conceptos es evidencial: la experiencia sirve como confirmatoria o disconfirmatoria de hipótesis como 'Eso es un GATO'. Una explicación así, no sería tratable en términos computacionales, por lo que vimos en §1 respecto de las propiedades globales, y quedaría por fuera del marco de una TRCM. Es, en parte, por ello, que Fodor niega que la adquisición de conceptos sea un aprendizaje. Contra esa postura, Fodor (1999, 2008) argumenta además que si se entiende el aprendizaje de ese modo, surgen fundamentalmente dos inconvenientes: (i) los conceptos no son hipótesis, sino los constituyentes de las hipótesis, es decir que se necesita de todas maneras una base de conceptos no aprendidos (innatos) para poder formular las hipótesis en cuestión y (ii) dado que los conceptos no son hipótesis, lo que se adquiere al confirmar hipótesis son creencias constituidas por conceptos, no conceptos en sí. Por estos motivos, Fodor niega que la adquisición de conceptos sea un aprendizaje y sostiene, en cambio, que es un proceso causal bruto, i.e. un mecanismo de disparo del concepto ante el estímulo correspondiente en el mundo. Así entendido, este proceso sería tratable en el marco de la TRCM.

Ahora bien, hay de hecho una relación innegable entre la experiencia y los conceptos (aunque no sea una relación de evidencia) y esto es algo que requiere explicación. Hay una relación entre las experiencias con picaportes, por ejemplo, y la adquisición de PICAPORTE: adquirimos el concepto PICAPORTE a partir de experiencias con picaportes y no a partir de experiencias con sombrillas, anteojos o alfileres. Este es el problema picaporte/PICAPORTE. Pues bien, según Fodor, los conceptos no son aprendidos a partir de la experiencia, pero son disparados (causados) por ella. Para explicar esta

relación, tal como se desarrolló en el capítulo 1, Fodor propone que la propiedad *ser picaporte* es una propiedad dependiente de la mente, en tanto que es una propiedad que resuena en mentes como las nuestras, causando que se instancie el concepto PICAPORTE. Sin embargo, esta no es toda la historia de la adquisición. Para que el mecanismo causal bruto funcione, es necesario que nuestras mentes puedan entrar en contacto con las propiedades en el mundo. Para esto, disponemos de ciertos MAS, entre los cuales la percepción se cuenta como el mecanismo principal. Otros mecanismos de este tipo pueden ser, por ejemplo, las inferencias teóricas o los instrumentos tecnológicos como sonares y radares. De modo que los MAS se introducen en la teoría para explicar cómo es que las propiedades causan la instanciación de conceptos.

Los MAS permiten que la mente entre en relación con las propiedades en el mundo, y así, posibilitan el enlace con las mismas. Como vimos, la relación entre los conceptos y las propiedades es de tipo nomológico. Sin embargo, esto no implica que esa relación deba ser básica, i.e. inmediata. Ese sería el caso si la mente humana pudiera entrar directamente en contacto con las propiedades, sin ningún tipo de mediación. Pero es un hecho acerca de los seres humanos que para que nuestras mentes entren en contacto con el mundo necesitamos algún tipo de mecanismo de acceso. Esta es la función que cumplen los MAS.³⁰ Así, las relaciones nómicas que confieren contenido a los conceptos se sustentan en estos mecanismos, pero ellos no son constitutivos del contenido de las representaciones mentales. Dada esta concepción, no es necesario tener ningún MAS en particular, sino tener *algún* mecanismo, y las diferencias de acceso entre distintos individuos a la propiedad en el mundo no atentan contra la publicidad de los conceptos –i.e., el hecho de que todos accedamos a sus contenidos. Esto es así porque, según el atomismo informacional, los MAS no son constitutivos del contenido, de modo que distintos MAS no modifican la información que transportan los conceptos.³¹ Lo único que determina el contenido es una relación nomológica con el mundo, la cual, aunque esté mediada por distintos mecanismos, se supone que es independiente de ellos.

³⁰ Esta función de mediación no es metafísicamente necesaria sino contingente. De manera que podríamos haber tenido un acceso semántico inmediato a las propiedades, o haber tenido otros tipos de MAS. Asimismo, es un hecho contingente acerca de nuestra constitución como seres humanos que la percepción sea el medio central para el acceso semántico (Fodor 1999).

³¹ Aunque podrían ser responsables de determinar qué vehículo transportará ese contenido, tal como argumenté en el capítulo 1.

Ahora bien, entre los mecanismos mediadores se incluye también la formación (o el aprendizaje) de prototipos, lo cual complica la explicación de la adquisición. Como hace explícito Fodor (2001, p. 144) al explicar las propiedades dependientes de la mente: “*ser picaporte* es tener la propiedad con la cual mentes como las nuestras se enlazan como consecuencia del tipo de experiencias a partir de las cuales nuestro tipo de mente aprende el prototipo de picaporte”. Según Fodor, la adquisición de PICAPORTE se produce en dos fases: por un lado, a partir de experiencias con picaportes se forma un prototipo (por mecanismos inductivos de generalización estadística). Y, por otro lado, opera una función de asociación de prototipos a conceptos, por la cual se produce el enlace entre la propiedad *picaporte* en el mundo y el vehículo PICAPORTE. De esta manera los prototipos funcionan como mediadores en la adquisición de conceptos (i.e., funcionan como MAS) pero, al igual que el resto de los MAS, no son constitutivos de los conceptos (Fodor 2001, 2008). Así se evitan los problemas que, como hemos visto en el capítulo 1, presentan los prototipos (en particular con la composicionalidad).

Estas aclaraciones, sin embargo, no explican cómo se produce el pasaje de prototipos a conceptos. En *LOT 2*, Fodor explica de manera similar a *Conceptos* la relación entre conceptos y prototipos, pero por suerte para el lector está presente Snark, que a cada paso tiene objeciones y preguntas.³² Snark pide las “aclaraciones pendientes” respecto de la relación entre conceptos, prototipos y experiencia. En primer lugar, pregunta por qué si la adquisición de conceptos no es una forma de inducción, se produce generalmente en paralelo a la adquisición de prototipos. Fodor explica, una vez más, que los conceptos no pueden ser prototipos porque los conceptos se combinan de manera composicional y los prototipos no. Sin embargo, al igual que en Fodor 2001, sugiere que los prototipos podrían participar en la primera etapa de la adquisición de conceptos. Habría una etapa inicial en la que se forma un prototipo, seguida por la adquisición del concepto, que consiste en un enlace con una propiedad. Snark, no satisfecho, pregunta cómo es que se pasa del *aprendizaje* del prototipo a la *adquisición* del concepto. Y Fodor responde: este es un proceso neurológico, no intencional, de modo que la adquisición de conceptos no es un proceso inferencial.

³² Snark es un personaje incorporado en *LOT2*, una criatura “repugnante e irritante”, que se encarga de hacer preguntas y no estar nunca de acuerdo con Fodor.

Fodor considera que la adquisición de un concepto, es decir, el pasaje del prototipo al concepto, no puede ser un proceso inferencial, y muestra que ninguna alternativa que pudiera instanciar dicho proceso resulta viable. Una posibilidad sería que se compare un prototipo con conceptos candidatos, pero de esta manera se presupone que ya se tienen los conceptos candidatos cuando lo que se quiere explicar es, precisamente, la adquisición de conceptos. Esta alternativa es evidentemente circular. Una segunda opción sería que exista un patrón de inferencia de acuerdo al cual dado su prototipo se infiere un concepto. El problema que señala Fodor respecto de esta posibilidad es que los conceptos y sus prototipos correspondientes pueden relacionarse de muchas maneras diferentes. Una relación puede ser la similitud, por ejemplo, para el caso de las instancias que caen bajo PARAGUAS, pero esta relación no se aplica en el caso del prototipo de, por ejemplo, PRESIDENTE. Un mismo prototipo podría incluso corresponder a distintos conceptos (Fodor sostiene que uno no puede simplemente observar un prototipo y saber a qué concepto corresponde).

Por estas razones, la explicación del pasaje del prototipo al concepto tiene que ser de otro tipo, no puede ser un proceso inferencial. Fodor apela entonces a la neurología y al tipo de mente que tenemos. Nuestras mentes son funciones de prototipos a conceptos, es decir que mapean prototipos en conceptos. Así, cuando aprendemos el prototipo de mesa, adquirimos MESA. Esta explicación, sin embargo, no esclarece demasiado, puesto que apela a una suerte de “armonía preestablecida” entre los conceptos y sus prototipos. Si un mismo prototipo puede corresponder a distintos conceptos, ¿qué es lo que determina el pasaje de dicho prototipo a uno u otro concepto? “Un proceso neurológico” es la única respuesta que ofrece Fodor.

Fodor incluso señala que hay una disposición innata para enlazarse con determinadas propiedades al aprender determinados prototipos. De esta manera, el equipamiento innato que se requiere para adquirir conceptos incluye, por un lado, ciertos mecanismos (como los mecanismos perceptivos) que ponen a nuestras mentes en contacto con las propiedades y les permiten así resonar (Fodor 1999) y, por otro lado, una disposición innata para enlazarse con propiedades al aprender sus prototipos (Fodor 2008). Esto indica que el aprendizaje de los prototipos no es suficiente para la adquisición de conceptos, sino que se requiere además una disposición innata para conectar esos prototipos

con los conceptos correspondientes. Ahora bien, según Fodor, esta disposición no es intencional, sino neurológica. Por un proceso neurológico los prototipos median la vinculación con las propiedades a las que nuestros conceptos refieren. Sin embargo, no queda demostrado que esta disposición innata no tenga nada de intencional, sino que se establece por estipulación que este hecho no es intencional.

Landau (2000) señala, respecto de la adquisición de conceptos, que el mecanismo de disparo de un concepto frente a instancias, que propone Fodor (1999), no es explicativo. Pues bien, tal como vimos, Fodor (2001, 2008) elabora algo más las características de este mecanismo y señala que está mediado por un aprendizaje. Sin embargo, se puede sostener la misma crítica respecto de la relación neurológica entre prototipos y conceptos: no explica cómo el hecho de aprender un prototipo permite que el concepto se enlace con la propiedad “correcta” especialmente si tenemos en cuenta que un mismo prototipo puede corresponder a más de una propiedad. Esta nueva versión de la adquisición avanza en dirección hacia una explicación, puesto que se propone incorporar nuevas herramientas, tales como los prototipos y las experiencias. Sin embargo, la introducción de los prototipos genera una nueva serie de interrogantes que es necesario resolver.

En especial, es necesario especificar cómo están vehiculizados los prototipos. No resulta descabellado pensar en los prototipos como RM. Pero si se tratara de construcciones codificadas en mentalés, esto atentaría directamente contra el atomismo del mentalés, puesto que se trataría de RM estructuradas. Es decir, si los prototipos fueran RM complejas del mentalés que sirven de mediadores en la adquisición de conceptos, esto significaría que para poseer un concepto, por ejemplo GATO, sería necesario poseer todos los conceptos que componen al prototipo y, de esta manera, GATO no sería atómico. Además, otro inconveniente con esta propuesta sería el riesgo de caer en una regresión al infinito, ya que para adquirir cada uno de los conceptos que constituyen a un prototipo harían falta, a su vez, otros prototipos mediadores constituidos también por conceptos. De modo que resulta más consistente con la teoría fodoriana interpretar que los prototipos son representaciones no conceptuales y que, en consecuencia, la adquisición de conceptos es un pasaje de representaciones no conceptuales a símbolos del LDP. Sin embargo, la propuesta de mediación de los prototipos requiere elaboración y la explicación de la adquisición todavía resuena en mentes como las de los críticos a cierta vaguedad.

2.2. Categorización

Fodor (1999, p. 46) sostiene que los conceptos “son categorías y se emplean, de manera rutinaria, como tales” y que es un requisito “no negociable” de una teoría de conceptos dar cuenta de esto. Sin embargo, como se mencionó en §1, no se puede explicar la categorización por medio de la TRCM dado que este es un proceso central de fijación de creencias por medio de inferencias abductivas, de modo que el atomismo debería incluir otro tipo de procesos mentales que puedan operar sobre las propiedades globales de las RM para poder satisfacer este requisito. Por otra parte, al postular que los conceptos son atómicos, es decir, que no tienen estructura interna, y que están constituidos independientemente de cualquier habilidad epistémica, no puede explicar la categorización sin recurrir a elementos no constitutivos de los conceptos (Laurence & Margolis 1999, Prinz 2002). Desarrollaré primero esta última dificultad, y luego retomaré el problema de los procesos globales.

Para explicar la aplicación, por ejemplo, del concepto GATO, dado el atomismo informacional, se debería recurrir o bien a las mismas leyes psicofísicas que establecen las relaciones entre los conceptos y el mundo, o bien a elementos no constitutivos del concepto. Se podría defender la primera alternativa sugiriendo que, dado que las leyes psicofísicas establecen que ante instancias de gatos en el mundo se disparará el concepto GATO, el categorizar a una instancia de gato como GATO, es simplemente representarlo como cayendo bajo el concepto GATO. Así, se podría explicar la categorización acudiendo a las mismas leyes psicofísicas que determinan el contenido de los conceptos. Sin embargo, podría ocurrir que estuviera oscuro, y alguien, frente a un gato, pensara LIEBRE. Podemos confundir gatos por liebres. ¿Significa esto que no poseemos el concepto GATO?

Fodor (1999) responde a esta pregunta afirmando que no existe ningún escenario tan claramente “gatuno” como para que la falla al categorizar erróneamente una instancia de gato bajo algún otro concepto implique la falta de posesión del concepto GATO. Esto es así porque Fodor diferencia entre las condiciones de posesión y las de aplicación de un concepto. Mientras que las primeras le corresponden al concepto de manera esencial, ya que poseer un concepto no es más que poseer una instancia de una representación que está nomológicamente conectada con una instancia de una propiedad en el mundo, las segundas no. Las condiciones de posesión dependen de la metafísica de los conceptos y son

esenciales en tanto dependen de las relaciones nomológicas constitutivas de los conceptos. Tener una representación nomológicamente conectada con la propiedad de *ser gato*, implica que se posee el concepto GATO. Así, las condiciones de posesión dependen de una relación semántica: la relación nomológica constitutiva del contenido de los conceptos.

En cambio, las condiciones de aplicación de un concepto dependen de relaciones epistémicas que no son constitutivas de los conceptos. Si bien la categorización establece, al igual que la representación, una relación entre conceptos y objetos, esta relación es distinta de la de representación (i.e. la relación nomológica que determina el contenido) puesto que no consiste en una mera relación causal entre un concepto y una propiedad, sino que en la categorización intervienen (en principio, de manera constitutiva) creencias, teorías y mecanismos perceptivos. Por este motivo, Fodor no puede explicar la categorización sin apelar a elementos no constitutivos de los conceptos.

Dado este escenario, reconstruiré en primer lugar la crítica al atomismo según la cual el hecho de que las características intrínsecas de los conceptos no sirvan para explicar la categorización implica una debilidad de la teoría (Prinz 2002), pero sostendré, siguiendo a Weinberg (2003), que los *desiderata* de una teoría de conceptos imponen que se dé cuenta de estos requisitos, pero no que se lo haga apelando exclusivamente a las características intrínsecas de los conceptos. En segundo lugar, desarrollaré una sugerencia de Fodor que podría permitir articular al atomismo informacional con las explicaciones de la psicología cognitiva acerca del proceso de categorización y, por lo tanto, con explicaciones de un proceso global.

Prinz (2002) señala que no poder explicar la categorización implica una debilidad importante para el atomismo conceptual, puesto que en las teorías de la psicología cognitiva se presta especial atención a los conceptos precisamente por su rol en la categorización, i.e. por su contribución en la determinación de que un elemento pertenece a una determinada categoría. Sin embargo, Prinz admite que la no satisfacción de este *desideratum* no es suficiente para descartar la teoría, aunque la deja en una posición débil frente a otras teorías que, pudiendo explicar el resto de los *desiderata*, pueden explicar también la categorización por medio de características esenciales de los conceptos. Frente a esta crítica, Weinberg (2003) objeta que Prinz “confunde las reglas del juego”. Una buena teoría de conceptos debería satisfacer ciertos requisitos explicativos, pero esto no implica que todas las

explicaciones deban basarse en las características metafísicas de los conceptos. Fodor se limita a señalar que las características metafísicas de los conceptos tienen que ser compatibles con aquello que las teorías psicológicas postulen para dar cuenta de nuestra vida mental. Así, reconoce el rol central que tiene en la vida mental la aplicación de conceptos como categorías y, por lo tanto, que aquello que postule una teoría de conceptos tiene que ser compatible con que los conceptos se apliquen cotidianamente en la categorización de objetos.

Tal como Fodor mismo sostiene: “Una teoría de conceptos tiene que explicar dos cosas: cómo los conceptos funcionan como categorías, y cómo una mente finita puede tener una capacidad conceptual infinita y sistemática” (1999, p. 149). Como vimos en el capítulo 1, los conceptos entendidos a la manera fodoriana permiten explicar la capacidad conceptual infinita y sistemática puesto que admiten la composicionalidad. Así, a partir de un conjunto finito de conceptos primitivos y reglas de combinación se pueden generar combinaciones infinitas y relacionadas entre sí. Ahora bien, respecto al requisito de explicar cómo los conceptos funcionan como categorías, la semántica informacional sólo dice que un concepto tiene como referente una clase de cosas: todas aquellas que instancian la propiedad con la que el concepto está nomológicamente relacionado. En este sentido, un concepto es una categoría.

Sin embargo, los *procesos* de categorización no pueden explicarse exhaustivamente sólo a partir de las representaciones conceptuales atómicas. Si esto fuera así, estos procesos tendrían que explicarse solamente a partir de leyes psicofísicas y esto no se sostiene. En la categorización intervienen también mecanismos perceptivos y creencias. Fodor (1999) no cree que sea posible formular una teoría de la individuación de las creencias, puesto que habría que recurrir para ello a la individuación por roles funcionales/causales, lo cual, por lo visto en §1, tiene la consecuencia inevitable del holismo.³³ Sin una teoría semejante, tampoco se pueden especificar las condiciones de verificación de los conceptos. Fodor diferencia entre las condiciones de satisfacción y las de verificación de los conceptos. Adopta una postura atomista respecto de las condiciones de satisfacción de un concepto, pero holista respecto de las condiciones de verificación. Es decir, los conceptos tienen

³³ Para individuar por el rol funcional/causal habría que tener en cuenta la intervención de las creencias en los procesos centrales, de modo que sería necesario establecer relaciones epistémicas entre la creencia que se quiere individuar y todo el resto del sistema.

esencialmente sus condiciones de satisfacción, y estas condiciones son atómicas, puesto que consisten en una relación mente-mundo que no depende de otras relaciones. Una propiedad instanciada cae bajo un concepto o no dependiendo de las relaciones nómicas. Las condiciones de verificación, en cambio, se refieren a las condiciones para *decidir* si un ítem cae bajo un concepto o no. Para cada instancia particular de aplicación de un concepto, estas relaciones están mediadas por teorías y creencias y son, por lo tanto, holistas (Fodor 1999). Estas relaciones no le corresponden al concepto de manera esencial, sino que al estar mediadas por teorías y creencias, varían de persona a persona, mientras que los conceptos son públicos y esencialmente idénticos (en tipo) en todas las mentes.

Ahora bien, a pesar de la distinción entre satisfacción y verificación, y de las dificultades para individuar creencias y, por lo tanto, para especificar condiciones de verificación, resulta explicativamente débil sostener que los conceptos se usan como categorías, pero que los procesos que permiten determinar si una instancia corresponde a una categoría, no forman parte de la teoría metafísica de los conceptos. El problema no es tanto que Fodor no puede explicar la categorización sin recurrir a mecanismos no constitutivos de los conceptos, sino que directamente excluye esta explicación de su teoría. Fodor sostiene que la aplicación de un concepto en una situación determinada, es decir, el reconocimiento de una instancia como cayendo bajo un concepto, es una forma de adquisición de una nueva creencia (por ejemplo reconocer a un gato como GATO implica formar o adquirir la creencia 'Ese es un gato') y, por lo tanto, no es meramente el disparo de un concepto frente a una instancia.

Ante las críticas por la falta de una explicación de la categorización (por ejemplo, Landau 2000), Fodor (2000) se desentiende del problema sosteniendo que no es un requisito para una teoría metafísica de conceptos explicar *cómo se aplican* los conceptos sino especificar *qué* es un concepto y cuáles son sus condiciones de posesión. Es decir, no es un requisito explicar qué mecanismos permiten determinar si cierto objeto cae bajo un concepto o no, sino explicar las propiedades metafísicas de los conceptos. La encargada de dar cuenta de este tipo de procesos sería la psicología cognitiva. Sin embargo, retomando las consideraciones de §1 respecto de los procesos globales o centrales, Fodor (1983) considera que los únicos sistemas que la psicología cognitiva logrará explicar satisfactoriamente serán los sistemas modulares que operan sobre propiedades locales de

las RM (ver nota 29). De modo que la exclusión del requisito de categorización de las teorías de conceptos es consistente con su enfoque de la arquitectura cognitiva y del tipo de explicación que las ciencias cognitivas pueden formular.

Ahora bien, esta perspectiva resulta quizás excesivamente desesperanzada, ya que hay teorías psicológicas disponibles que intentan explicar la categorización recurriendo a representaciones de rasgos asociadas a cada categoría y a una función de similitud que compara el ejemplar que se quiere categorizar y un concepto determinado (Laurence & Margolis 1999, Prinz 2002). Estas teorías explican la categorización como una comparación entre los rasgos percibidos y los rasgos asociados a la categoría a la que el elemento corresponde.³⁴ Fodor (1999) sostiene que la metafísica de los conceptos debe ser compatible con las explicaciones de la psicología acerca de las funciones de los conceptos, pero para que el atomismo informacional fuera compatible con éstas, sería necesario que explicara la articulación entre las representaciones conceptuales atómicas y la información asociada a ellas (los rasgos que determinan que esa categoría se asigne a un elemento o no).

Tal como vimos en §2.1, Fodor acepta que los conceptos tienen prototipos y que el aprendizaje de los prototipos es una etapa de la adquisición de conceptos. De modo que se podría recurrir a estos prototipos asociados a los conceptos para explicar la categorización. Esta sugerencia no parece descabellada puesto que Fodor mismo admite que los prototipos resultan adecuados (o *bastante* adecuados) para explicar la categorización (Fodor & Lepore 1996). Sin embargo, tal como vimos en §1, la articulación entre conceptos y prototipos no está clara. De hecho, en palabras de Fodor (2008, p. 168) "... no se sabe prácticamente nada acerca de la relación entre los prototipos y los conceptos correspondientes, excepto que son cosas diferentes".

Por otra parte, Fodor (2008) también recurre a la metáfora de los archivos mentales para sugerir un modo en que los conceptos pueden estar asociados a conocimiento sobre las propiedades que representan. Sostiene que parte de nuestra mente podría estar organizada en forma de archivos mentales etiquetados por conceptos y se refiere a estos archivos como

³⁴ Tal como veremos en el capítulo 5, no hay consenso entre los psicólogos acerca del tipo de estructura conceptual que permite explicar la categorización. Podría tratarse de una estructura de rasgos prototípicos, ejemplares, teorías o una combinación de ellas. Dado que Fodor admite que los conceptos tienen un prototipo, me ocuparé sólo de la articulación entre los conceptos, tal como Fodor los entiende, y los prototipos.

conjuntos de creencias en mentalés. La información prototípica que sirve para explicar la categorización podría ser parte de estos archivos y así estar codificada o transportada por oraciones en mentalés. Esto parece plausible, puesto que, al igual que los teóricos de las teorías de prototipos que ofrecen descripciones lingüísticas para caracterizar a este tipo de estructuras, se podría postular una descripción análoga en mentalés. Sin embargo, tal como vimos en en §2.1, esto entraría en tensión con la tesis del atomismo del mentalés. Los prototipos no parecen ser el tipo de cosa que pueda ser transportada por los símbolos del mentalés ya que mientras los primeros poseen estructura interna, los segundos son atómicos.

A pesar de esta dificultad, podría interpretarse la metáfora de los archivos mentales como un elemento que permitiría articular la teoría fodoriana con las explicaciones psicológicas. Las explicaciones psicológicas apelan a información (fundamentalmente, prototipos y teorías) para explicar la categorización. Si Fodor puede mostrar que esa información no es constitutiva de los conceptos, sino que solamente está asociada a ellos, y que el proceso de categorización termina con la asignación de una etiqueta a un ítem del mundo (aunque lo haga en virtud de la información que contienen los archivos, y no en función de la etiqueta en sí), entonces su teoría podría responder a la crítica respecto de la imposibilidad de categorizar utilizando conceptos atómicos. La sugerencia de los archivos es reciente y todavía requiere desarrollo, pero indica una manera posible de defender a la teoría fodoriana. Sin embargo, para incorporar esta sugerencia a la teoría será necesario especificar la taxonomía de entidades mentales, puesto que habría que especificar qué tipo de representaciones son los prototipos y qué tipo de representaciones se almacenan en los archivos.

Por otro lado, la articulación del atomismo informacional con algún modelo de la psicología cognitiva que explique la categorización, caerá inevitablemente por fuera de la TRCM. Sin embargo, considero que esto no es una desventaja, sino que implicaría un avance en la dirección de explicar los procesos centrales, de modo que es una alternativa que valdría la pena profundizar, en especial considerando que Fodor pretende que su teoría de conceptos sea compatible con las explicaciones de la psicología cognitiva.

3. Conclusiones

La TRCM tiene numerosas ventajas. Resulta adecuada para explicar los procesos deductivos de inferencia y ofrece una explicación de la intencionalidad de los pensamientos que permite vincular a los mismos con el comportamiento intencional. Por otra parte, es una teoría que integra una explicación de la metafísica de las RM con la idea de que los procesos de pensamiento son transiciones causales gobernadas por propiedades locales. Sin embargo tiene limitaciones, en particular respecto de las inferencias inductivas y abductivas y todo aquello que dependa de propiedades globales. De modo que hay cierto tipo de procesos que quedan por fuera de esta teoría, como la categorización, que no es explicable de manera computacional. Sin embargo, éste es un requisito central de una teoría de conceptos y no debería quedar sin explicación.

Por otra parte, tanto la explicación que ofrece Fodor de la adquisición como la posible articulación entre su teoría de conceptos y una explicación psicológica de la categorización -que es algo a lo que Fodor aspira- exhiben la necesidad de elaborar y esclarecer la relación que hay entre conceptos y prototipos. Por un lado, Fodor sugiere que los prototipos cumplen una función al enlazarnos con las propiedades en la adquisición de conceptos y, por otro lado, admite que resultan satisfactorios para explicar la categorización. Sin embargo, no termina de especificar de qué tipo de representaciones se trata ni cómo se vinculan con los conceptos.

Capítulo 3

El neo-empirismo de conceptos.

La satisfacción de los requisitos de intencionalidad, alcance y composicionalidad

En los capítulos 1 y 2 reconstruí las tesis principales del atomismo conceptual, y evalué la satisfacción de los requisitos fundamentales. En el capítulo 1, mostré que esta teoría, al postular que los conceptos son símbolos de un LDP, resulta especialmente adecuada para dar cuenta de la composicionalidad y el alcance. Por otra parte sostuve que la explicación de la intencionalidad, si bien resulta atractiva, genera dificultades para la individuación de conceptos. En el capítulo 2 reconstruí la TRCM y sostuve que esta teoría resulta adecuada para explicar los procesos deductivos de razonamiento, pero tiene dificultades para explicar los procesos de adquisición y categorización de manera computacional.

En este capítulo y el próximo me ocuparé del neo-empirismo de conceptos. Siguiendo una estructura similar a la del capítulo 1, en este capítulo expondré las tesis principales del neo-empirismo respecto de las características de los conceptos y analizaré si puede satisfacer los requisitos de intencionalidad, alcance y composicionalidad. Intentaré mostrar que esta teoría tiene dificultades con estos tres requisitos.

1. Características de los conceptos

Según el neo-empirismo de conceptos, los conceptos son proxitipos, i.e. particulares mentales internamente estructurados, formados por copias o combinaciones de copias de representaciones perceptivas (RP). Así, esta teoría toma del empirismo clásico la tesis de que todo lo que está en el pensamiento proviene de los sentidos. De aquí el nombre de “neo-empirismo”.³⁵ Prinz (2002, 2005) enfatiza que los proxitipos son conjuntos de representaciones que *se usan* para detectar instancias de categorías. Se puede caracterizar a esta teoría como la conjunción de tres tesis centrales, las cuales desarrollaré luego de su presentación:

³⁵ Retomaré esta tesis de la prioridad perceptiva en el próximo capítulo, en relación al requisito de adquisición.

Empirismo de vehículos: los conceptos son copias (o combinaciones de copias) de representaciones perceptivas.

Tesis semántica: el contenido de los conceptos se determina por medio de relaciones nomológicas con propiedades en el mundo (contenido intencional o real) y, al mismo tiempo, los conceptos representan las propiedades a partir de las cuales detectamos las categorías (contenido cognitivo o nominal).

Sensibilidad al contexto: las representaciones conceptuales son estructuras altamente variables.

Empirismo de vehículos

El neo-empirismo de conceptos es fundamentalmente una tesis acerca de los vehículos del pensamiento (Prinz 2002, p. 109). Según esta tesis, los conceptos son “copias almacenadas de representaciones producidas por la percepción” (Prinz 2002, p. 272) y el formato que transporta el contenido de los conceptos es el propio de cada uno de los sistemas perceptivos. Prinz caracteriza a estos últimos como sistemas de *input* dedicados. Se trata de sistemas que presentan un conjunto distintivo de operaciones y representaciones alojados en una vía neural propia, que responden a *inputs* específicos de dominio (por ejemplo, la visión responde a longitudes de onda, mientras que el olfato responde a las formas de las moléculas). Las RP son los *outputs* de cada uno de estos sistemas. De esta manera, el formato de los conceptos es modal, i.e. específico de cada sistema de *input* (hay representaciones visuales, auditivas, etc.).

Esta caracterización de las RM se opone a concebirlas como amodales o simbólicas, tales como las que postula Fodor, que, como vimos, son representaciones análogas a los símbolos de un lenguaje natural. Prinz (2002, 2005) rechaza la idea de que los conceptos son representaciones amodales que funcionan como etiquetas. A partir de un argumento de parsimonia, sostiene que postular un nivel de representaciones amodales es superfluo y que todas las funciones conceptuales se pueden explicar apelando a conceptos estructurados a partir de RP, prescindiendo de las etiquetas amodales. Tal como reconstruye la teoría fodoriana, habría un nivel de mecanismos detectores (i.e., los MAS) que serían mediadores entre las propiedades en el mundo y los indicadores de esas propiedades (i.e., estos últimos serían los conceptos fodorianos). Contra Fodor, Prinz argumenta que los mecanismos de

detección realizan todo el “trabajo” conceptual, de manera que los indicadores serían superfluos y por eso los conceptos *son* mecanismos detectores.³⁶ El argumento de parsimonia señala que es innecesaria la multiplicidad de representaciones -i.e. es innecesario postular detectores e indicadores- y que se debe identificar a los conceptos con los conjuntos de RP que se usan en la detección de categorías.

Ahora bien, las copias de RP que usamos para detectar instancias de categorías se almacenan en la memoria a largo plazo formando redes de representaciones ligadas entre sí, de modo que la información en un proxitipo está articulada a partir de distintas asociaciones entre representaciones. Prinz no desarrolla en profundidad los principios a partir de los cuales se producen estas asociaciones, pero menciona dos: la co-instanciación de representaciones “en un mismo objeto” y la similitud entre representaciones. El primero da lugar a enlaces *jerárquicos*, que permiten “acercar” o “alejarse” las representaciones (focalizando partes específicas), *transformacionales*, que agrupan transformaciones posibles de una representación en otra (por ejemplo, un animal en distintas posiciones), *de unión*, que asocian representaciones de distintas modalidades, y *situacionales*, que agrupan representaciones de distintos objetos que generalmente co-ocurren (por ejemplo, de un gorila y una banana). El segundo principio da lugar a los enlaces predicativos, i.e. la disposición a transferir características de una representación a otra (Prinz 2002, p. 145).

Por otra parte, Prinz (2002) sostiene que las RP sirven de vehículo a distintos tipos de información. Las redes de memoria a largo plazo pueden tener una parte prototípica, otra con información acerca de relaciones causales e, incluso, incluir representaciones de palabras en lenguaje natural. Es discutible si en la teoría de Prinz los prototipos, ejemplares y teorías son partes de cada concepto –entendiendo así que su teoría es una teoría híbrida (Machery 2009)- o cada estructura constituye un concepto distinto –entendiendo así que su teoría es pluralista respecto de las estructuras conceptuales, i.e. que hay, por ejemplo, distintos conceptos PERRO con distintas estructuras y, así, con distinto tipo de información. Tal como lo entiendo, en la mayoría de los contextos Prinz parece referirse a los prototipos, ejemplares y teorías como partes o componentes de los conceptos.³⁷

³⁶ Prinz se refiere a los conceptos tanto en términos de RM como mecanismos.

³⁷ Prinz (en prensa, p. 1) sostiene explícitamente que los prototipos son componentes de los conceptos.

Tesis semántica

Respecto al aspecto semántico de los conceptos, Prinz adopta una versión de la teoría semántica de doble factor, y sostiene que los proxitipos refieren de dos maneras diferentes, es decir, tienen dos tipos de contenido distintos: contenido intencional o real y contenido cognitivo o nominal. El contenido real refiere a esencias reales (aunque ocultas) de las cosas. El segundo, en cambio, refiere a las propiedades a partir de las cuales identificamos las instancias en el mundo que caen bajo un concepto. Para explicar el contenido real de los conceptos Prinz adopta una semántica informacional, según la cual el contenido está determinado por relaciones nomológicas entre instancias de los conceptos e instancias de las propiedades a las cuales aquellos refieren. Según su versión de la semántica informacional, *X* es el contenido intencional de un concepto *C* si se satisfacen dos condiciones:

Condición etiológica: una instancia de *X* fue una causa incipiente de *C*.

Condición de covariación nomológica: las instancias de *X* covarían nomológicamente con instancias de *C*.

La primera condición hace referencia a la historia causal del concepto. Los conceptos refieren a sus causas incipientes, i.e. a los elementos que de hecho causaron las primeras instancias de ese concepto.³⁸ De este modo, Prinz apela a la historia causal actual y se distancia de versiones de la semántica informacional, como la de Fodor, según las cuales las relaciones nomológicas están sustentadas puramente por contrafácticos. Esta condición se propone para resolver los problemas del error y del contenido disyuntivo, que vimos en el capítulo 1. Si MARIPOSA fue causado por encuentros perceptivos con mariposas y se aplica a una polilla, su contenido no es *polilla o mariposa*, puesto que *polilla* no se encuentra entre sus causas incipientes.

La condición de covariación nomológica contribuye a delimitar el contenido, y evita así los llamados problemas del respecto y de la cadena causal. El primer problema consiste

³⁸ Es discutible cómo interpretar aquí el plural de “primeras”, puesto que podría querer decir que el primer encuentro con cada propiedad determina la adquisición del concepto o que son una serie de encuentros (los primeros) con cada propiedad. Me ocuparé de esto en el capítulo 4, en relación al requisito de adquisición.

en la dificultad de determinar, por ejemplo, en caso de que una representación de GATO sea causada por un gato, si esta representación refiere a *gato*, o, como podría ser, a *animal*, o a *cuadrúpedo*. La covariación evita este problema, ya que permite sostener que GATO refiere a *gato* y no a *animal* ni a *cuadrúpedo*, porque en mundos posibles donde no hubiera gatos, otros animales o cuadrúpedos no causarían una instancia de GATO, de modo que GATO covaría con *gatos* y no con la clase de todos los animales ni de todos los cuadrúpedos. El segundo problema se relaciona con los distintos elementos que intervienen en la compleja cadena causal que lleva a la instanciación de un concepto. ¿Por qué GATO refiere a una propiedad en el mundo y no a imágenes en la retina? La condición de covariación nomológica resuelve este problema, ya que en un mundo posible donde los gatos tuvieran una apariencia diferente, las imágenes de gatos en la retina del mundo actual no causarían GATO. Sin embargo, las instancias de gatos seguirían causando GATO, de modo que el contenido de GATO covaría nomológicamente con las instancias de gatos y no con las imágenes en la retina y, por eso, su contenido es la propiedad en el mundo de *ser gato*. La condición de covariación nomológica, al igual que en la semántica informacional fodoriana, permite resolver estos dos problemas.

Ahora bien, Prinz no considera que el contenido de los conceptos se reduzca al contenido intencional, ya que éste no resulta suficiente para individuar conceptos en los casos Frege o los casos de tierra gemela. Tal como vimos en el capítulo 1, estos son casos en los que, respectivamente, dos conceptos distintos tienen el mismo contenido intencional, o un concepto tiene dos contenidos intencionales distintos en el mundo actual y en un mundo posible. Al igual que Fodor, Prinz pretende que una teoría de conceptos sea parte de las teorías psicológicas que explican el comportamiento apelando a los contenidos de los estados mentales. Puesto que los deseos y las creencias heredan los contenidos de los conceptos que los componen, es necesario poder individuar los conceptos de manera fina para articularlos en una explicación adecuada del comportamiento.

De modo que para explicar los casos Frege y los casos de tierra gemela, Prinz introduce otro tipo de contenido, el contenido cognitivo o nominal. Mientras que el contenido real refiere a las esencias reales de las categorías en el mundo, el contenido nominal refiere a las propiedades (también en el mundo) por medio de las cuales identificamos esas categorías, esto es, refiere a las apariencias de las instancias de esas

categorías. Por ejemplo, el contenido real de ORO es *Au*³⁹ y su contenido nominal incluye las propiedades de apariencia del oro, como *brillante*, *amarillo* y *maleable*. Es importante observar que el contenido nominal también se explica por medio de una semántica informacional.

La articulación entre los dos tipos de contenido no resulta sencilla. En principio, el contenido nominal “aproxima” el contenido real y, de esta manera, sirve para detectar ese contenido. Es decir, el contenido nominal sería un intermediario entre los conceptos y sus contenidos reales. Sin embargo, Prinz también señala que los proxitipos *constituyen* el contenido nominal, puesto que consisten en un conjunto de RP y las RP primitivas se individualizan por las apariencias que detectan. Me ocuparé de la relación entre el contenido real y nominal en §2.1 al tratar el requisito de intencionalidad.

Sensibilidad al contexto

La tercera tesis afirma que los conceptos son construcciones altamente variables, dado que la representación que se activa en un contexto no será la misma que se active en otro. Prinz (2002) explica esto en relación al *uso* de los conceptos. Puesto que los conceptos se usan como mecanismos para detectar instancias de categorías, la representación que se activa en cada caso está relacionada con el contexto. No será la misma RP la que usemos para rastrear un perro guardián o un perro en el ártico, aunque ambas son proxitipos del concepto PERRO porque refieren a la misma propiedad en el mundo (es decir, tienen el mismo contenido real). De acuerdo a esta propuesta, las RP se almacenan en redes de memoria a largo plazo y cada concepto puede tener diferentes proxitipos en la memoria de trabajo según la información que se active, es decir, el mismo contenido intencional puede ser codificado en proxitipos distintos, dependiendo del contexto. En este sentido, los conceptos son “construcciones altamente variables en la memoria de trabajo” (Weinberg *et al.* 2003, p. 297). Pero, de esta manera, los proxitipos oscilan entre dos niveles distintos: la memoria de trabajo –como construcciones temporarias- y la memoria a largo plazo –como redes más o menos estables. ¿Los conceptos son RP almacenadas o construcciones variables, sensibles al contexto? Esto requiere elucidación.

³⁹ *Au* recoge aquí las características esenciales del oro, tales como su número atómico.

En primer lugar, si los conceptos fueran idénticos a construcciones variables derivadas de estructuras mayores de información que están almacenadas en la memoria a largo plazo y se activan temporariamente en la memoria de trabajo, entonces dos proxitipos distintos deberían ser considerados dos conceptos distintos. Pero la propuesta de Prinz no sigue esta dirección. Prinz admite que distintos proxitipos pueden representar *el mismo* concepto, como en el caso de los distintos proxitipos de PERRO. En segundo lugar, Prinz afirma que los conceptos no pueden ser completamente identificados con construcciones temporarias en la memoria de trabajo porque eso significaría que sólo poseemos los conceptos que están activos en la memoria de trabajo en un determinado momento. De este modo no seríamos capaces de almacenar conceptos.

Por otra parte, los conceptos tampoco se identifican con las redes de memoria a largo plazo. Prinz (2005) caracteriza a los conceptos como bases complejas de datos, lo cual parece ubicar a los conceptos en el nivel de las redes de memoria a largo plazo. Sin embargo, ofrece también argumentos muy convincentes acerca de por qué un concepto no puede ser idéntico a una red entera. En primer lugar, apela a la necesidad de distinguir entre el conocimiento estable (aquel que se conserva aunque no estemos pensándolo a cada momento, es decir, mientras no está activo en la memoria de trabajo) y los pensamientos ocurrientes. En segundo lugar, Prinz (2002) señala que dado que la memoria de trabajo no tiene capacidad suficiente para activar una red de memoria entera, los pensamientos en tanto estados ocurrientes no pueden estar constituidos por redes enteras de memoria. Más aún, dada la capacidad limitada de la memoria de trabajo, entender a los conceptos como redes de memoria a largo plazo admitiría la posibilidad de poseer conceptos lo suficientemente complejos como para exceder la capacidad de la memoria de trabajo. De esta manera, dichos conceptos nunca podrían ser activados y así habría conceptos impensables.

Para evitar estos problemas, Prinz apela a una definición modal de proxitipos. Estos son “representaciones mentales que están o *pueden ser* activadas en la memoria de trabajo” (Prinz 2002, p. 149). Esta definición resuelve, por un lado, el problema del almacenamiento, ya que la noción de concepto no queda limitada a aquello que está activo en un momento determinado y, por otro lado, elude los perturbadores conceptos impensables, ya que, por definición, un concepto es aquello que *puede* ser activado en la

memoria de trabajo. Sin embargo, esta definición oscilante enfrenta dificultades a la hora de dar cuenta de la individuación de conceptos, como analizaré a continuación, a propósito del requisito de intencionalidad.

2. Satisfacción de los requisitos

2.1 Intencionalidad

Como vimos en relación al atomismo de conceptos, la semántica informacional no resulta suficiente para dar cuenta de la individuación de conceptos. Mientras que Fodor recurre a las propiedades sintácticas de los conceptos para resolver esta cuestión, Prinz apela a otro tipo de contenido, i.e. el contenido cognitivo, adoptando una semántica de doble factor. Sin embargo, esto genera nuevos problemas para la teoría.⁴⁰ Esta semántica de doble factor introduce el riesgo de que el contenido real quede reducido a contenido nominal, es decir, que la referencia de los conceptos se reduzca a meras apariencias de las categorías. Prinz (2002, p. 124) afirma que las representaciones que nos permiten detectar categorías intervienen en las relaciones que determinan el contenido. Así, el contenido intencional estaría determinado, en parte, por el contenido cognitivo, puesto que las representaciones a partir de las cuales detectamos las categorías constituyen el contenido cognitivo. Esto ocurre en parte debido a las dificultades que acarrea identificar a los conceptos con proxitipos. Por otro lado, esta semántica no logra resolver los problemas de individuación sin apelar a representaciones que no sean perceptivas o indicadores, los cuales, tal como vimos en §1, Prinz considera superfluos.

El problema de la individuación de conceptos en esta teoría radica en cómo entender la relación entre conceptos y proxitipos. Como vimos, los proxitipos se definen como construcciones variables, sensibles al contexto, que pueden activarse en la memoria de trabajo. Ahora bien, los conceptos no pueden ser idénticos a construcciones variables que son temporariamente activadas en la memoria de trabajo. Si esto fuera así, Prinz tendría que admitir que dos proxitipos distintos son dos conceptos distintos. Sin embargo, sostiene que distintos proxitipos pueden representar *el mismo* concepto en distintos contextos. De esta manera, no tenemos un concepto diferente cada vez que pensamos acerca de perros, sino que usamos, según esta propuesta, distintos proxitipos correspondientes al mismo concepto.

⁴⁰ Desarrollé parte de esta crítica en Haimovici (2009b).

Si distintos pensamientos, tales como que una tortuga está en una playa y que una tortuga está en un jardín están constituidos por el mismo concepto TORTUGA, a pesar de estar vehiculizados por proxitipos distintos, se sigue que los conceptos no pueden ser idénticos a proxitipos y que no pueden individuarse por sus vehículos, sino más bien por sus contenidos.

Pues bien, el contenido real no resulta suficiente para la individuación, puesto que no permite dar cuenta de los casos Frege, de modo que todo parece señalar al contenido nominal como el responsable último de la individuación de conceptos. Sin embargo, el contenido nominal equivale a apariencias. Si los conceptos se individuaran por sus apariencias, no se podría dar cuenta de los casos de tierra gemela ni de los casos de gemelos en esta tierra (introducidos en el capítulo 1). Una solución posible es apelar a la individuación por tipos, de modo que distintos casos de proxitipos sean idénticos en tipo. Esta es la estrategia que adopta Prinz, de manera que debería especificar cuándo dos representaciones en la memoria de trabajo corresponderían al mismo tipo de concepto. Pero, ¿a un tipo de concepto en tanto tipo real o nominal? En principio, los dos tipos de contenido no son independientes, sino que los contenidos reales están determinados, en parte, por las representaciones que conforman el contenido nominal, y el contenido nominal, a su vez, es una “aproximación” al contenido real (Prinz 2002, p. 282). Sin embargo, sólo es posible individuar tipos reales y tipos nominales de manera independiente, es decir, un concepto individuado como tipo nominal es un proxitipo, pero individuado como tipo real es un conjunto de proxitipos distintos que comparten el mismo contenido real. La ilustración 1 ejemplifica esta relación entre conceptos y proxitipos.

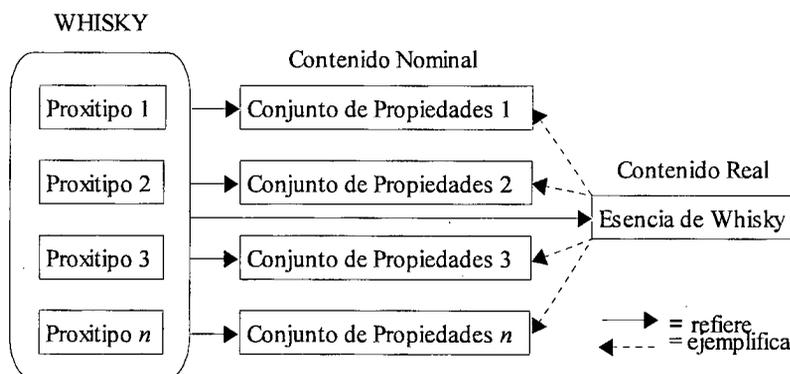


Ilustración 1: Un concepto como un conjunto de proxitipos, tomado de Prinz (2002, p. 280).

Cada conjunto de rasgos que nos permite detectar una categoría al activarse en la memoria de trabajo constituye un proxitipo. Y cada uno de esos conjuntos de rasgos conforma un contenido nominal. En este sentido, cada proxitipo *constituye* un contenido nominal, de manera que “en tanto tipos nominales, los conceptos son proxitipos” (Prinz 2002, p. 279). Ahora bien, los conceptos no pueden identificarse únicamente con el contenido nominal, por lo que la identidad entre proxitipos y conceptos no puede mantenerse cuando se considera a los conceptos *qua* tipos reales: distintos proxitipos pueden corresponder al mismo contenido real y distintos contenidos reales pueden corresponder al mismo proxitipo (Prinz 2002, p. 279). El problema de la individuación evidencia el problema de la relación entre conceptos y proxitipos, ya que muestra que las nociones de proxitipo y concepto no son las mismas, puesto que no hay una correspondencia uno a uno entre ellas. Los conceptos *qua* tipos nominales pueden ser entendidos como idénticos a los proxitipos, pero los conceptos *qua* tipos reales no. En consecuencia, insistir en la identidad entre conceptos y proxitipos implicaría reducir los conceptos a su contenido cognitivo, excluyendo el contenido intencional, y esto sería, para Prinz, ir en contra de su propia semántica de los conceptos.⁴¹

Por otra parte, es cuestionable que Prinz pueda dar cuenta de la individuación de proxitipos en tanto tipos reales. Prinz identifica dos representaciones como manifestaciones del mismo concepto “si están predicativamente ligadas a la misma red de memoria” (Weinberg *et al.* 2003). Esto presupone que la individuación de estas redes es posible, lo cual es cuestionable. Tal como señala Ryder (Weinberg *et al.* 2003), la superposición entre redes, debido a las múltiples asociaciones entre representaciones genera dudas respecto de la posibilidad de identificación de una red correspondiente a un único concepto. Los enlaces que Prinz propone (a partir de los principios de coinstanciación y similitud que vimos en §1) no parecen suficientes para individuar redes de representaciones ligadas.

Una adaptación del ejemplo de Ryder serviría para ilustrar esta crítica. Supongamos que activo en mi memoria de trabajo un proxitipo cuyo contenido nominal es una apariencia de Collie y su contenido real es *perro*. Este proxitipo puede estar asociado en la memoria a largo plazo con múltiples representaciones, por ejemplo, PLAZA,

⁴¹ De Rosa (2005) llega a la misma conclusión a partir de otras consideraciones en torno a la relación entre proxitipos y conceptos.

VETERINARIO, LASSIE y muchos otros proxitipos (además de los de las otras razas de perros). Prinz afirma que el proxitipo de apariencia de Collie deriva del concepto PERRO y no del concepto PLAZA. Según Prinz, esto es así porque la representación de Collie está predicativamente ligada a PERRO, mientras que la relación entre COLLIE y PLAZA se da mediante otro tipo de enlace. Sin embargo, se puede objetar, siguiendo a Ryder, que Prinz está apelando a una diferencia entre enlaces intra-redes e inter-redes, pero el criterio ofrecido para definir los enlaces predicativos -i.e. la disposición a transferir características de una representación a otra- no resulta suficiente para una distinción tan fina. ¿Por qué esta disposición se daría únicamente entre representaciones de la misma categoría? Esto no está justificado en la teoría. Además, los enlaces predicativos se basan en la similitud perceptiva, pero es cuestionable qué características similares tienen, por ejemplo, un chihuahua y un gran danés, además de ser ambos precisamente, perros. Los enlaces predicativos no parecen tener una base puramente perceptiva, con lo cual no sólo no son suficientes para dar cuenta de la individuación de redes, sino que también es difícil fundamentar las relaciones predicativas mismas recurriendo únicamente a RP (Sarnecki 2004).

Frente a estas dificultades, resulta tentador apelar a cierta organización de las representaciones en “archivos mentales” en los cuales las RP se agrupan para constituir conceptos. Prinz utiliza la metáfora de las carpetas de archivos cuando considera la posibilidad de que un mismo contenido cognitivo corresponda a dos conceptos diferentes, sin que haya ninguna RP asociada a alguno de ellos que permita distinguirlos. El ejemplo que propone es el de una persona que cree que hay dos individuos distintos llamados Farrakhan, pero no tiene ninguna información asociada a ninguno de los dos. La explicación propuesta para este caso asigna un rol fundamental a los archivos mentales en los cuales se almacenan las representaciones: “Una persona en esta situación tiene representaciones distintas en la memoria a largo plazo (piense en dos carpetas mentales sin contenido), pero sus contenidos cognitivos son idénticos” (Prinz 2002, p. 271). En este caso, los contenidos cognitivos son idénticos, puesto que no hay RP distintas asociadas a ninguno de los dos conceptos.

El ejemplo anterior, al igual que las dificultades para distinguir entre enlaces intra e inter-redes, aporta razones para sostener que la organización del nivel de memoria a largo

plazo en archivos mentales resulta necesaria para explicar la individuación de conceptos. Pero más aún, en este ejemplo, Prinz admite que hay conceptos distintos almacenados en distintos archivos. Considero entonces que Prinz está aquí recurriendo a un nivel distinto del de los detectores, tal como el de los indicadores que propone el atomismo informacional, el cual como vimos en §1, Prinz considera superfluo. Aunque la alusión a archivos mentales para organizar el nivel de la memoria a largo plazo sea solamente metafórica, la necesidad de apelar a las mismas denota que probablemente el nivel de los indicadores no sea tan superfluo como señalaba el argumento de parsimonia y que algún tipo de indicador sea necesario para individuar conceptos.

2.2 Alcance

El desafío para el neo-empirismo en relación al requisito de alcance consiste en mostrar que se pueden representar distintos tipos de conceptos, entre ellos, los abstractos y teóricos, apelando únicamente a copias de RP. Para esto, Prinz propone diversas estrategias que pueden clasificarse en dos grupos: las que contribuyen a explicar únicamente el contenido cognitivo de los conceptos y las que explican su contenido real.

Entre las primeras, Prinz señala, por un lado, el uso de ciertas habilidades verbales y, por el otro, siguiendo a Lakoff y Johnson (1980), la posibilidad de que los conceptos abstractos sean proyecciones metafóricas a partir de RP básicas relacionadas directamente con la experiencia. Las habilidades verbales, tales como saber en qué contexto usar cada palabra o con qué otras palabras está asociada una palabra en particular, no son suficientes para establecer una relación entre los conceptos y sus referentes en el mundo, ya que únicamente reflejan relaciones entre palabras. Por eso no contribuyen al contenido intencional, sino únicamente al contenido cognitivo y, en este sentido, según Prinz nos permiten razonar acerca de propiedades intangibles. Por otra parte, la proyección metafórica de RP, como Prinz mismo señala, excluye un componente central de los conceptos: su contenido real. Según esta estrategia, se establece una relación de semejanza entre las RP y el concepto en cuestión, pero esta relación no es de equivalencia. De esta manera, las proyecciones metafóricas aproximan el contenido real, pero no lo constituyen. Es decir, esta estrategia, al igual que la apelación a las habilidades verbales, alude al

contenido cognitivo de este tipo de conceptos, pero no da cuenta de cómo estos se relacionan con el mundo –i.e. no da cuenta del contenido real.

Para aplicar la semántica informacional a los conceptos abstractos y teóricos es necesario apelar a una estrategia que los conecte con sus referentes reales. Es decir, para no reducirlos a un puro contenido cognitivo es necesario que las representaciones que constituyen los conceptos como mecanismos de detección estén causalmente conectadas con las propiedades que representan. Para establecer esta conexión, Prinz señala una tercera estrategia: ciertos conceptos están representados por características observables que están asociadas contingentemente a las categorías que representan y que sirven como “rastreadoras de signos”. Habría distintos tipos de rastreo de signos, tales como la identificación de los rasgos superficiales o apariencias de los miembros de una categoría, el uso de instrumentos científicos que sirven como herramientas mediadoras en la percepción, o el rastreo por el uso de palabras del lenguaje natural. Dado que esta es la única estrategia que explica la conexión entre los conceptos abstractos o teóricos y las propiedades en el mundo a las que refieren, si esta estrategia fallara, este tipo de conceptos quedaría reducido a un puro contenido cognitivo. Como última estrategia, Prinz también sugiere que ciertos supuestos contraejemplos no corresponden verdaderamente a conceptos, sino a reglas u operaciones. La negación, por ejemplo, sería una variación de una operación de contraste entre las representaciones y el mundo, “una manera especial de medir similitudes entre conjuntos de rasgos” (Prinz 2002, p. 183).⁴² Considero que estas estrategias no son suficientes para explicar cómo se representan los conceptos teóricos y abstractos por medio de RP. Intentaré mostrar esto tomando como ejemplos a ELECTRÓN y VERDAD.

Para dar cuenta del concepto ELECTRÓN, Prinz divide a las mentes que tienen representaciones de electrones entre mentes científicas y mentes no científicas. En el caso de los científicos, Prinz opta por la estrategia del “rastreo de signos” y sugiere que este concepto se puede representar por medio de RP tales como las de trazos de niebla observables en experimentos científicos al hacer pasar electrones por una cámara que

⁴² No me ocuparé en profundidad de este tipo de conceptos, sin embargo, quiero señalar que Prinz no los trata de manera uniforme, sino que por un lado sugiere que la negación es una operación pero, al mismo tiempo, hace referencia a NO GATO como una representación que tiene un prototipo -i.e. que está estructurada a partir de los rasgos típicos de la categoría (Prinz 2002, p. 287). Si NO fuera sólo un mecanismo especial para medir similitudes, no tendría sentido hacer referencia a un conjunto de rasgos típicos de NO GATO.

contiene vapor de un gas.⁴³ Por otra parte, el concepto ELECTRÓN que posee la comunidad no científica que domina el uso de la palabra “electrón” aun desconociendo los experimentos con cámaras de niebla, se explica a partir de sus habilidades verbales. De esta manera, su concepto es deferente, tiene únicamente contenido cognitivo y refiere en virtud de que hay expertos (los científicos) encargados de “anclar” el contenido del concepto en el mundo.

Ahora bien, esta explicación presenta dos inconvenientes. En primer lugar, para toda la población no científica que nunca oyó hablar de cámaras de niebla, la representación del concepto ELECTRÓN no sería más que la de la palabra “electrón”. En ausencia de otros rasgos perceptivos, ELECTRÓN no tendría por contenido cognitivo más que las asociaciones con otras palabras implicadas en el dominio verbal de la palabra “electrón” y no tendría contenido real en absoluto. En segundo lugar, si en las mentes de los científicos la representación de trazos de niebla sirve de vehículo perceptivo para ELECTRÓN, Prinz no explica cómo puede representarse perceptivamente el vínculo entre los trazos de niebla y los electrones que no es un vínculo de identidad sino de causa-efecto. Y puesto que el concepto mismo de CAUSACIÓN presenta sus propios problemas, es poco plausible que pueda incorporarse con ventajas explicativas al conjunto de RP correspondientes a ELECTRÓN.

Respecto de los conceptos deferentes, su adquisición no es causada por entrar en contacto con las propiedades a las que refieren, sino que se adquieren por medio de representaciones “cuyos significados han sido previamente fijados por otros” (Prinz 2002, p. 252). El contenido intencional de estos conceptos consiste, de esta manera, en el contenido de las representaciones a partir de las cuales fueron adquiridos. En el caso de ELECTRÓN, la mayoría de la gente entra en contacto con este concepto a partir de una definición como “partícula subatómica cargada negativamente”, pero Prinz no pretende sostener un retorno al definicionismo. El contenido intencional de ELECTRÓN no está constituido por una definición, sino que está fijado por los métodos que los científicos utilizan para reconocer apariencias de electrones. Específicamente, los científicos

⁴³ La cámara de niebla contiene vapor de un gas saturado, que se condensa ante perturbaciones mínimas, tales como el aumento de presión o temperatura. Al introducir en la cámara partículas con energía, el gas se ioniza y se condensa alrededor de los núcleos que interactuaron con las partículas. Por esta razón se puede observar un camino o trazo de vapor condensado con una forma característica.

representan perceptivamente los electrones a partir de los trazos que dejan en cámaras de niebla. De este modo, Prinz aparentemente resuelve la primera objeción, es decir, parece evitar el encierro en un círculo de definiciones y relaciones entre palabras, ofreciendo una explicación de la adquisición del contenido real y cognitivo de conceptos como ELECTRÓN. Sin embargo, esta estrategia supone que hay un mecanismo adecuado que conecta las palabras del lenguaje natural con sus referentes y, de esta manera, no resuelve el problema, sino que lo desplaza hacia una teoría de la referencia del lenguaje natural.

Y, concediendo que se adopte, tal como Prinz supone, una teoría semántica del lenguaje natural que remite a los expertos científicos para fijar causalmente la referencia de términos como “electrón”, también hay que suponer que los expertos conectan de manera adecuada su propio concepto ELECTRÓN con su contenido real, y, como señalo en la segunda objeción, esto resulta difícil de cumplir si su concepto está constituido únicamente por RP. En el caso de los científicos que representan ELECTRÓN usando como vehículo las representaciones de los trazos en una cámara de niebla, surge la pregunta por la diferencia entre los conceptos ELECTRÓN y TRAZO DE NIEBLA. El problema es, nuevamente, el de la individuación de conceptos. Si las RP que sirven de vehículo a ambos conceptos son las mismas y rastrean las mismas propiedades en el mundo, entonces ¿en qué sentido se puede sostener que son dos conceptos distintos? Prinz responde a esta pregunta apelando al comportamiento contrafáctico de uno y otro concepto. En mundos posibles en los cuales dispusiéramos de otros mecanismos o técnicas para detectar electrones (por ejemplo, si pudiéramos observar los electrones usando unos lentes especiales), ELECTRÓN tendría vehículos perceptivos diferentes. TRAZO DE NIEBLA, en cambio, tendría el mismo vehículo en otros mundos posibles. Es decir que ELECTRÓN varía contrafácticamente, a diferencia de TRAZO DE NIEBLA. Sin embargo, esta respuesta confirma que en el mundo actual los vehículos perceptivos de ELECTRÓN y TRAZO DE NIEBLA son idénticos.

Sospecho que los científicos, sin embargo, no confunden uno y otro concepto y tampoco los usan indistintamente. Es decir, recurren a algo más que las RP para individuarlos. Veo dos alternativas para explicar qué es lo que distingue a ambos conceptos; ambas resultan problemáticas para la teoría de proxitipos. En primer lugar, podría sostenerse que la diferencia entre uno y otro concepto radica únicamente en las

representaciones de las palabras del lenguaje natural correspondientes a uno y otro. De esta manera, se daría apoyo a la tesis del atomismo amodal respecto de los conceptos, admitiendo que las etiquetas no son prescindibles, sino que sirven como individualadoras. La estrategia de Prinz de incluir las palabras del lenguaje natural en las representaciones conceptuales es una forma de incorporar, de manera encubierta, un sistema de símbolos etiquetadores para los conceptos. Es decir, en el mundo actual, los detectores no son suficientes para la individuación de conceptos sin apelar a un sistema simbólico como el lenguaje natural.

En segundo lugar, podría sostenerse que los científicos no confunden ELECTRÓN y TRAZO DE NIEBLA porque, independientemente de los vehículos del lenguaje natural, se establece una relación causal entre los electrones y los trazos de niebla observables. Los trazos de niebla son el resultado de la interacción entre las partículas de vapor y las partículas introducidas en la cámara. Esto obliga a incorporar en el concepto ELECTRÓN un vínculo o enlace causal que lo conecte con TRAZO DE NIEBLA, pero el enlace en sí no es una RP. Prinz ofrece una explicación de cómo podría representarse perceptivamente el concepto CAUSACIÓN a partir de representaciones de contigüidad y sucesión. Sin embargo, en el caso de ELECTRÓN, las relaciones de contigüidad y sucesión se dan entre entidades imperceptibles y sus consecuencias observables. Es decir que aun si se introdujera con éxito una representación perceptiva de la relación causal entre electrones y trazos de niebla, ésta requeriría una representación de ELECTRÓN (puesto que se necesita representar la relación entre ELECTRÓN y TRAZO DE NIEBLA). De modo que en este caso, la representación perceptiva ELECTRÓN presupone la representación de una relación causal y la representación de la relación causal, a su vez, presupone la representación de ELECTRÓN. Finalmente, es probable que los científicos que diseñaron los experimentos con cámaras de niebla ya pensarán en electrones antes de observar los trazos en dichas cámaras (Weinberg 2003). Esto sugiere que tenían otras representaciones que les permitían pensar acerca de electrones.

Pasando al concepto abstracto de VERDAD, Prinz lo explica apelando a una operación que relaciona o compara expectativas derivadas de nuestras creencias con las representaciones del mundo obtenidas en la experiencia. VERDAD se representaría simulando una relación de correspondencia entre representaciones. Así, se reduce el

concepto VERDAD a una relación de confirmación de creencias por medio de RP. Pero, como señala Sarnecki (2004), VERDAD no es equivalente a CONFIRMACIÓN. La relación de VERDAD podría explicarse como una relación de correspondencia entre creencias/oraciones y el mundo. Sin embargo, en la propuesta de Prinz lo que se estaría contrastando serían creencias/oraciones y RP del mundo, de modo que VERDAD sería así una relación entre creencias y RP. De esta manera, se dejaría afuera, precisamente, la concepción de VERDAD como una correspondencia con algo en el mundo. Por otra parte, se recurre a otros conceptos que difícilmente se puedan representar perceptivamente, tales como el de CORRESPONDENCIA o CORRECCIÓN, y se desplaza de esta manera el problema de la representación perceptiva de un concepto abstracto a la representación de otros conceptos abstractos. ¿Cómo sería una representación perceptiva de la operación de correspondencia? ¿Cómo se representa perceptivamente la correspondencia entre dos RP distintas, o entre una creencia y una experiencia? Hay aún otro riesgo, señalado por Sarnecki (2004): la circularidad. Según la caracterización de Prinz, VERDAD se construiría a partir de la simulación de una operación de correspondencia que le permite a la mente asegurar la corrección de sus creencias. De esta manera, la relación entre VERDAD, CORRESPONDENCIA y CORRECCIÓN resultará circular.

Los ejemplos de aplicación de las estrategias sugeridas por Prinz muestran que la representación de conceptos teóricos y abstractos a partir de RP sigue siendo un desafío para el neo-empirismo, puesto que para ampliar el alcance de la teoría, los proxitipos deben incorporar como elementos constitutivos de los conceptos algo que no es en sí mismo perceptivo, ya sea enlaces entre representaciones o símbolos de un lenguaje natural.

2.3 Composicionalidad

Como se vio en §1, los proxitipos son altamente variables y su activación resulta sensible al contexto. ¿Pueden entonces dar lugar al pensamiento productivo y sistemático? Es un hecho aceptado en ciencias cognitivas que para explicar la productividad y sistematicidad del pensamiento es necesario apelar a sistemas de representaciones que admitan composicionalidad, es decir, sistemas con un conjunto de conceptos simples que admiten combinación para formar otros más complejos, cuya semántica y sintaxis queda

exhaustivamente determinada por la semántica de los conceptos constituyentes y las reglas sintácticas de combinación (Fodor & Pylyshyn 1988, Fodor 1999).

Así, hay acuerdo respecto de que nuestros pensamientos requieren la composicionalidad de los conceptos que los constituyen, pero ésta puede entenderse de distintas maneras. Prinz (2002, en prensa) propone una lectura debilitada, frente a la composicionalidad estricta, que hace hincapié en el aspecto modal de este requisito. Lo que Prinz discute es la interpretación de que la composicionalidad es *la única* forma en que combinamos conceptos. El requisito de composicionalidad en relación a la productividad, solamente establece que debemos tener la capacidad de combinar conceptos de manera composicional, pero no que deba ser ésta la única manera. Respecto de la sistematicidad, sólo es necesario mostrar que se tiene la capacidad de tener ciertos pensamientos si se tienen otros, pero no es necesario tenerlos de hecho. Así, lo que sostiene Prinz es que es suficiente mostrar que los proxitipos se *pueden* combinar de manera composicional, no es necesario que esto sea así en todos los casos de combinación de conceptos.

Prinz (2002) elabora un modelo de combinación de conceptos en tres etapas, el modelo de *Recuperación Composición y Análisis* (RCA), que incorpora distintas estrategias de combinación, entre las cuales se encuentra la composicionalidad. Y para mostrar que este tipo de combinación es posible, defiende la composicionalidad de los prototipos. Los proxitipos no son idénticos a prototipos, pero pueden tener, entre otras, una estructura prototípica, i.e. una estructura que representa las características típicas de una categoría, específicamente las que resultan diagnósticas, sobresalientes, o estadísticamente frecuentes. En especial, los proxitipos por *default* –i.e. los que se activan en ausencia de un contexto específico-, son estructuras estables de tipo prototípico. De manera que el peso de la defensa recae sobre la posibilidad de combinar prototipos.

Como sostendré más adelante, considero que el debilitamiento del requisito es una estrategia adecuada. Sin embargo, aun aceptando esta estrategia, la defensa indirecta de la composicionalidad de los proxitipos a través de la composicionalidad de los prototipos no resulta suficiente. Por un lado, esta defensa no consigue hacer frente a las críticas que se han esgrimido contra la composicionalidad de los prototipos y, por otro lado, el modelo no resulta suficiente para mostrar la composicionalidad de los proxitipos en general. Para que

la defensa indirecta funcione, Prinz debería mostrar o bien que todo proxitipo tiene una parte prototípica o bien que existe un modelo de composición para los proxitipos que no exhiben estructura prototípica, pero no muestra ninguna de estas dos cosas. Además, los argumentos que ofrece Prinz son a favor de la composicionalidad *semántica* de los prototipos, es decir, presentan un modelo de combinación del contenido de los conceptos. Así, Prinz se concentra especialmente en mostrar que la información estadística –*qua* tipo de información y tipo de estructura conceptual- se puede combinar de manera productiva y sistemática. Sin embargo, no se ocupa del aspecto *sintáctico* de la combinación de conceptos, es decir, de la combinación entre los vehículos de los conceptos. Y puesto que el neo-empirismo es una tesis acerca de la naturaleza de los vehículos de los conceptos, para que esta tesis sea viable es importante mostrar que el tipo de vehículo que postula admite la composicionalidad sintáctica. En lo que sigue, analizaré el requisito de composicionalidad debilitado, luego expondré el modelo RCA y finalmente desarrollaré las críticas mencionadas.

Prinz (2002, en prensa) interpreta el requisito de composicionalidad de manera modal. Tanto la productividad como la sistematicidad son capacidades, y la composicionalidad es la mejor explicación para ellas. Esto implica solamente que se debe tener la capacidad de combinar conceptos de modo composicional. Corresponde aclarar que el requisito de composicionalidad no suele entenderse de manera universal como sosteniendo que los conceptos se componen *siempre*. Ni siquiera Fodor y Lepore (1996), tenaces defensores de este requisito, niegan los casos de modismos o expresiones idiomáticas. “Caer en la cuenta” o “al pie de la letra” son expresiones cuyo contenido no depende de la composicionalidad semántica ni de la estructura sintáctica. El contenido de este tipo de expresiones debe adquirirse como un bloque, es decir, como si correspondiera a un concepto léxico. Sin embargo, éste no puede ser el caso general, puesto que de esa manera o bien nuestro sistema representacional debería ser infinito, o bien no tendríamos la capacidad para generar infinitos pensamientos nuevos. Puesto que no se da ninguna de las dos cosas, las expresiones idiomáticas no pueden ser el caso general, sino la excepción. Pues bien, Prinz sostiene que en la mayoría de los casos, la combinación de conceptos recurre a representaciones almacenadas en la memoria y a conocimiento de trasfondo, y que la composicionalidad se aplica únicamente cuando no hay otro conocimiento disponible.

Así, para garantizar la productividad y la sistematicidad es suficiente mostrar que esta estrategia se puede utilizar en ciertos casos.

Hasta aquí, no encuentro objeciones. Sin embargo, el peso de la defensa de esta estrategia de combinación recae en la posibilidad de componer proxitipos y considero que esto presenta diversos inconvenientes. El modelo RCA pretende, por un lado, explicar cómo interviene cada parte de un proxitipo en la combinación de conceptos y, por otro, hacer frente a ciertas críticas a la composición de los prototipos. En especial, Prinz se ocupa de responder a objeciones relativas a las propiedades emergentes de los prototipos. Al combinar ciertos prototipos pueden surgir propiedades nuevas que no estaban en los prototipos iniciales. Por ejemplo CUCHARA GRANDE tiene como rasgo prototípico el ser de madera aunque éste no sea un rasgo típico ni de las cucharas ni de las cosas grandes. Por otro lado, hay casos en los que los prototipos que se combinan pierden propiedades que no son prototípicas del concepto compuesto. Este es el caso de HOMBRE ORQUESTA que es típicamente un hombre con muchos instrumentos (en general un bombo, platillos, una armónica, un acordeón y una trompeta) y cuyo ejemplar típico no se aleja mucho de Bert en *Mary Poppins*. Este concepto complejo hereda de ORQUESTA algo así como el atributo VARIOS INSTRUMENTOS, pero no el atributo COMPUESTA POR VARIOS MÚSICOS. Estos contraejemplos se proponen para mostrar que si se trata a los conceptos compuestos como prototipos, su contenido no puede explicarse por combinación de prototipos de otros conceptos más simples (Fodor y Lepore 1996, Fodor 1999, Connolly *et al.* 2007).

El RCA tiene tres etapas. La primera es la de *recuperación*, en la cual se busca en la memoria información relevante para representar un concepto compuesto. En esta etapa se puede obtener una representación del concepto compuesto de dos formas distintas, ninguna de las cuales apela a la composicionalidad. Por un lado, podrían reactivarse representaciones de ejemplares del concepto compuesto. Esta sería una manera no composicional de arribar a ese concepto, ya que no se parte de conceptos más simples, sino que directamente se activa un ejemplar. Por otro lado, podría realizarse un proceso de cruzamiento de listas de rasgos para ver si se encuentra una representación de un ejemplar que caiga bajo los dos conceptos que se quiere combinar. Por ejemplo, si se quiere representar CUCHARA GRANDE puede buscarse un ejemplar que tenga los rasgos

correspondientes a CUCHARA y el rasgo GRANDE. De esta manera se explica la aparición de rasgos emergentes: estos rasgos provienen de las representaciones particulares de ejemplares en la memoria.

Si esta etapa falla, i.e. si no se encuentran ejemplares del concepto compuesto en la memoria, se pasa a una segunda etapa, la de *composición*. Esta etapa permite combinar conceptos siguiendo reglas de combinación, algunas de las cuales permitirían satisfacer el requisito de composicionalidad. Prinz presenta una heurística para combinar conceptos con estructura prototípica basada en los procesos de alineación e integración, en la que se puede reconocer la influencia del *Modelo de Modificación Selectiva* de Smith *et al.* (1988). En el modelo que Prinz propone, la alineación consiste en poner en paralelo los rasgos comunes (o similares) a los dos conceptos, mientras que la integración consiste en ajustar los valores de esos rasgos comunes a ambos.⁴⁴ Esta etapa podría ser, según Prinz, puramente composicional (sin recurrir a representaciones de ejemplares en la memoria ni a conocimiento de trasfondo), cuestión a la que volveré más adelante.

Por último, hay una etapa de *análisis* en la cual se evalúa si el resultado de la combinación es coherente. Para esto entran en juego el conocimiento de trasfondo y los componentes estructurados como teorías que pueden incluir los proxitipos, “se deben salvar las brechas, se deben extraer inferencias, y se deben eliminar o explicar las inconsistencias aparentes” (Prinz 2002, p. 306). Para ilustrar lo que ocurre en esta etapa, Prinz toma un ejemplo de Murphy que resulta muy esclarecedor: SIERRA ANTIGUA tiene la propiedad emergente OXIDADA. Murphy explica esta propiedad como el resultado de una inferencia estadística: “las sierras son típicamente de metal; los objetos antiguos son viejos; y el metal viejo típicamente está oxidado; de allí que una sierra antigua deba estar oxidada” (Prinz 2002, p. 307). Así, el proceso de análisis también permite explicar las propiedades emergentes, que pueden ser resultado, por ejemplo, de inferencias a partir de conocimiento de trasfondo.

Como intentaré mostrar a continuación, considero que este modelo no resulta suficiente para mostrar que los proxitipos se pueden combinar de modo composicional. La defensa que Prinz ofrece de la composicionalidad de los proxitipos consiste en una defensa

⁴⁴ El modelo supone que los distintos rasgos de las representaciones están ponderados, es decir, tienen valores estadísticos. Veremos esto en más detalle en el próximo capítulo, en relación a la categorización.

indirecta: se ocupa de responder a las objeciones respecto de la composicionalidad de los prototipos. ¿Cómo justifica esta estrategia? Prinz introduce la idea de los proxitipos por *default* para responder al requisito de publicidad. Si pretendemos que los conceptos tengan un papel en la explicación del comportamiento y la comunicación, entonces deben ser compartibles. Para poder predecir, por ejemplo, similitudes en el comportamiento en relación a los perros, es necesario que compartamos el concepto PERRO. Respecto de la comunicación, ésta funciona porque asociamos los mismos conceptos con las palabras, así nos entendemos porque cuando usamos, por ejemplo, la palabra “perro”, la usamos como asociada al mismo concepto PERRO (Prinz 2002, p. 14).

El problema que aparentemente tendrían los proxitipos para satisfacer el requisito de publicidad es que son sensibles al contexto. La manera en que se representa BÓCHA, por ejemplo, será diferente si se trata del dominio que tiene Teodoro jugando a las bochas o de una bocha de helado de chocolate. Los proxitipos por *default* son las representaciones de una categoría que se activarían si no se diera ningún contexto particular. Los rasgos contenidos por los proxitipos por *default* dependen de la frecuencia con la que representamos una categoría como teniendo ese rasgo. Así, el contenido de estos proxitipos está estructurado en rasgos salientes, típicos y diagnósticos ponderados, al igual que las estructuras de los prototipos y, por lo tanto, se tornan blanco de las bien conocidas críticas a la composicionalidad de los prototipos.

Una de estas críticas consiste en señalar casos de conceptos sin prototipo. En ciertos casos, al combinar conceptos que tienen prototipo, el concepto resultante no lo tiene. Fodor (1981) propone ejemplos como CIUDADES AMERICANAS SITUADAS EN LA COSTA ESTE UN POCO AL SUR DE TENNESSE y ABUELAS LA MAYORÍA DE CUYOS NIETOS SON DENTISTAS y otros ejemplos usando operaciones booleanas como ROSA SI ES CUADRADO, NO GATO y NÚMERO NO PRIMO (Fodor & Lepore 1996). Prinz intenta resolver esta objeción de manera sencilla: tener un prototipo es tener una estructura “en grados de typicalidad”, es decir, debe haber instancias que son más típicas de ese concepto que otras. De esta manera, para mostrar que un concepto tiene un prototipo, sólo es necesario mostrar que hay instancias más típicas de ese concepto que otras. Prinz propone entonces que, por ejemplo, para el concepto ABUELAS LA MAYORÍA DE CUYOS NIETOS SON DENTISTAS, una mujer que tiene 8 nietos, de los cuales 6 son

dentistas será un ejemplar más típico de este concepto que una mujer que tiene 5 nietos, de los cuales 3 son dentistas. Y la misma respuesta valdría para NO GATO o para ROSA SI ES UN CUADRADO. Un puma sería un ejemplar mucho menos típico de NO GATO que una mesa, y un cuadrado rosa sería un ejemplar mucho más típico de ROSA SI ES UN CUADRADO que un cilindro verde o una pelota violeta.

Esta respuesta apela a intuiciones. Sin embargo, la pregunta por si este tipo de conceptos presenta una estructura prototípica o “en grados” debería zanjarse empíricamente.⁴⁵ Prinz reconoce esto, pero considera que un breve ejercicio de introspección resulta suficiente para dar un indicio de que estos conceptos presentan una estructura en grados. Sin embargo, considero que las intuiciones respecto de cuestiones empíricas no resultan confiables ni pueden aportar argumentos suficientes a favor de tesis empíricas. Y, aún si decidiéramos conceder la intuición a la que Prinz apela, es cuestionable el tipo de gradación que ofrece para conceptos como los vistos. ¿Qué resulta más típico de ABUELAS LA MAYORÍA DE CUYOS NIETOS SON DENTISTAS? Prinz parece considerar que toda la tipicidad de este concepto pasa por una estructura en grados de MAYORÍA (6/8 resulta una instancia de mayoría más prototípica que 3/5, aunque sólo es una mayoría “mayor” y no por eso necesariamente más prototípica). De manera que no queda claro que conceptos como NO GATO o ABUELAS LA MAYORÍA DE CUYOS NIETOS SON DENTISTAS puedan considerarse como prototipos y menos aún como el resultado de componer otros prototipos.

Incluso en los casos de conceptos complejos respecto de los cuales no se discute que tengan un prototipo, se puede objetar que ese prototipo no es el resultado de la combinación de otros prototipos siguiendo reglas de composición. El requisito de composicionalidad establece que la semántica de los conceptos complejos debe quedar exhaustivamente determinada por los contenidos de los conceptos constituyentes, junto con las reglas de combinación. Se han formulado reglas específicas para la combinación de prototipos (Smith *et al.* 1988). Sin embargo, puede objetarse que las explicaciones de estos modelos, aunque resultan atractivas, siguen sonando un tanto *ad hoc*. ¿Cuántos rasgos tiene un prototipo? ¿Qué pasa si al combinar dos conceptos uno parece modificar un rasgo que el

⁴⁵ Como vimos en el capítulo 1, Fodor tiene exactamente la intuición contraria, i.e. que un puma y una mesa son igualmente típicos de NO GATO (porque ninguno lo es).

otro concepto no tenía, o si no tienen rasgos en común? En palabras de Murphy (2002), la combinación de prototipos parece ser más *interactiva* que lo que estos modelos describen, resultando difícil de reducir a un conjunto de reglas. Así, la defensa de que los prototipos se pueden combinar sin recurrir a información de trasfondo o a representaciones de otro tipo, tales como los ejemplares, no resulta decisiva y deja interrogantes abiertos. De modo que Prinz no logra su objetivo de mostrar que los prototipos se pueden combinar composicionalmente, más allá de algunos casos muy sencillos.

Por otra parte, Prinz enfatiza que los proxitipos son construcciones altamente variables en la memoria de trabajo y que lo que se activa puede ser tanto una estructura prototípica como un modelo mental o una imagen de una palabra del lenguaje natural. Señala que: “Un proxitipo puede ser una representación multimodal detallada, un modelo visual simple, o incluso una representación de una palabra (e.g. una imagen auditiva de la palabra “perro”). El contexto determina qué proxitipo se usa en la memoria de trabajo en cualquier ocasión dada” (Prinz 2002, p. 149). Esto difícilmente sea compatible con la afirmación de que estas estructuras se pueden combinar de manera sistemática. La sistematicidad del pensamiento es la capacidad de tener ciertos pensamientos *si* se tienen otros o, dicho de otra manera, el hecho de que no es posible ser capaz de pensar aRb y no ser capaz de pensar bRa . Como vimos en el capítulo 1, Fodor y Pylyshyn (1988) señalan que la mejor manera de explicar este fenómeno es apelando a la estructura del pensamiento. Pero esto no es todo. También hay que suponer que la contribución que hace a en aRb y en bRa es la misma. Sin embargo, en el caso del neo-empirismo, se trata de representaciones sensibles al contexto, precisamente esta variabilidad consiste en que hacen contribuciones distintas en contextos distintos. Esto trae nuevamente el problema de la individuación de conceptos. Si se pudiera mostrar que dos proxitipos distintos son iguales en tipo, se podría sostener que en ese sentido contribuyen con lo mismo a la representación compleja. Pero, como sostuve en §2.1, esto no es posible. De modo que no se puede explicar la sistematicidad a partir de estructuras como los proxitipos.

Supongamos por mor del argumento de Prinz que cuando se activa una palabra o una estructura de tipo prototípico en la memoria de trabajo, estas representaciones son combinables de manera productiva y sistemática. ¿Qué sucede con el resto de las estructuras que puede presentar un proxitipo? ¿Se pueden combinar de esta manera distintas

creencias o distintos ejemplares? Prinz afirma que el hecho de mostrar que los prototipos pueden satisfacer el requisito de composicionalidad es suficiente para mostrar que los proxitipos también pueden satisfacerlo. Esto no resulta evidente, pero no ofrece un argumento a favor de esa afirmación. Para que esta afirmación sea viable, Prinz debería mostrar que todo proxitipo tiene una parte prototípica, o que otro tipo de estructuras (como los ejemplares o las teorías) admiten composicionalidad, pero no hace ninguna de estas dos cosas, de manera que su defensa no resulta suficiente para mostrar que los proxitipos admiten la composicionalidad.

Mi última objeción está relacionada con los vehículos. La teoría de Prinz omite una explicación respecto de cómo se combinan los vehículos de los conceptos. ¿Las RP tienen propiedades formales sobre las cuales puedan operar reglas de combinación? Se puede esgrimir un argumento según el cual las RP admiten composicionalidad semántica, pero no sintáctica puesto que no exhiben una estructura de constituyentes canónicos, sino tan sólo de partes homogéneas (Fodor 2007, 2008). Así, la contribución que realiza cada parte a la representación completa es del mismo tipo. Es decir, a diferencia de las representaciones *discursivas*, las RP no exhiben predicación ni estructura lógica de ningún tipo. Resulta llamativo que en el neo-empirismo no se trata esta cuestión, siendo una tesis, tal como se afirmó en §1, acerca de los vehículos conceptuales.

3. Conclusiones

En este capítulo sostuve que el neo-empirismo de conceptos, al identificar conceptos con proxitipos, es decir, con conjuntos de RP, tiene dificultades para cumplir con los requisitos de intencionalidad, alcance y composicionalidad. Respecto del contenido, no puede dar cuenta de la individuación de los conceptos sin apelar a un tipo de representación distinta de las RP. Tampoco resulta adecuado para la representación de conceptos abstractos y teóricos, y por ello la teoría tendría un alcance limitado. Por último, no está claro que este tipo de representaciones admita la composicionalidad estricta. Con este panorama, el neo-empirismo no parece prometedor. Sin embargo, en el próximo capítulo veremos algunas de sus ventajas, en relación a los requisitos de adquisición y categorización.

Capítulo 4

El neo-empirismo de conceptos.

La satisfacción de los requisitos de adquisición y categorización

En el capítulo 3 sostuve, que el neo-empirismo tiene dificultades con el requisito de intencionalidad porque no consigue, en particular, dar cuenta de la individuación de conceptos sin incorporar representaciones distintas de las RP -ya sea enlaces entre RP o indicadores amodales-, no puede para dar cuenta de la representación de conceptos abstractos y teóricos apelando únicamente a RP y, por último, tampoco consigue satisfacer el requisito de composicionalidad en el sentido estricto.

En este capítulo desarrollaré la teoría del neo-empirismo de conceptos respecto de los procesos mentales, según la cual pensar consiste en reactivar estados perceptivos. En primer lugar, señalaré que no está claro que esta propuesta pueda dar cuenta de los distintos procesos de pensamiento -fundamentalmente el razonamiento deductivo- prescindiendo de representaciones amodales. En segundo lugar, abordaré la explicación del neo-empirismo de los requisitos de adquisición y categorización. Señalaré las ventajas de estas explicaciones, aunque cuestionaré, en el caso de la categorización, la compatibilidad entre algunos mecanismos cognitivos propuestos y las RP.

1. La Teoría de la Reactivación de Estados Perceptivos

El neo-empirismo de conceptos sostiene que pensar consiste en desplegar nuevamente RP de manera *off-line*, es decir, reactivar RP en ausencia de los estímulos externos que oportunamente las ocasionaron. Activar un proxitipo en la memoria de trabajo equivale así a recrear un estado perceptivo semejante al que se tendría frente a aquello que ese proxitipo representa. Así, los conceptos serían copias de RP y el pensamiento consistiría en la reactivación o simulación de RP conceptuales. Prinz no define en qué consiste ser una copia, sino que sugiere dos alternativas posibles: podría ser una duplicación de la RP original, almacenada en un sistema distinto a los sistemas perceptivos; o podría ser un registro (también en un sistema distinto a los perceptivos) capaz de regenerar las RP en sus sistemas perceptivos originales en ocasiones subsiguientes (Prinz 2002, pp. 108-109).

El formato de las RP es característico de cada uno de los sistemas perceptivos, es decir que las RP son modales. Los conceptos, a su vez, consisten en conjuntos de RP que pueden ser de distintas modalidades. La representación PERRO, por ejemplo, podría incluir RP auditivas de ladridos, RP táctiles y visuales de distintos pelajes, etc. (Prinz 2002, p. 147). Estas RP conceptuales están enlazadas en redes de memoria a largo plazo, y pueden ser reactivadas para detectar categorías o para pensar acerca de esas categorías en ausencia de los estímulos correspondientes. Ahora bien, un concepto sería, de esta manera, una representación multimodal, esto es, un conjunto de RP de distintas modalidades. Según la Teoría de Reactivación de Estados Perceptivos (TREP), pensar consiste en desplegar estos conjuntos de RP. Abordaré esta teoría teniendo en cuenta, en primer lugar, cómo explica la reactivación conjunta de RP de distintas modalidades -y señalaré que para eso necesita recurrir a representaciones amodales- y, en segundo lugar, cómo da cuenta de las AP por medio de simulaciones y cómo se combinan distintos conceptos en una simulación -y señalaré que esto trae dificultades para dar cuenta del razonamiento formal.

En relación a la reactivación de RP de distintas modalidades, al pensar, por ejemplo, que Larry recibió un pastelazo en la cara, podríamos reactivar un conjunto de representaciones que incluya, entre otras, RP táctiles de algo frío y cremoso en la cara, RP visuales de la escena y RP auditivas de risas o quejas. De este modo, los pensamientos no pueden ser modales, puesto que no están limitados a una única modalidad. Deben ser, por lo menos, multimodales. Pero entonces, ¿puede sostenerse que las representaciones que se activan al pensar son solamente las propias de cada sistema de *input*? Para explicar la reactivación simultánea o conjunta de RP de distintas modalidades es necesario apelar a registros que articulen las distintas RP. Prinz recurre para esto a las llamadas zonas de convergencia, noción que toma de Damasio (1989). Estas son poblaciones neuronales que tienen la función de “dirigir” la reactivación de patrones de actividad neuronal de menor nivel (i.e. más cercanos a los sistemas perceptivos, que son periféricos). Las zonas de convergencia tienen la propiedad de guardar registros de la co-ocurrencia de activaciones neuronales en distintos sistemas perceptivos. Así, cuando una población particular de neuronas de este tipo se torna sensible a la co-ocurrencia de un patrón de actividad particular, puede servir para reactivar ese mismo patrón de manera endógena -i.e. bajo control del organismo-, en ausencia del estímulo perceptivo que causa su activación de

manera “automática”. Las zonas de convergencia son, de este modo, las que almacenan registros de los enlaces entre RP de distintas modalidades y controlan la reactivación conjunta de esas RP.

Prinz recurre a las zonas de convergencia con dos objetivos distintos. Por un lado, le permiten explicar el registro de la co-ocurrencia de RP de distintas modalidades y resolver así el problema que se conoce en la literatura como problema del enlace (*binding problem*). Este problema consiste en que cada sentido registra de manera independiente los estímulos de cada modalidad, formando RP separadas y si no hubiera representaciones que integraran las RP de las distintas modalidades, entonces no habría manera de determinar si, por ejemplo, el conjunto de rasgos CUADRADO, VERDE, CÍRCULO, ROSA representa un cuadrado verde y un círculo rosa o un cuadrado rosa y un círculo verde. Por otro lado, Prinz utiliza la evidencia de las zonas de convergencia como evidencia proveniente de las neurociencias a favor de su neo-empirismo, en tanto que apoyaría la tesis de que pensar consiste en reactivar RP. Las zonas de convergencia permitirían explicar, a nivel neuronal, cómo opera el registro y la reactivación de este tipo de representaciones.

Sin embargo, en el nivel representacional, las zonas de convergencia pueden jugar en contra de esta misma tesis neo-empirista, puesto que si bien apoyan la tesis de que al pensar se reactivan RP, ponen en duda que los vehículos del pensamiento sean únicamente este tipo de representaciones. No puede decirse que las zonas de convergencia contengan representaciones modales, puesto que estas zonas guardan registros de la activación de RP de distintas modalidades (Damasio, de hecho, se refiere a ellas como zonas amodales, y más aún, hace mención explícita del LDP). Tal como señala Weiskopf (2007), si las representaciones que se forman en estas poblaciones neuronales calificaran como conceptos, habría que decir que el empirismo incorpora representaciones amodales. Prinz admite que estas áreas podrían calificar como amodales, pero sostiene que no son los vehículos reales del pensamiento, sino que sólo sirven para iniciar la actividad en las distintas áreas de modalidad específica.

Según la teoría de proxitipos, los criterios para que una RP sea un concepto son: (i) estar bajo control endógeno -esto es lo que diferencia a los conceptos de las meras RP, como veremos a propósito de la adquisición- y (ii) ser representaciones que se usan para detectar instancias de categorías. Si las zonas de convergencia satisfacen estos dos criterios,

entonces también deberían ser vehículos del pensamiento. Pues bien, las zonas de convergencia satisfacen (i), puesto que deben activarse ellas mismas en ausencia de los estímulos externos para reactivar las RP modales correspondientes. Respecto de (ii), si las zonas de convergencia son las que permiten recuperar la ocurrencia conjunta de rasgos en un objeto, entonces resultan fundamentales para las tareas de categorización (Weiskopf 2007). En consecuencia, las representaciones de las zonas de convergencia deberían considerarse también como vehículos conceptuales.

Así, si las zonas de convergencia resuelven el problema del enlace, entonces deben ser parte constitutiva de los conceptos. De no ser así, los conceptos serían conjuntos de rasgos inconexos y no habría manera de distinguir entre conceptos que tienen los mismos rasgos pero articulados de distinta manera. Es decir, Prinz debería admitir que estas representaciones son parte constitutiva de los conceptos. Y puesto que estas zonas son amodales, entonces los proxitipos no son exclusivamente modales y, por lo tanto, tampoco los pensamientos. Estas observaciones no atentan contra la tesis de la reactivación de RP en los procesos de pensamiento, sino que apuntan a la necesidad de incluir representaciones amodales en estos procesos.

Por otra parte, Machery (2007) señala un problema metodológico de la TREP, respecto de la evidencia empírica a favor de la reactivación de RP en los procesos de pensamiento. La evidencia empírica que cita Prinz (2002), que proviene de distintas disciplinas científicas, sólo sirve como evidencia en contra de modelos específicos propuestos para tareas específicas. Machery observa, de manera muy aguda, que ni el enfoque modal ni el amodal por sí mismos arrojan ninguna predicción testeable, sino que deben ser complementados con modelos específicos para cada tarea. Por ejemplo, frente a la tarea de listar los rasgos de un concepto, el modelo de combinación de conceptos que adoptan Barsalou *et al.* (1999), en un marco de teoría modal, predice que los individuos listarán más rasgos internos de los conceptos al incorporarse un modificador que los “revele” perceptivamente. Este modelo predice que al listar los rasgos de SANDÍA, por ejemplo, no se incluirán las semillas o el color interno, sino sólo los rasgos externos, mientras que para MEDIA SANDÍA deberían incluirse aspectos internos. Los partidarios del enfoque empirista suponen que los modelos de teorías amodales, en cambio, deberían predecir el mismo comportamiento en ambos casos, i.e. listas de rasgos similares con y sin

el modificador MEDIA. Los resultados obtenidos por Baralou *et al.* en este tipo de estudios confirmaron las predicciones de su modelo. Sin embargo, como señala Machery, estos procedimientos sólo pueden mostrar que las predicciones de un modelo se ajustan mejor que las de otro para explicar el fenómeno que se está observando, pero no refutan a todos los modelos amodales que se proponen para todas las tareas y, por lo tanto, tampoco a las teorías amodales en general.⁴⁶ Machery sostiene, constructivamente, que habría una forma de arribar a resultados más confiables testeando los distintos modelos alternativos de teorías amodales que se formulan para una misma tarea. Si ninguno resultara adecuado, esto daría razones para adoptar una teoría modal. Pero esto todavía no se llevó a cabo, de manera que los resultados no permiten zanjar la cuestión respecto de los vehículos.

En relación al segundo aspecto de la TREP, esto es, cómo se combinan los distintos conceptos en una simulación y cómo se da cuenta de las distintas AP, utilizaré un ejemplo que ofrece Prinz. Boris desea cazar un antílope gordo. ¿Cómo describe la TREP el estado mental de Boris? Prinz (2002, p. 150) afirma que Boris simula la percepción de “una persona con un rifle persiguiendo un animal con forma de antílope con una barriga descomunal”. Como vimos en el capítulo 2 en relación a la TRCM, para poder incorporar una teoría metafísica del pensamiento a las explicaciones intencionales de la psicología, es necesario poder dar cuenta de distintas AP. Prinz (2002, p. 150), al igual que Fodor, apela para esto a su rol funcional, aunque no ofrece una explicación exhaustiva. A grandes rasgos, una creencia es el tipo de simulación que consideramos que está en correspondencia con el mundo, mientras que un deseo es una simulación que nos dispone a buscar o intentar realizar una situación que se corresponda con su contenido.

Por otra parte, es necesario poder diferenciar las distintas simulaciones en virtud de su contenido, para lo cual la explicación de Prinz no es muy precisa. La individuación de los contenidos de la simulación tiene los mismos problemas que vimos en el capítulo anterior en relación a la individuación y a la combinación de conceptos. Por ejemplo, tener un pensamiento proposicional como que los aviones tienen alas, equivaldría a reactivar

⁴⁶ Además, Machery (2006) señala que hay modelos de teorías amodales que pueden acomodar estos resultados. Podrían hacerlo incorporando, por ejemplo, una cláusula pragmática de comunicación, según la cual, al interpretar una expresión compuesta como “media sandía” se asume que el hablante está intentando transmitir la mayor cantidad posible de información específica, de modo que se interpreta que MEDIA SANDÍA no tiene los mismos rasgos que SANDÍA. Un modelo como éste arrojaría la misma predicción que el de Barsalou *et al.*

representaciones perceptivas de un avión y vincularlas con una simulación de alas. Sin embargo, no queda claro cómo esta simulación recoge la idea de que los aviones *tienen* alas, en lugar de que los aviones *son* alas, y resulta dudoso que esta diferencia se pueda establecer sin apelar a un sistema de representaciones lingüístico o simbólico (Machery 2006).

Una individuación de ese tipo resulta necesaria para poder dar cuenta del pensamiento proposicional y de los procesos de razonamiento formal. El problema es que las simulaciones no permiten distinguir distintos tipos de predicados, ya que no es distinguible o “separable” la contribución que hace cada proxitipo a la simulación completa. Para este tipo de individuación sería necesario que las RP tuvieran una estructura sintáctica que, como vimos en el capítulo anterior, no tienen. Como argumenta Fodor (2007, 2008), la relación de predicación se representa adecuadamente por medio de representaciones discursivas, las cuales permiten distinguir constituyentes sintácticos canónicos. Como vimos en los capítulos 1 y 2 en relación al atomismo informacional, las representaciones discursivas tienen una forma lógica que hace explícita la estructura de constituyentes, puesto que explicita cómo contribuye cada parte a la interpretación del símbolo entero. Las RP, en cambio, sólo admiten una distinción de partes, i.e. cada parte contribuye de la misma manera a la interpretación, de modo que no tienen una estructura de constituyentes canónicos y, por lo tanto, no pueden representar las relaciones predicativas.

Prinz intenta explicar la predicación entre RP por medio de un enlace predicativo con redes de memoria. Supongamos que se activa una simulación que consiste en una RP de un insecto de ocho patas caminando por la pared. Esta será una simulación de una araña si está predicativamente ligada a una red de memoria a largo plazo de ARAÑA, es decir, si, como vimos en el capítulo 3, hay una disposición a transferir a esa RP particular las propiedades o rasgos almacenados en la red de memoria de ARAÑA -tales como que teje telas de araña en las cuales atrapa moscas, es silenciosa, algunas son venenosas, etc. Ahora bien, en este caso, el enlace predicativo es equivalente a la relación de pertenencia a una categoría. En el ejemplo de los aviones, en cambio, que los aviones tienen alas no puede implicar un vínculo de predicación equivalente a la pertenencia a una categoría, ya que no se está representando que los aviones pertenecen a la categoría ALAS (ni que las alas pertenecen a la categoría AVIONES). Y se quiere, precisamente, poder distinguir el

pensamiento de que los aviones *tienen* alas del pensamiento de que los aviones *son* alas. De modo que para relacionar AVIÓN y ALAS, hay que recurrir a otro tipo de enlace. Podría ser al enlace de unión, que vimos en el capítulo 3, que permite asociar RP de distintas modalidades. Sin embargo, esta alternativa supone que lo que relaciona o une a las distintas RP en una simulación son enlaces del mismo tipo que los que asocian a las distintas RP en una misma red en la memoria a largo plazo. Pero esto no es algo que se especifique en la teoría de proxitipos, sino que Prinz simplemente sostiene que en una simulación se reactivan de manera conjunta distintas RP que estaban almacenadas en distintas redes de la memoria a largo plazo. Por otra parte, recurrir a distintos enlaces entre RP para explicar cómo se pueden representar de distinto modo SER y TENER, supondría que no hay un proxitipo para SER ni TENER, lo cual significaría que no hay un concepto SER ni un concepto TENER.

Así, para poder distinguir entre el pensamiento de que los aviones tienen alas y el pensamiento de que las alas tienen aviones no es suficiente con la simple asociación de simulaciones de RP (Machery 2006). Esto requiere representaciones que establezcan distinciones más finas en la relación de predicación, tales como las representaciones discursivas. En consecuencia, la individuación de pensamientos proposicionales parece requerir representaciones simbólicas o lingüísticas. Pues bien, Prinz admite que los proxitipos incluyen RP de las palabras del lenguaje natural y que los pensamientos pueden ser puramente verbales. Dado que los proxitipos no admiten forma lógica, para explicar tanto la individuación de pensamientos proposicionales como el razonamiento formal Prinz debe recurrir a una forma de oracionalismo, esto es, un sistema simbólico que permita las transiciones entre formas lógicas. Pero, al incluir las representaciones simbólicas del lenguaje natural, Prinz apela a un sistema de representaciones amodal. De esta manera, Prinz parece negar su propio argumento de parsimonia (introducido en el capítulo 3), según el cual las etiquetas amodales o simbólicas son superfluas.

Las críticas presentadas pretenden mostrar que sin apelar a representaciones amodales, parte de nuestra vida mental quedaría sin explicación. Sin embargo, algunos aspectos de la TREP resultan ventajosos. En especial, esta teoría ofrece una explicación satisfactoria de la adquisición y sugiere un mecanismo plausible de categorización, como se verá en la siguiente sección.

2. Satisfacción de los requisitos

2.1 Adquisición

Una de las mayores ventajas del neo-empirismo es poder explicar la adquisición de conceptos como resultado de la experiencia a partir de las interacciones perceptivas con el entorno. Esta teoría retoma del empirismo clásico la tesis de la prioridad perceptiva y sostiene que los conceptos tienen un origen perceptivo: nada está en el intelecto que no esté antes en los sentidos. De esta manera establece una continuidad entre la percepción y la cognición. La prioridad de la percepción con respecto a la cognición es de tipo causal y establece que “los conceptos no llegarían a existir, si no hubiera RP disponibles para ser copiadas o combinadas” (Prinz 2002, p. 108).

Tal como vimos antes, y en el capítulo 3, los conceptos son copias de RP que se almacenan en redes de memoria a largo plazo y pueden ser reactivadas en la memoria de trabajo. Esto podría sugerir que adquirir RP equivale a adquirir conceptos. Sin embargo, Prinz introduce una distinción entre las meras RP y las RP conceptuales (los conceptos). Tal como vimos en §1, los conceptos son RP (o copias de ellas) que están bajo control endógeno del organismo. Es decir que se ha adquirido un concepto cuando se tiene la capacidad para reactivar de manera autónoma las representaciones que conforman a ese concepto. Las RP están bajo control del entorno en el sentido en que dependen de que se esté en contacto con determinados estímulos para que se activen. Los conceptos, en cambio, se pueden activar por control endógeno, esto es, en ausencia de los estímulos relevantes en el entorno. Podemos, por ejemplo, desplegar o reactivar las RP que necesitamos para planear una fiesta, sin necesidad de tener en el entorno luces de colores, música o invitados. De este modo, puede decirse que la adquisición de conceptos para el neo-empirismo es un pasaje de RP pre-conceptuales a RP conceptuales, y en este sentido, el formato de representación en ambos casos es el mismo.

Integrando las distintas consideraciones que ofrece Prinz, la adquisición de conceptos puede reconstruirse en tres etapas:

1. *Activación de RP*: formación de RP a partir de la interacción con el entorno.
2. *Adquisición de conceptos*: las RP pasan a estar bajo control endógeno.

3. *Aprendizaje*: los conceptos (o RP conceptuales) se van modificando con el tiempo a partir del contacto con instancias de las categorías y con nueva información.

La primera etapa es la de formación de RP. Estas son representaciones modales que se activan o se forman automáticamente cuando estamos en contacto con el entorno. La segunda etapa consiste en el pasaje al control endógeno. En esta etapa aparecen los conceptos. Puede entenderse que se ha adquirido un concepto cuando se almacena una copia de una RP que puede ser luego reactivada en ausencia de los estímulos correspondientes. Prinz (2002, p. 252) se refiere a la adquisición como un proceso instantáneo y afirma que “hay un punto en que un concepto comienza a existir”. Si se interpreta esto como que hay un instante en que una RP (o conjunto de RP) pasa a control endógeno, entonces no se desestima la posibilidad de que en algunos casos se requiera más de un encuentro con la propiedad en cuestión para que las RP pasen a estar bajo control del organismo.

La tercera etapa se introduce para diferenciar, por un lado, el proceso instantáneo de adquisición de, por el otro, las modificaciones posteriores que pueden sufrir los conceptos. Prinz diferencia así adquisición de aprendizaje. Mientras que la adquisición es un proceso instantáneo, el aprendizaje puede ser progresivo y gradual. Los conceptos comienzan a existir cuando cierto conjunto de RP pasa a estar bajo control endógeno, es decir, cuando se adquiere la capacidad de coactivar en la memoria de trabajo un conjunto de RP, almacenadas en la memoria a largo plazo, que corresponden a una determinada categoría. Luego, a partir de encuentros posteriores con instancias de esa misma categoría y de otras, se forman nuevas RP que pueden asociarse al concepto inicial y así ir modificándolo. En esta etapa se crean nuevos enlaces entre representaciones y se pueden modificar los pesos diagnósticos asignados a cada rasgo. Considero que es en esta etapa donde se produce el aprendizaje de los prototipos, el cual requiere el registro de regularidades o tendencias centrales de las categorías por mecanismos inductivos y, por lo tanto, no puede ser instantáneo. En el caso límite en que un concepto se haya adquirido a partir del contacto con, e.g. una liebre, y luego no se aplique nunca a liebres, sino, por ejemplo, a conejos, entonces el concepto LIEBRE se ha transformado y su contenido ya no es su causa

incipiente *liebre*, sino aquello con lo que en efecto covaría nomológicamente, en este ejemplo, *conejo*.

Prinz (2005) sostiene que todas las representaciones conceptuales son aprendidas y no hay representaciones conceptuales innatas. Considero relevante analizar qué significado tienen, en la teoría, estas dos afirmaciones puesto que hay distintos modos de interpretar tanto el innatismo como el aprendizaje. Respecto del innatismo, dado que los conceptos están constituidos por RP, Prinz podría querer decir que no hay ningún tipo de representación innata, ni siquiera las RP. Sin embargo, Prinz (2002) no adopta esta tesis, sino que acepta la posibilidad de que contemos con un equipamiento innato de RP. Esto indica que la afirmación contra el innatismo se refiere a que no hay representaciones *conceptuales* innatas. De modo que si en la primera etapa de la adquisición, esto es, en la etapa puramente perceptiva, entrara en juego una gran carga de RP innatas, esto no atendería contra la tesis empirista, puesto que no implicaría que se tienen conceptos innatos. El neo-empirismo así entendido no requiere oponerse terminantemente a alguna forma de innatismo, siempre y cuando las representaciones innatas no sean representaciones conceptuales.

Ahora bien, ¿qué significa que los conceptos son aprendidos? El aprendizaje podría entenderse como un mecanismo de formulación y testeo de hipótesis. En este caso, la experiencia sería evidencia a favor o en contra de esas hipótesis. Sin embargo, si la adquisición de conceptos consiste en el almacenamiento de RP que pueden ser reactivadas sin necesidad de que estén presentes los estímulos que las causaron, esto no parece incluir un mecanismo semejante a la formulación de hipótesis. El pasaje a control endógeno podría basarse en un mecanismo causal, al igual que el que da origen a las RP que se disparan o activan frente a los estímulos del entorno. En cambio, la etapa de aprendizaje (siguiendo a Prinz 2002), incluiría procesos tales como la formulación y testeo de hipótesis.

De este modo, si la reconstrucción en tres etapas es correcta, la adquisición de conceptos propiamente dicha no es un aprendizaje, sino que se asemeja más al mecanismo de disparo propuesto por Fodor, que hemos visto en el capítulo 2. De hecho, esto puede deberse a que es el tipo de mecanismo consistente con una semántica informacional. Si los conceptos refieren puramente en virtud de relaciones nomológicas causales (en lugar de referir por similitud o por descripción, como podría ser el caso), es lógico que su

adquisición se produzca por un mecanismo causal. La teoría de Prinz resulta innovadora porque propone una explicación de cómo se podría implementar ese mecanismo causal. Los encuentros con instancias de categorías causan que se activen RP en nuestros sistemas de *input* dedicados y hay un mecanismo que causa que se almacenen copias de estas RP para ser reactivadas en ocasiones posteriores. Por otra parte, una vez que se van adquiriendo conceptos, operan los mecanismos de aprendizaje. Se pueden aprender características de los mismos almacenando otras representaciones asociadas e ir “afinando” las categorías ajustando los pesos de los distintos rasgos en las representaciones prototípicas. Así esta teoría tiene la ventaja de explicar el disparo de los conceptos como una consecuencia de la interacción con instancias de la categoría y, con ello, la formación de representaciones en los sistemas perceptivos. Además, el cambio conceptual y el desarrollo de los conceptos se explican por mecanismos de aprendizaje.

Sin embargo, es poco plausible que esta explicación pueda extenderse a todos los conceptos que poseemos, puesto que, como argumenté en el capítulo 3 en relación al requisito de alcance, los conceptos abstractos y teóricos, así como todos los conceptos que son deferentes, no pueden adquirirse por encuentros perceptivos con sus instancias. De modo que la adquisición de este tipo de conceptos debe explicarse de otra manera. En estos casos, los conceptos se adquieren por medio de descripciones o imágenes, cuyo contenido intencional ya ha sido fijado previamente. En algunos casos de conceptos deferentes, esta explicación funciona, puesto que esas imágenes o descripciones están ancladas en el mundo, de modo que al adquirir, por ejemplo, el concepto OLMO por medio de imágenes y descripciones lo que se adquiere es un concepto acerca de una categoría en el mundo, y no solamente IMAGEN DE OLMO. Sin embargo, esta explicación tiene un límite en el caso de los conceptos abstractos y teóricos puesto que, como sostuve en el capítulo anterior, no permite distinguir entre, por ejemplo, ELECTRÓN y las RP de los trazos de niebla.

2.2 Categorización

La explicación de la categorización es otro aspecto atractivo del neo-empirismo. Prinz incluye dos fenómenos distintos como parte de este requisito. Incluye, por un lado, la *producción* de categorías, que es la anticipación, por inducción, de los rasgos que tendrá un elemento si cae bajo una determinada categoría, y, por otro lado, la *identificación* de

categorías, es decir, la determinación de que un elemento cae bajo una categoría. Según la teoría de los proxitipos, estos procesos recurren a distinto tipo de información.

La producción de categorías se efectúa a partir de las representaciones de los proxitipos por *default*, es decir que los rasgos que se anticipan son en general los que típicamente se listan para una categoría cuando no se especifica ningún contexto particular. En caso de que haya más información disponible, la inducción de rasgos puede apelar a esa información y anticipar rasgos que no pertenecen al proxitipo por *default*. Esta explicación, como Prinz (2002) mismo señala, resulta trivial puesto que la producción de categorías y el uso de los proxitipos por *default* consisten en el mismo fenómeno. Prinz apela a los proxitipos por *default* para explicar la publicidad basándose en la intuición de que hay rasgos que generalmente asignamos a las categorías en ausencia de un contexto específico. Estos proxitipos se definen operacionalmente en función, precisamente, de la producción de categorías, de modo que no pueden utilizarse para explicar este fenómeno, sino que más bien forman parte del mismo.

Respecto de la identificación de categorías, el neo-empirismo ofrece, a diferencia del atomismo informacional, herramientas para explicarla.⁴⁷ Prinz propone que los procesos de categorización consisten en una comparación entre el conjunto de rasgos por medio de los cuales representamos al objeto que percibimos (o que escuchamos describir o nombrar) y conjuntos de rasgos de las redes de memoria a largo plazo. Esta comparación se puede modelar por medio de una función que arroja una medida de la similitud entre ambos conjuntos de rasgos. Puesto que los proxitipos incluyen estructuras de tipo prototípico constituidas por un conjunto de rasgos con valores o pesos diagnósticos asignados, la categorización se podría explicar como una función que compara los rasgos de la representación que se quiere categorizar con los de una representación prototípica almacenada en una red de memoria a largo plazo.

Para ilustrar la función de similitud, Prinz recurre a modelos de las teorías de prototipos. Hay distintos modelos teóricos que representan a los prototipos de distintas maneras. Tomaré el mismo que adopta Prinz, que representa a los prototipos como un conjunto de rasgos que se dividen entre atributos y valores. Los valores son los rasgos

⁴⁷ Tal como se vio en el capítulo 2, el atomismo informacional no explica este proceso. Para poder explicarlo, haría falta articular los conceptos que postula con otro tipo de representaciones, tales como los prototipos, y con una explicación de los procesos globales de pensamiento.

posibles que pueden exhibir las instancias de esa categoría para cada uno de los atributos. Por ejemplo, el prototipo de MANZANA puede incluir, entre otros, los atributos COLOR y FORMA, y el atributo COLOR, a su vez, puede incluir los valores ROJA, VERDE y AMARILLA. Según este modelo, todos los valores y atributos están ponderados. En el caso de los atributos, los pesos se distribuyen de acuerdo a la utilidad diagnóstica de cada uno, i.e. en función de cuánto contribuyen en la tarea de reconocer elementos pertenecientes a la categoría. Por ejemplo, el atributo COLOR puede no ser de mucha utilidad a la hora de determinar si algo es un perro, mientras que resulta relevante en el reconocimiento de la arena (si de lo que se trata es de diferenciarla, por ejemplo, de tierra). En el caso de los valores, los pesos se asignan en función de cuán sobresaliente resulta ese rasgo, esto es, en función de la frecuencia con la que ese rasgo ocurre en instancias de ese concepto o en función de cuán perceptible resulta. En el caso del prototipo de MANZANA, puesto que la mayoría de las manzanas que encontramos son rojas, ROJA tendrá un peso mayor que VERDE y éste, a su vez, uno mayor que AMARILLA (Smith *et al.* 1988). La tabla 1 ofrece un ejemplo esquemático de una representación prototípica de MANZANA.

MANZANA			
ATRIBUTO (a)	PESO	VALOR (v)	PESO
COLOR	1	ROJO	25
		VERDE	5
		AMARILLO	
FORMA	0.5	REDONDA	15
		CUADRADA	
		CILÍNDRICA	5
TEXTURA	0.25	SUAVE	25
		RUGOSA	5
		DESIGUAL	

Tabla 1: prototipo de MANZANA, tomado de Smith *et al.* (1988, p.490)

Dado este modelo de representación prototípica, Smith *et al.* adoptan una fórmula de Tversky que arroja una medida de similitud computando tanto los rasgos comunes como los no comunes entre dos representaciones. Para aplicar la fórmula, resulta razonable suponer, siguiendo a Smith *et al.*, que la representación que se quiere categorizar es como una representación de un prototipo, pero sin pesos diagnósticos asignados a los atributos. El peso diagnóstico que se asignará a cada uno de estos atributos para computar la similitud, dependerá de los pesos diagnósticos de cada atributo en la representación prototípica. Cada rasgo -i.e. cada valor-, por otra parte, adopta como peso la suma de los pesos de todos los valores para el atributo al que ese valor corresponde.⁴⁸ Smith *et al.* adoptan la fórmula de contraste de Tversky para medir la similitud entre una representación I de una instancia y un prototipo P.

$$(1) \quad \text{Sim}(I,P) = S_i [af_i (P \cap I) - bf_i (P - I) - cf_i (I - P)]$$

En (1), i se refiere a un determinado atributo, $P \cap I$ determina el conjunto de rasgos que I tiene en común con P para ese atributo, $P - I$ señala el conjunto de rasgos distintivos de P para el atributo i -es decir, rasgos que son característicos del prototipo y no están presentes en la instancia-, mientras que $I - P$ señala el conjunto de rasgos distintivos de I para el atributo i -rasgos presentes en la instancia I y no en P. Por otra parte, f_i multiplica el peso para cada conjunto de rasgos por el peso diagnóstico del atributo i . En el ejemplo, Smith *et al.* asumen que el valor de a , b y c es 1, para simplificar la ecuación, pero se podrían asignar distintos valores a estas variables para otorgar distinta relevancia al conjunto de rasgos en común y a los conjuntos de rasgos distintivos.

Siguiendo con el ejemplo de Smith *et al.*, una instancia A de una manzana verde podría representarse como un conjunto de rasgos ordenados como pares atributo/valor, pero sin pesos. Para comparar esa representación con el prototipo MANZANA, debe asignarse pesos a esos rasgos (valores), en función del peso que tengan en el prototipo, como muestra la tabla 2.

⁴⁸ Esta es una simplificación que este modelo asume. Podría pasar que una representación tuviera en común con un prototipo más de un valor para un determinado atributo y, en ese caso, no tendría sentido asignar a cada uno la suma total de los valores. Este es un aspecto teórico muy específico de las teorías de prototipos que excede el marco de esta tesis, por lo tanto adoptaré el modelo simplificado sin cuestionarlo.

Atributo	Valor	Peso
COLOR	ROJO	30
FORMA	REDONDA	20
TEXTURA	SUAVE	30

Tabla 2: Representación de una manzana roja

La tabla 3, por último, ilustra la medida de similitud que tendría esa representación de MANZANA ROJA con el prototipo MANZANA de la tabla 1. Cada fila de la tabla representa el cómputo de la función para un atributo siguiendo el orden de atributos de la tabla 1 y explica paso por paso la aplicación de la fórmula.

Función Sim (I,P)	Explicación
$1 \cdot (25 - 5 - 5)$	Para el atributo color, I y P tienen el rasgo ROJO en común con un peso de 25, mientras que P tiene como distintivo el rasgo VERDE con un peso de 5 y I tiene como distintivo el rasgo ROJO con un peso de 5 (puesto que el peso para este rasgo en el prototipo es de 25). Esto se multiplica por 1 que es el peso diagnóstico del atributo de color.
$0.50 \cdot (15 - 5 - 5)$	Se aplica el mismo procedimiento para el atributo forma
$0.25 \cdot (25 - 5 - 5)$	y para el atributo textura.
$15 + 2.5 + 3.75$	Se suman los valores obtenidos para cada atributo.
$= 21$	Esta es la medida de similitud entre I y P.

Tabla 3: Medida de la similitud entre una representación de una instancia particular y el prototipo de MANZANA.

Para que la medida de similitud que arrojan este tipo de fórmulas tenga algún significado, es necesario establecer un umbral que las instancias deben superar para caer bajo un determinado prototipo. Por otro lado, para normalizar los resultados, los pesos asignados a cada atributo pueden adoptar valores entre 0 y 1, y la suma de todos los pesos diagnósticos ser igual a 1, de manera que la medida de similitud siempre dé un resultado entre 0 y 1.

La idea de este ejemplo es simplemente ilustrar cómo puede hacerse uso del aparato teórico de algunos modelos elaborados para los prototipos para dar cuenta de un proceso como la categorización, ya que Prinz incorpora una operación de contraste para la identificación de categorías. Pues bien, dado que la teoría de proxitipos cuenta con herramientas como la función de similitud, está en mejores condiciones que el atomismo informacional para explicar la categorización. Sin embargo, la explicación de Prinz exhibe ciertas dificultades. Encuentro principalmente dos objeciones.

En primer lugar, se necesitaría evidencia a favor de que las RP pueden codificar la información de tipo estadístico que hace falta para que operen funciones como la de Tversky. Como alternativa, la teoría de proxitipos podría hacer uso de una operación de contraste como las que se han propuesto en los modelos de categorización de ejemplares (por ejemplo, Medin & Schaffer 1978). Los ejemplares son representaciones de instancias específicas con las que se ha estado en contacto, de modo que no incluyen información de tipo estadístico. Según las teorías que postulan este tipo de representaciones, los conceptos son conjuntos de ejemplares -i.e. una categoría estará representada por medio de un conjunto de representaciones de instancias particulares de la misma. Las funciones propuestas para contrastar instancias y ejemplares almacenados en la memoria asignan valores a los rasgos que estas representaciones tienen en común y a los que no comparten, pero esto es parte del mecanismo de comparación, no es necesario que las representaciones en sí estén ponderadas. Sin embargo, esta solución no parece ser lo que Prinz tiene en mente. Prinz (2002, p. 161) sostiene que la operación que permite identificar categorías “suma los pesos diagnósticos de todos los rasgos que concuerdan con una representación”, de modo que la objeción respecto de la información estadística seguiría en pie. Es posible que la dificultad para ver cómo se podría representar este tipo de información usando RP se deba a una falta de imaginación (Prinz 2002, p. 148), pero Prinz debería ofrecer un argumento o una sugerencia que ayude a la imaginación y no lo hace.

En segundo lugar, no está claro con cuáles RM de las redes de la memoria a largo plazo opera la función de contraste o similitud. Si hubiera que comparar todas las RM hasta encontrar la que mejor resultado arroja, probablemente nos moriríamos de hambre antes de determinar que algo pertenece a la categoría COMESTIBLE (o a alguna categoría específica como TOMATE o PAN). Prinz (2002, p. 162) ofrece como ejemplo que si un

objeto percibido es más similar a una representación en una red de PERRO “que a cualquier otra representación” entonces es identificado como un perro. Esto indica que Prinz está sugiriendo la comparación del objeto percibido con *todas* las representaciones de *todas* las redes, lo cual no suena plausible. Es poco probable que se pueda categorizar un elemento en un tiempo razonable si debe contrastarse su representación con todas las representaciones que se tienen almacenadas.

3. Conclusiones

Como vimos, la TREP no resulta adecuada para dar cuenta de los procesos de razonamiento formales, puesto que las RP no admiten forma lógica. Además, no está claro que se pueda explicar el pensamiento en general prescindiendo por completo de las representaciones amodales. Respecto a la adquisición, la explicación de la teoría resulta atractiva puesto que explica el origen de los conceptos a partir de bases perceptivas sin necesidad de apelar a conceptos innatos y, al mismo tiempo, sin apelar a una relación de evidencia entre los conceptos y la experiencia. Por otra parte, también ofrece una explicación adecuada del aprendizaje de nuevas relaciones conceptuales y la incorporación de nueva información en una etapa posterior a la adquisición. Respecto a la categorización, el neo-empirismo ofrece herramientas para explicar este proceso. Sin embargo, no muestra que estas herramientas sean compatibles con las RP. Además, faltaría una explicación de cómo la función de contraste selecciona las RP a comparar.

Capítulo 5

Propuesta pluralista de conceptos

En los capítulos anteriores vimos que las dos teorías actualmente en debate tienen dificultades para dar cuenta de los requisitos fundamentales para toda teoría de conceptos. Estas dificultades parecen tener su origen en la apelación a un único tipo de estructura o vehículo. En el caso del atomismo conceptual, hice hincapié en la necesidad de relacionar la propuesta de los símbolos amodales con otro tipo de estructura -como los prototipos- para poder articular la teoría de conceptos con las explicaciones psicológicas de la categorización, así como para dar cuenta de lo que en otras teorías se llama contenido cognitivo. En el caso del neo-empirismo de conceptos, sostuve que es necesario incorporar algún tipo de indicador amodal, puesto que con las RP que sirven como detectores no es suficiente para dar cuenta de requisitos fundamentales como la composicionalidad y el alcance. Por otra parte, respecto de la explicación de los procesos de pensamiento, sostuve que ninguna de estas teorías resulta adecuada. Mientras que una propuesta computacional como la de Fodor (la TRCM) tiene dificultades para explicar los procesos que operan sobre propiedades globales -como la categorización-, una propuesta de simulación o reactivación de estados perceptivos como la de Prinz (la TREP) tiene dificultades para individuar pensamientos y para explicar cualquier tipo de razonamiento formal, puesto que no puede dar cuenta de la forma lógica.

Como anticipé en la introducción, el debate en torno a la estructura de los conceptos ha dado origen recientemente a propuestas alternativas eliminativistas (Machery 2009) y pluralistas (Weiskopf 2009a, 2009b). Estas propuestas toman como punto de partida el reconocimiento de las ventajas y limitaciones que ofrecen los distintos tipos de estructuras a la hora de satisfacer cada uno de los requisitos.⁴⁹ Dado que ninguna de las estructuras está en condiciones de satisfacerlos a todos, estas vertientes teóricas proponen conservar las ventajas de cada tipo de estructura para poder así explicar los distintos fenómenos de los cuales se ocupan las teorías de conceptos.

⁴⁹ Me referiré por el momento a “estructuras” en general, aunque más adelante estableceré la distinción entre distintos tipos de estructuras semánticas (como las de rasgos prototípicos o basadas en relaciones causales) y distintos tipos de vehículos (RP o representaciones simbólicas).

En este capítulo esbozaré una propuesta de tipo pluralista. Para esto será necesario, en primer lugar, revisar el criterio de pertenencia a la clase concepto. Sostendré que la mejor estrategia para defender un pluralismo de estructuras consiste en adoptar una noción funcional de concepto que permita abarcar las distintas estructuras propuestas. En segundo lugar, repasaré las razones para sostener que distintos tipos de representaciones resultan adecuados para explicar distintos fenómenos relacionados con los roles funcionales que los conceptos deben satisfacer. Revisaré ciertos estudios empíricos así como ciertos argumentos *a priori* a favor de los distintos tipos de representaciones. Por último, señalaré las ventajas que ofrece una propuesta de tipo pluralista y sugeriré cómo pueden enfrentarse los nuevos desafíos que el pluralismo acarrea, tales como qué tipos de mecanismos pueden operar sobre las distintas estructuras y cómo se determina la activación de una estructura u otra.

1. Los conceptos como una clase funcional

Generalmente, el criterio para determinar qué tipos de entidades son los conceptos es un criterio que apela al rol funcional. Es decir, se establecen una serie de roles que los conceptos deben satisfacer en nuestra vida mental y se postula luego cuál es la entidad que mejor satisface todos esos roles. Las diferencias entre las entidades que postula cada teoría tienen que ver en algunos casos con los roles explicativos seleccionados.⁵⁰ Así, por ejemplo Fodor (1999) establece que los conceptos son las clases de cosas que funcionan como causas y efectos de los procesos mentales, son los que permiten explicar gracias a su composicionalidad la productividad y sistematicidad del pensamiento, y son utilizados como categorías. En este sentido, establece una serie de funciones que los conceptos deben cumplir. Los argumentos que conectan luego estos roles con el LDP pretenden mostrar que lo único que puede satisfacer todas esas funciones de manera conjunta son las RM

⁵⁰ Los roles funcionales no son equivalentes a los *desiderata* de las teorías de conceptos, aunque pueden superponerse. Los roles funcionales son los que los conceptos deben satisfacer, mientras que los *desiderata* deben ser explicados por las teorías. Como sostuve en el capítulo 2, siguiendo a Weinberg, los *desiderata* son fenómenos que la teoría debe explicar, pero no necesariamente apelando sólo a conceptos. En cambio, los roles funcionales que delimitan la clase de los conceptos deben ser satisfechos por los conceptos y no por otro tipo de entidades que las teorías puedan postular más allá de los conceptos. De todas maneras, como sostendré en §3, así como los roles funcionales permiten seleccionar o delimitar la clase de los conceptos, los *desiderata* imponen restricciones sobre el tipo de entidades y mecanismos que una teoría debe postular para poder explicarlos.

simbólicas estructuradas a la manera de los lenguajes naturales. De modo que arriba así a una caracterización ontológica de los conceptos: estos son particulares mentales, físicamente instanciados, que son símbolos de un LDP.

Prinz también adopta un criterio funcional, pero focaliza el rol de los conceptos en la categorización. A pesar de que establece una extensa lista de *desiderata* para la teoría, el criterio que predomina para diferenciar conceptos de otras entidades es el de ser utilizados en la detección de categorías -junto con la condición de estar bajo control endógeno. Este también es un criterio de tipo funcional, ya que establece los roles que los conceptos deben cumplir: deben funcionar como mecanismos detectores de categorías y deben estar bajo control del organismo. Ahora bien, para satisfacer estos roles Prinz también postula particulares mentales, pero de un tipo diferente al que propone Fodor: los proxitipos que, como vimos en el capítulo 3, consisten en copias de RP. El pluralismo de Weiskopf (2009b), a su vez, incluye como roles explicativos de los conceptos el hecho de que son utilizados para categorizar y que admiten la combinación productiva y sistemática en estructuras más extensas y complejas. El eliminativismo de Machery (2009), por otra parte, sostiene que los conceptos son los que se usan por *default* en procesos cognitivos superiores y se incluyen entre estos procesos la categorización, la deducción y la inducción.⁵¹

De este modo los distintos autores fijan un criterio para delimitar la clase de los conceptos, es decir, para determinar qué entidades serán conceptos. Una cuestión diferente, pero relacionada con esta, es si los conceptos constituyen una clase natural. Es decir, una vez establecidos los roles funcionales que los conceptos deben satisfacer y una vez halladas las entidades que los satisfacen, ¿hemos recortado así una clase natural? ¿O hemos delimitado de esta manera un conjunto de entidades mentales análogo a 'cosas que hay en la cocina', el cual no constituye una clase natural? Es fundamental observar que la cuestión del criterio para seleccionar conceptos y la pregunta por si los conceptos constituyen una clase natural son distintas. Todos los autores comienzan por establecer una serie de roles funcionales que los conceptos deben satisfacer y algunos, como Fodor y Prinz, defienden luego que las entidades que ocupan o satisfacen ese rol constituyen una clase natural.

⁵¹ Siguiendo el uso general en psicología, Machery diferencia entre procesos cognitivos inferiores y superiores. Los inferiores incluyen las competencias perceptivas y motrices, mientras que los procesos superiores no toman estímulos perceptivos como *input* y tampoco emiten *outputs* motores.

Pues bien, en relación a la pregunta por si los conceptos constituyen una clase natural, en este caso la cuestión importante es si las entidades agrupadas comparten un conjunto de propiedades relevantes en común, de modo que se puedan establecer generalizaciones fructíferas para la ciencia sobre todo ese conjunto (Boyd 1991, Machery 2009). En relación a este punto, como desarrollaré en §2, los autores que adhieren a un eliminativismo sostienen que las entidades que satisfacen los roles funcionales de los conceptos no comparten propiedades relevantes, por lo que no constituyen una clase natural y, por lo tanto, el término “concepto” debe ser eliminado del vocabulario científico. En cambio, autores, como Fodor o Prinz, que sostienen que los conceptos son una clase natural postulan generalmente un único tipo de entidad con un conjunto de propiedades metafísicas en común y argumentan que por lo tanto están describiendo una clase natural. Entre estas dos posturas puede encontrarse, sin embargo, una tercera, que admite que las entidades que satisfacen esos roles funcionales no constituyen una clase natural -i.e. no se pretende cortar así al mundo “por sus articulaciones”. No obstante, se pueden establecer generalizaciones fructíferas sobre esas entidades, de modo que conservar la noción de concepto es útil para la ciencia. Esta postura sostiene así que los conceptos constituyen una clase funcional y considero que es la que un enfoque pluralista como el que delinearé debe adoptar.

2. El abandono de un único tipo de estructura

En los últimos años, se han propuesto diversos enfoques que se proponen incorporar distintos tipos de estructuras a la clase concepto. La teoría neo-empirista de proxitipos, tal como vimos, es una de ellas. Sin embargo, esta teoría incorpora distintos tipos de estructuras como partes de cada concepto, en lugar de sostener que cada una de ellas constituye un concepto. Por otro lado, postula un único tipo de vehículo -las RP-, lo cual, tal como vimos en el capítulo 3, le genera numerosas dificultades para dar cuenta de los requisitos de intencionalidad, alcance y composicionalidad. Weiskopf (2009b) señala como problemática para las teorías de conceptos la adhesión a dos tesis: la tesis de que los conceptos conforman una clase uniforme de entidades psicológicas -tesis de la uniformidad- y la tesis de que, a grandes rasgos, hay *un* concepto asociado a cada ítem léxico, es decir, que se tiene, por ejemplo, *el* concepto PEZ, *el* concepto VERDE -tesis de la singularidad. En este apartado repasaré propuestas que abandonan tanto la tesis de la

uniformidad como la de singularidad.⁵² En primer lugar, veremos el eliminativismo que propone Machery (2005, 2009) y luego el pluralismo que postula Weiskopf (2009a, 2009b).

Machery (2005, 2009) revisa gran parte de la evidencia que se ha propuesto en la literatura a favor de la existencia de prototipos, ejemplares y teorías -de la cual me ocuparé en §3- y considera que todos estos estudios proveen razones suficientes para sostener la realidad psicológica tanto de los prototipos, como de los ejemplares y las teorías. Sostiene que delimitando la clase de los conceptos a partir del rol funcional que estos deben satisfacer -i.e., tal como señalé en §1, “concepto” es aquello que es utilizado por *default* en los procesos cognitivos superiores-, se seguiría que el dominio de la clase concepto son los prototipos, ejemplares y teorías, puesto que éstas son, según las teorías de la psicología cognitiva, las estructuras sobre las que operan los procesos cognitivos superiores. Pues bien, Machery señala que para constituir una clase natural, los conceptos deberían compartir un conjunto de propiedades científicamente relevantes más allá de las definitorias de la clase y que los prototipos, ejemplares y teorías no las comparten. Dado que estas estructuras no tienen propiedades relevantes en común, por lo tanto no constituyen una clase natural. En consecuencia el término “concepto” debe eliminarse del vocabulario de la psicología y ser reemplazado por los términos correspondientes a cada una de las estructuras, i.e. debe ser reemplazado por “prototipos”, “ejemplares” y “teorías”.

Ahora bien, las diferencias que Machery observa entre estas entidades (y que le permiten concluir un eliminativismo) incluyen el hecho de que tienen estructuras semánticas distintas -i.e. prototipos, ejemplares y teorías están estructurados en función de información de distinto tipo que se organiza de distintas maneras, e.g. como conjuntos de rasgos o como conjuntos de proposiciones- y el hecho de participar en procesos cognitivos distintos. Sin embargo, considero que ambas observaciones resultan insuficientes para dar apoyo a su conclusión. En relación a la primera, podría argumentarse que estas entidades comparten un mismo tipo de vehículo y que esta es una característica en común relevante

⁵² En ciertos pasajes parece que Prinz (2002) abandona la tesis de la singularidad. Por ejemplo: “Si cada proxitipo que puede ser derivado de una red de representaciones de perro califica como un concepto PERRO, entonces todos tenemos incontables conceptos PERRO, incluyendo muchos que no son compartidos por otros individuos” (p. 152). Sin embargo, este tratamiento no es constante, ya que en otros pasajes sostiene que cada proxitipo es una representación distinta del mismo concepto-tipo. Por ejemplo: “En el enfoque que estoy considerando, el mismo concepto estará constituido por representaciones diferentes en distintas ocasiones” (Prinz 2005, p. 6).

más allá de la que define a la clase (no favoreceré una uniformidad de vehículos, pero sería una estrategia posible para bloquear la conclusión eliminativista). Y en relación a los procesos cognitivos en los que intervienen, Machery mismo ofrece ejemplos de procesos que podrían incorporar a estas distintas entidades. En particular, al tratar los procesos de combinación de conceptos, sugiere la adopción de un modelo en el cual “un proceso único que toma distintos tipos de conceptos como *input*, subyace a la creación de conceptos complejos” (Machery 2009, p. 212).⁵³ De allí se sigue que podría haber generalizaciones interesantes para la ciencia, relacionadas con los procesos cognitivos, que podrían formularse sobre prototipos, ejemplares y teorías, a pesar de las diferencias en su estructura semántica.

Además, aun si se acepta que los conceptos no constituyen una clase natural, el eliminativismo no se sigue necesariamente. Puede concederse que los conceptos no constituyen una clase natural, sino una clase funcional, e igualmente sostener que resulta fructífero para la ciencia conservar el término, puesto que de este modo se pueden realizar generalizaciones sobre los distintos elementos de la clase concepto que resultan relevantes para explicar el funcionamiento de la mente. Esta es la postura que adopta Weiskopf (2009b), quien sostiene que dados los roles que los conceptos deberían satisfacer, la única manera de dar cuenta de todos los fenómenos que los conceptos deberían explicar es apelando a un pluralismo de estructuras. Así, concuerda con Machery en que los conceptos no constituyen una clase natural, pero sostiene, sin embargo, tal como se mencionó, que conservar el término “concepto” es útil para la ciencia, puesto que pueden establecerse generalizaciones fructíferas sobre esta clase de entidades.

Weiskopf sostiene que: (i) Hay procesos mentales que dependen de la forma lógica y que operan sobre la clase de los conceptos en general, (ii) La combinación conceptual opera sobre distintos tipos de estructuras, (iii) Los conceptos tienen modos comunes de adquisición y (iv) Los procesos que permiten el almacenamiento en la memoria a largo plazo y la posterior recuperación, manipulan distintos tipos de representaciones. Considero que estos son buenos ejemplos de generalizaciones fructíferas que alientan a rechazar un eliminativismo y a conservar el término “concepto” en la psicología. Sin embargo, señalaré

⁵³ El modelo que Machery sugiere es bastante similar al modelo RCA que propone Prinz y vimos en el capítulo 3.

una consecuencia de (i) con la que Weiskopf debería comprometerse y sugeriré una ligera modificación de (iii). Me ocuparé de todas ellas a continuación, comenzando por la última.

Dado que, en principio, el sistema de memoria a largo plazo en el que se almacenan los distintos tipos de representaciones es uno sólo, resulta necesario que los procesos que permiten almacenar y luego recuperar las representaciones sean capaces de manipular todos los distintos tipos de estructuras que somos capaces de almacenar. Es decir, deberían poder formularse generalizaciones acerca del almacenamiento y la recuperación que abarquen a todos estos tipos de representaciones. Por otra parte, en relación a (iii), Weiskopf sostiene que es posible que dispongamos de mecanismos que nos permiten adquirir distintos tipos de conceptos. Como ejemplo, señala que es habitual explicar la adquisición de prototipos por abstracción a partir de representaciones de ejemplares, o la adquisición de teorías a partir de prototipos, utilizando distintos mecanismos mediados en general por el lenguaje natural. Sin embargo creo que, basándose en estos ejemplos, una forma más clara de expresar esta generalización sería decir que hay mecanismos de adquisición que manipulan distintos tipos de conceptos, lo cual, como sostendré en §4, no significa que estos distintos conceptos se adquieran del mismo modo.

En relación a (ii), Weiskopf ofrece una generalización interesante: es posible combinar la información de distintos conceptos para formar una representación compleja, de modo que el proceso de combinación conceptual debe poder recibir como *input* los distintos tipos de representaciones. Por último, respecto de (i), i.e. la forma lógica, Weiskopf señala que si los pensamientos tienen estructura lógica, los procesos que operan en virtud de ella deben funcionar independientemente del tipo de representaciones específicas que los conforman. Por ejemplo, a partir del pensamiento de que el niño tiene sed y no hay naranjas, podemos inferir el pensamiento de que el niño tiene sed o el pensamiento de que no hay naranjas, independientemente de cómo representemos al niño sediento o a las naranjas. Sin embargo, al sostener que los procesos que operan en virtud de la forma lógica de los pensamientos son “ciegos” al tipo de estructura de cada concepto, considero que Weiskopf se ve forzado a comprometerse con un único tipo de vehículo conceptual, i.e. un vehículo de tipo simbólico puesto que, como vimos en los capítulos 2 y 4, es el único que permitiría las inferencias basadas en la forma lógica. De este modo, Weiskopf debería adoptar una uniformidad de vehículos sobre los cuales sería posible

establecer todas las generalizaciones que señala. Así, su pluralismo sería únicamente semántico.

En lo que sigue revisaré la evidencia y los argumentos a favor de la existencia de distintos tipos de estructuras semánticas, así como también de distintos tipos de vehículos, y presentaré las ventajas explicativas que ofrece adoptar un pluralismo tanto semántico como de vehículos.

3. Requisitos y representaciones

En los capítulos 1 a 4 vimos que algunos requisitos imponen ciertas restricciones sobre el tipo de representaciones o estructuras que permiten explicarlos. Las teorías de Prinz y Fodor son un buen punto de partida para ver este tipo de restricciones, puesto que ambas tienen dificultades por lo menos con alguna de ellas, en particular, con las restricciones que imponen los requisitos de composicionalidad, adquisición y categorización. Como vimos, la categorización sólo puede explicarse apelando a cierta estructura semántica, ya sea una estructura constitutiva de los conceptos (como en el caso del neo-empirismo) o sólo asociada a ellos (como en el caso del atomismo). Esto es así dado que para explicar este proceso es necesario apelar a una estructura de rasgos sobre la que pueda operar una función de similitud o a una estructura de relaciones causales que permita explicar la categorización como una forma de razonamiento -un razonamiento deductivo o una inferencia a la mejor explicación. La composicionalidad, por otra parte, impone una restricción sobre los vehículos, ya que para admitir una combinación productiva y sistemática estos deben tener estructura sintáctica, i.e. deben ser simbólicos. Por otra parte, la adquisición, tal como vimos, impone la incorporación de ciertos MAS -basados en última instancia en mecanismos perceptivos-, para poder explicar cómo nuestras mentes entran en contacto con las propiedades en el mundo. De modo que esto demanda una articulación entre los conceptos y los mecanismos perceptivos, que son los que posibilitan el acceso al mundo y así a los contenidos de los conceptos.

Dadas estas consideraciones, los requisitos mismos parecen apuntar a distintos tipos de representaciones, es decir, parecen sugerir un pluralismo. No sólo el análisis *a priori* de los requisitos sino también la evidencia empírica de distintas investigaciones, que veremos más adelante, parece apuntar en esta dirección. Ahora bien, este podría ser un *pluralismo*

semántico o *de estructuras*, esto es, la propuesta de que la información del contenido de los conceptos puede estar estructurada de distintas maneras (e.g. según rasgos definicionales o estadísticos, y en consecuencia serán respectivamente definiciones o, por ejemplo, prototipos) o un *pluralismo de vehículos*, esto es, la propuesta de que el contenido puede estar vehiculizado por distintos tipos de representaciones -en particular por RP y representaciones simbólicas. Estas dos tesis son independientes, de manera que adoptar una no implica un compromiso con la otra. Es decir, se podría adoptar un pluralismo semántico y una uniformidad de vehículos o un pluralismo de vehículos que tengan la misma estructura semántica.⁵⁴ Denomino a mi propuesta *pluralismo extendido* porque abarca ambos tipos de pluralismo, a favor de los cuales considero que hay tanto argumentos como evidencia empírica.

Respecto al pluralismo semántico, se puede elaborar un argumento a su favor a partir de una inferencia a la mejor explicación que toma en cuenta los fenómenos observados en las investigaciones en torno a los procesos de categorización, los cuales han sido extensamente estudiados por los psicólogos. Los resultados de los distintos experimentos en torno a las tareas de categorización no pueden ser acomodados por un único tipo de estructura. Es decir, no hay un único tipo de estructura que sea la mejor disponible para dar cuenta de todos los fenómenos que se observan en estas tareas (revisiones de estos fenómenos se encuentran, por ejemplo, en Laurence & Margolis 1999, Murphy 2002, Prinz 2002, Machery 2009). De modo que se puede sostener que la mejor explicación de estos fenómenos tomados en conjunto consiste en postular distintos tipos de estructuras semánticas. A continuación abordaré brevemente algunos de los fenómenos que esta tesis permitiría explicar y señalaré también brevemente las dificultades que tienen los distintos enfoques para dar cuenta de ellos.

Los enfoques de prototipos que surgieron en 1970 se proponían entre otras cosas explicar los efectos de tipicalidad. Bajo esta denominación, en general, se engloba a un conjunto de fenómenos observados en relación a la categorización. Por un lado, se observó que los individuos clasifican de manera uniforme distintos objetos como instancias más o menos típicas de una categoría (Rosch 1975) y que hay una correlación entre ese grado de

⁵⁴ Si bien las dos tesis son independientes entre sí, algunas combinaciones particulares entre estructuras semánticas y vehículos podrían no ser posibles o traer dificultades.

tipicalidad y el tiempo de reacción en la categorización -las instancias más típicas de una categoría son categorizadas más rápido que las menos típicas (Rosch & Mervis 1975, Hampton 1997). Los defensores de las teorías de prototipos sostienen que las representaciones estructuradas en función de los rasgos típicos y salientes de una categoría, i.e. los prototipos, son las más adecuadas para dar cuenta de estos fenómenos. Sin embargo, Barsalou (1987) muestra que los juicios de tipicalidad no son uniformes, sino que varían de un individuo a otro y en un mismo individuo en distintos momentos.⁵⁵ De modo que los efectos de tipicalidad constituyen evidencia a favor de los prototipos, aunque no evidencia decisiva.

Por otra parte, otros estudios empíricos muestran efectos inversos a los de tipicalidad. Hay casos en los cuales se categoriza más rápido una instancia atípica de una categoría que una prototípica. Esto se puede explicar apelando a ejemplares (Prinz 2002), ya que se puede atribuir este desempeño a que se está familiarizado con instancias atípicas de una cierta categoría y por lo tanto se las puede clasificar velozmente, de modo que algunos autores denominan a este tipo de fenómenos “efectos de ejemplares” (Machery 2009). Sin embargo, esto puede generar confusión, dado que no son efectos de los ejemplares en sí, sino fenómenos de la categorización que aparentemente requieren ejemplares para ser explicados, de modo que para evitar confusiones y establecer el contraste con los efectos de tipicalidad, me referiré a ellos como “efectos de atipicalidad”.

Los defensores de la teoría-teoría, a su vez, sostienen que los procesos de categorización se explican mejor con modelos de inferencias a la mejor explicación que apelando a cálculos de similitud (tal como sostienen los teóricos de prototipos y ejemplares). Hay distintos modelos correspondientes a este enfoque. Según algunos, los conceptos son términos teóricos que se identifican por los roles conceptuales que cumplen en las teorías en las que figuran (Carey 1991).⁵⁶ Otras versiones postulan que los conceptos están estructurados como modelos causales, estableciendo relaciones de dependencia entre distintas propiedades (Rehder 2003). Así, se supone que los conceptos almacenan información causal acerca de aquello que representan. Este enfoque ofrece estudios empíricos que intentan mostrar que la categorización se explica mejor apelando a un

⁵⁵ Para una breve discusión de los resultados obtenidos por Barsalou, véase Hampton (2006).

⁵⁶ Aunque en algunos casos los autores oscilan entre esta definición y la afirmación de que los conceptos son teorías, i.e. conjuntos de proposiciones (Laurence & Margolis 1999, Machery 2009).

proceso inferencial que opera sobre conocimiento de tipo causal, que recurriendo a una comparación entre aquello que se quiere categorizar y un concepto. (Rehder 2003, Keil 1989, Rehder & Burnett 2005). Por ejemplo, si vemos un animal gris y pequeño chapoteando alrededor de un ganso, podemos categorizarlo como un ganso (aunque no sea similar a las representaciones de gansos de las que disponemos), en virtud de un razonamiento respecto de que las crías son más pequeñas que sus padres y generalmente están cerca de ellos (Murphy 2002). Me referiré a este tipo de fenómenos como “efectos causales”.⁵⁷

Todos estos hallazgos empíricos dan apoyo a la tesis de que no se puede dar cuenta de todos los fenómenos que exhibe la categorización apelando a un único tipo de estructura. Las teorías de prototipos y ejemplares no pueden acomodar los efectos causales, mientras que la teoría-teoría no puede acomodar los efectos de tipicidad y atipicidad. La teoría de prototipos no puede acomodar los efectos de atipicidad, y la teoría de ejemplares debe apelar a una representación de tipo prototípico para acomodar los efectos de tipicidad (la cual se puede incorporar al enfoque de ejemplares por un mecanismo de abstracción de la tendencia central del conjunto de ejemplares almacenados que genere así una representación prototípica). Una vez aceptado esto, hay dos actitudes posibles que se pueden adoptar. Se puede seguir buscando una estructura alternativa que consiga acomodar todos estos fenómenos, o asumir que, por lo menos en el caso de la categorización, esta tarea se resuelve haciendo uso de distintos tipos de estructuras. Esto último sería equivalente a afirmar que la mejor explicación de los fenómenos observados en la categorización requiere postular un pluralismo de estructuras semánticas.

Respecto del pluralismo de vehículos, hay múltiples experimentos que intentan mostrar que los vehículos del pensamiento son perceptivos, pero también hay poderosos argumentos a favor de los vehículos simbólicos. Barsalou y otros han realizado numerosos estudios con la intención de mostrar que los conceptos están estructurados como conjuntos de RP de distintas modalidades. Barsalou *et al.* (2003) se concentraron en la tarea de verificación de propiedades de un concepto. Los sujetos de este estudio debían responder si, por ejemplo, *dulce* es una propiedad de AZÚCAR. Como resultado, registraron un aumento en el tiempo de reacción para verificar propiedades de un concepto cuando había un cambio

⁵⁷ Una buena revisión de este tipo de efectos en la categorización se encuentra en Weiskopf (2009b).

de modalidad con respecto al estímulo anterior -e.g. los sujetos respondían más velozmente que *ruidosa* es una propiedad de LICUADORA si el estímulo inmediatamente anterior había sido el par HOJAS-*crujientes* que si antes habían verificado una propiedad de una modalidad distinta, como en el caso de ARÁNDANO-*ácido*. Barsalou *et al.* (2004) llevaron a cabo un experimento análogo en el cual los sujetos también debían verificar propiedades de un concepto. En este estudio cada concepto se presentaba dos veces, con propiedades de la misma modalidad en los dos casos o con propiedades de modalidades distintas (separados por verificaciones de propiedades de otros conceptos). Arribaron a conclusiones similares, registrando que los porcentajes de error y los tiempos de reacción eran mayores cuando la segunda propiedad correspondía a una modalidad diferente.

Estos estudios ofrecen evidencia a favor de una base perceptiva del pensamiento, por lo menos para ciertos conceptos y ciertas tareas, como la de verificación de propiedades. Sin embargo, como vimos en el capítulo 3 en relación a las dificultades de la teoría de proxitipos con el requisito de alcance, la base perceptiva no consigue acomodar conceptos abstractos y teóricos, sino que resulta adecuada únicamente para conceptos de objetos perceptibles. Por otra parte, y como Barsalou y sus colegas bien admiten, un modelo amodal de representaciones puede acomodar los resultados citados si postula que la información conceptual está organizada espacialmente en el cerebro por modalidades. Como afirma Machery (2006), una teoría amodal podría sostener que la instanciación cerebral de los conceptos que corresponden a propiedades perceptivas está ubicada cerca de los sistemas perceptivos de esa modalidad. Por ejemplo, AMARILLO estaría localizado cerca del sistema visual y AGUDO estaría ubicado cerca del sistema auditivo. De este modo, se podría explicar el fenómeno observado en el estudio de Barsalou *et al.* que sugiere un efecto de *priming* entre activaciones de conceptos de la misma modalidad, porque los conceptos de una misma modalidad estarían más cerca uno del otro que de conceptos de otras modalidades.

Como vimos en el capítulo 4, Machery (2006, 2007) critica que los estudios en torno a los vehículos sólo sirven para descartar modelos específicos diseñados para tareas específicas. Sin embargo, resulta consistente con un enfoque pluralista que un tipo de vehículo sea el más adecuado para resolver ciertas tareas, aunque no para todas. Podría ser que para verificar propiedades perceptivas de los conceptos recurramos a RP de modalidad

específica, como sugieren los estudios de Barsalou *et al.* El problema sería utilizar esta evidencia para descartar las representaciones amodales *tout court*. Lo que propongo, en cambio, es incorporar esta evidencia a favor de que ciertos fenómenos relacionados con ciertas tareas –en particular con la verificación de propiedades– se explican mejor apelando a RP que a representaciones amodales.⁵⁸ Por otra parte, a favor de los vehículos simbólicos amodales, se pueden esgrimir los argumentos que vimos en los capítulos 1 y 2 en torno al tipo de vehículo que admite composicionalidad sintáctica y que permite realizar razonamientos de tipo formal (Fodor 2007, 2008). Pueden señalarse, asimismo, las ventajas que ofrece este tipo de vehículo para la representación de conceptos abstractos y teóricos.⁵⁹

A la luz de estas consideraciones, creo que habría que articular y defender una versión extendida del pluralismo sosteniendo, por un lado, un pluralismo semántico que afirma que los conceptos incluyen prototipos, ejemplares, teorías y etiquetas atómicas y, por otro lado, un pluralismo de vehículos que propone tanto vehículos simbólicos como perceptivos. Adoptando la distinción entre las teorías filosóficas y psicológicas de conceptos que mencioné en la introducción, para explicar los fenómenos que más interés han despertado entre los filósofos, i.e. la intencionalidad y la composicionalidad, los conceptos de tipo simbólico y sin estructura interna resultan ser los más adecuados. Sin embargo, para explicar los fenómenos de los que más extensamente se ocupan los psicólogos, i.e. la categorización y la adquisición, los conceptos estructurados resultan más explicativos. Respecto de la adquisición, considero que los vehículos perceptivos permiten establecer un vínculo entre la percepción y la cognición, de modo que resulta ventajoso incluir este tipo de vehículos en la teoría. Y en relación a la categorización, lo importante para poder dar cuenta de ella es la estructura semántica de los conceptos, cualquiera sea el vehículo que se proponga para transportarla.

Una propuesta de pluralismo extendido resulta atractiva porque promete heredar las ventajas explicativas de las distintas teorías que incorpora. Sin embargo, de la misma manera, presenta el riesgo de heredar todas sus dificultades. En lo que sigue intentaré

⁵⁸ Machery (2010) señala que teóricos amodales como Fodor o Pylyshyn también reconocen el uso de imaginaria para resolver ciertas tareas. Por ejemplo, si nos piden que contemos cuántas ventanas tiene nuestra casa, es probable que activemos una representación visual de nuestra casa y contemos a partir de aquella la cantidad de ventanas.

⁵⁹ Machery (2006) revisa brevemente la evidencia empírica a favor de las representaciones amodales de la cardinalidad de conjuntos de entidades.

señalar las ventajas de esta propuesta y sugerir cómo podría hacer frente a las dificultades, en especial, a las críticas que formulé en los capítulos anteriores respecto del atomismo informacional y el neo-empirismo de conceptos.

4. El poder explicativo del pluralismo extendido

En esta sección intentaré mostrar las ventajas explicativas del esbozo de enfoque pluralista que propongo en relación a cada uno de los requisitos fundamentales. Este no pretende constituir un desarrollo exhaustivo y acabado que explique cómo dar cuenta de cada uno de estos requisitos, sino una sugerencia especulativa de cómo sería posible satisfacerlos desde la propuesta del pluralismo extendido.

Categorización

Empecemos por el requisito que ofrece motivos más claros para adoptar un pluralismo: el requisito de categorización. Los estudios empíricos citados sugieren que no hay un único tipo de estructura que pueda acomodar todos los efectos que se observan en las tareas de categorización, de modo que, para resolver esto, se postulan distintas estructuras. Sin embargo, ¿cómo se explicará qué estructura se utiliza en cada caso? Constituye un desafío de las propuestas pluralistas (y eliminativistas) explicar qué es lo que determina la activación de cada tipo de estructura. Un buen indicio de la que podría ser la clave para responder esta pregunta está en los mismos estudios citados. Las tareas que se proponen a los sujetos en los distintos estudios no son las mismas y no involucran en todos los casos el mismo tipo de categorías. Presumiblemente, la tareas que involucran casos actuales no sean de la misma naturaleza que las que involucran casos hipotéticos. Y, presumiblemente, la categorización de primer orden o de categorías básicas -como MANZANA o ROJO- sea distinta a la de segundo orden, como en el caso de FRUTA o COLOR.

En los estudios de Rosch y Mervis (1975), se preguntaba a los sujetos específicamente por la tipicidad en relación a una categoría de segundo orden para analizar la relación entre la similitud que exhiben distintos miembros de una categoría y su grado de tipicidad -e.g. una manzana es una instancia más típica de FRUTA que una aceituna y una manzana exhibe más rasgos típicos de las frutas que una aceituna. En los

estudios de Keil (1989), en cambio, las tareas involucraban categorías de primer orden, pero en situaciones contrafácticas. Se planteaba a los sujetos una situación en la que un grupo de doctores había realizado ciertas operaciones para cambiar la apariencia de distintos elementos, por ejemplo, habían modificado a un caballo para que pareciera una zebra (pintándolo como las zebras y demás), y lo habían entrenado luego para que se comportara como una zebra (para que comiera pasto en lugar de avena, etc.) y los individuos debían determinar a qué categoría pertenecía el elemento luego de la transformación.⁶⁰

Resulta razonable que para responder a una pregunta por la typicalidad recurramos a representaciones de tipo prototípico, mientras que para categorizar una instancia en el caso hipotético de una transformación utilicemos información de tipo causal o la idea de esencias ocultas. Lo que intento sugerir es que una respuesta posible a qué tipo de estructura se activa para categorizar podría depender del orden o nivel de la categoría en cuestión y de las condiciones específicas de cada tarea.

Intencionalidad

El requisito más difícil de resolver para una propuesta pluralista es el de intencionalidad. En primer lugar, un enfoque de este tipo obliga a desprenderse de la idea tradicional de que tenemos “el concepto PERRO” o “el concepto VERDURA” y a postular en cambio una multiplicidad de conceptos PERRO o VERDURA. Sin embargo, a la hora de explicar el contenido de estos conceptos, es relevante el hecho de que, por ejemplo, el prototipo PERRO_P, la teoría PERRO_T y el símbolo atómico PERRO_A, son todos acerca de la *perritud*. Es decir, todas estas representaciones son acerca de la misma propiedad. Pero, a la vez, están estructuradas a partir de información de distinto tipo, de manera que respecto del contenido cognitivo, estas estructuras no serían equivalentes.

⁶⁰ Estos estudios fueron realizados con niños de distintas edades y con ejemplos de clases naturales y artificiales. Las modificaciones involucraban siempre el cambio de apariencia de un animal a otro animal (como el caso de caballo/zebra) y de un artefacto a otro artefacto (por ejemplo tacho de basura/silla), no se producían pasajes de clase natural a artificial ni viceversa. Una de las predicciones del estudio era que para las clases naturales los niños más pequeños (alrededor de 5 años), que todavía no disponen de teorías, juzgarían que las instancias cambiaron de categoría, mientras que los niños más grandes (entre 7 y 10 años) no juzgarían el cambio de categoría por el cambio de apariencia, mientras que para las clases artificiales todos juzgarían el cambio de categoría. Los resultados confirmaron estas predicciones.

El problema que se debe enfrentar consiste, entonces, en que los distintos tipos de estructuras semánticas traen consigo distintos modos de referir. Un símbolo atómico refiere en virtud de leyes causales que lo conectan con las instancias de la propiedad que representa. Un prototipo, en cambio, refiere a sus instancias por similitud, dado que representa las características estadísticamente más frecuentes de la misma. Estas características no constituyen un conjunto de condiciones necesarias y suficientes para caer bajo la categoría, de modo que un prototipo no refiere únicamente a las instancias que tienen todas esas características, sino que refiere a todas las instancias que superen un *umbral* de similitud con ese conjunto de representaciones. En el caso de los ejemplares, cada uno refiere a una instancia particular de una determinada categoría. Esto puede explicarse apelando a una relación causal entre este ejemplar y la instancia que lo causó o a una relación de similitud entre ambos.⁶¹ En las teorías que postulan únicamente ejemplares, se supone que cada conjunto de ejemplares corresponde a una categoría en general en virtud de la similitud de todas las instancias de la categoría con uno u otro ejemplar almacenado. Las teorías, por otra parte, no son suficientes por sí mismas para dar cuenta del contenido intencional. Si bien se puede sostener que disponemos de conocimiento acerca de relaciones causales y de esencias ocultas de las categorías, las teorías no siempre describen o especifican cuáles son esas esencias y la información causal que codifican no resulta suficiente para individuar categorías. En algunos casos, las teorías especifican las esencias de manera circular, por ejemplo, al identificar a los tigres como aquellos que tienen esencia de tigre y son crías de tigres (Prinz 2002). Esta será información acerca de tigres sólo si se dispone de otra representación que permita fijar la referencia de TIGRE. Esto significa que, aunque las teorías codifican información acerca de la pertenencia a categorías, adoptar únicamente teorías como conceptos no permitiría dar cuenta del requisito de intencionalidad.

Entonces, ¿cómo articular estas distintas maneras de referir y resolver la dificultad de las teorías con este requisito? Un enfoque pluralista estaría proponiendo que hay, por un lado, conceptos amodales que refieren por relaciones causales, por otro lado, prototipos y ejemplares que refieren por similitud de rasgos y, por último, teorías que codifican

⁶¹ Algunos modelos incluyen también información general acerca de todas las instancias de la categoría, la cual se obtiene por abstracción de las representaciones de instancias. Por ejemplo, para el concepto ROPA un rasgo de este tipo sería que 'es algo para vestir' (Smith y Medin 1981).

información sobre las categorías pero que no refieren a ellas por sí mismas (sino en virtud de que incorporan otras representaciones que las conectan con el mundo). Esto parece señalar que la mejor forma de articular estos distintos tipos de contenido es adoptando una teoría semántica de doble factor. Pero, como vimos en relación a la teoría de prototipos en el capítulo 3, esto no está exento de dificultades. Será necesario ofrecer un enfoque de cómo se articulan todos estos tipos de contenido.

Un esbozo de salida posible se puede ilustrar mediante la metáfora de los archivos, a la que me referí en los capítulos anteriores y que es cada vez más utilizada (Prinz 2002, Fodor 2008, Weiskopf 2009a). La idea de esta metáfora es que nuestro sistema conceptual estaría organizado en archivos, cada uno correspondiente a una propiedad. Es decir, los distintos conceptos estarían agrupados o asociados en virtud de referir a la misma propiedad (aunque esta referencia no sea del mismo tipo). Vimos que la semántica informacional permite dar cuenta del contenido intencional, de modo que la relación con el mundo estaría dada por medio de relaciones nomológicas con las etiquetas. Ahora bien, respecto del resto de las estructuras semánticas que permiten almacenar información sobre la pertenencia a categorías, considero que también puede adoptarse una semántica informacional para los ejemplares (ya que su manera de referir puede explicarse en virtud de relaciones causales con sus instancias), pero no para los prototipos ni las teorías. El contenido de estas representaciones no es producto de una mera relación causal, sino evidencial/racional, es decir, constituyen conocimiento acerca de las propiedades. En un sentido, este conocimiento es *acerca de* la propiedad de modo que también puede tomarse como refiriendo a propiedades, aunque no en virtud de relaciones meramente causales (y con la salvedad señalada, de que es preciso que haya representaciones conectadas con el mundo y asociadas a las teorías para que éstas tengan contenido intencional).

De modo que la propuesta que defiendo consiste en individuar conceptos en virtud de su contenido cognitivo -i.e. por cómo representan aquello que representan. Así, por ejemplo, $PERRO_P$ y $PERRO_E$ serían conceptos distintos. Y se diferencian del símbolo $PERRO_A$, en tanto este último no tiene contenido cognitivo, sino únicamente intencional. Por otro lado, los archivos se individuarían en función de la propiedad que representan. En este sentido se puede agrupar a $PERRO_A$, $PERRO_P$ y $PERRO_E$, en virtud de que refieren a la propiedad *perritud*. De esta manera, los archivos serían conjuntos de conceptos que

refieren a la misma propiedad. Estos archivos deben poder incluir tanto distintos tipos de estructura como distintos tipos de vehículo, ya que, como vimos, ciertas tareas conceptuales demandan la activación de RP y otras de símbolos. La organización que propongo agrupa las representaciones que portan información acerca de la misma propiedad, independientemente de su vehículo.

Esta propuesta enfrenta, sin embargo, fundamentalmente dos dificultades. En primer lugar, es necesario ofrecer una explicación de cómo se determinan las asociaciones de los conceptos a cada archivo, lo cual no resulta sencillo. Y, en segundo lugar, si los archivos se individuaban sólo por su relación con el mundo no se podrían explicar los casos Frege. Esto resultaría paradójico, ya que la incorporación de la información en los archivos -que sería contenido cognitivo- se hace en general precisamente para explicar este tipo de casos. Ambas dificultades están relacionadas, como veremos a continuación. El problema es el que enfrentan las teorías de doble factor: cómo articular el contenido cognitivo y el intencional. Respecto al primer punto, una respuesta posible es que las conexiones que asocian conceptos en un mismo archivo son de tipo epistémico, es decir, se basan en que tomamos a esos conceptos *como* refiriendo a la misma propiedad. Así, no sería el mundo el que determina causalmente esas asociaciones, sino nuestras creencias acerca de cómo es el mundo. De este modo, la asociación de conceptos a archivos estaría dada en función de cómo concebimos el mundo.

Esta respuesta puede contribuir a su vez a resolver la segunda dificultad, ya que la única manera de dar cuenta de los casos Frege es apelando a contenidos cognitivos distintos en archivos distintos. De este modo, podría haber dos archivos distintos que refieran a la misma propiedad en el mundo, pero serían archivos distintos porque contienen distintos contenidos cognitivos, es decir, contienen distintos ejemplares, teorías y prototipos. Por ejemplo, para los conceptos SUPERMAN y CLARK KENT, que refieren a la misma propiedad en el mundo, habría dos archivos con las etiquetas S y CK respectivamente. Estas etiquetas están nomológicamente conectadas con la misma propiedad en el mundo. A su vez, cada archivo tendría un conjunto de conceptos distintos, porque tendría ejemplares, prototipos y teorías distintos, cada uno representando distintos aspectos de la categoría. Por ejemplo, el archivo S incluiría un ejemplar de un individuo con un traje azul y los rasgos del prototipo de un superhéroe, mientras que el archivo CK incluiría un ejemplar de un

individuo con anteojos y los rasgos del prototipo de periodista. Por otra parte, habría que agregar asimismo que la apertura de uno u otro archivo en los primeros encuentros con la propiedad también debe depender de factores epistémicos, tales como qué MAS interviene en cada caso -de otro modo, esta propuesta tendría el mismo problema que el atomismo informacional, para el cual, como sostuve en el capítulo 1, no hay manera de explicar cómo una propiedad está nomológicamente relacionada con dos símbolos distintos.

Esta propuesta tiene puntos de contacto con los enfoques de Fodor (2008) y Prinz (2002). Prinz propone una metáfora de archivos sin etiquetas, esto es, estructuras semánticas agrupadas pero sin símbolos amodales. Como vimos en el capítulo 3, muchas de las dificultades de la teoría de proxitipos se resolverían si incorporara representaciones amodales que sirvan de etiquetas para esos archivos. Mi mayor crítica a la satisfacción del requisito de intencionalidad para el neo-empirismo giró en torno a la necesidad de incorporar RM amodales para poder dar cuenta de la individuación de conceptos. Lo que propongo es conservar la metáfora de los archivos, pero incorporando a los mismos las etiquetas amodales del atomismo informacional que permiten dar cuenta de casos en los que se tienen dos conceptos distintos con el mismo contenido cognitivo asociado (como el caso de dos individuos llamados Farrakhan que vimos en el capítulo 3). En estos casos, lo que permitiría diferenciarlos serían dos etiquetas amodales distintas, y la apertura de dos archivos distintos para la misma propiedad en el mundo (el mismo individuo) debería explicarse apelando a la intervención de distintos factores epistémicos en la apertura de uno y otro archivo.

La propuesta fodoriana, por otra parte, consiste en 'asociar' las etiquetas con archivos que almacenan información codificada en mentalés. Esta propuesta, como vimos en el capítulo 2, no permite almacenar información prototípica, ejemplares, ni RP, sino sólo conocimiento de tipo proposicional. De modo que permitiría únicamente incorporar teorías. Los archivos que propongo se diferencian de esta propuesta en tanto que incorporan otros tipos de estructuras y vehículos. Y, fundamentalmente, se diferencia del atomismo informacional, en que todas estas estructuras serían constitutivos de los conceptos y no meras 'asociaciones' a los mismos.

Composicionalidad

En relación a la composicionalidad, considero que se debe entender este requisito de manera debilitada, tal como defiende Prinz (2002) y hemos visto en el capítulo 3. Así, lo que habría que mostrar es que los conceptos pueden combinarse de manera productiva y sistemática, pero no es necesario que siempre sea así. Respecto de este requisito, el pluralismo ofrece una clara ventaja. Las mayores críticas a la composicionalidad de los prototipos, ejemplares y teorías provienen del atomismo informacional. Pero dado que el pluralismo extendido incorpora RM de tipo simbólico, que pueden ser atómicas y análogas a las palabras de un lenguaje natural tal como postula Fodor, puede recurrir a ellas para defenderse de estas críticas.

Por otra parte, el modelo RCA de Prinz es una buena sugerencia de cómo podrían contribuir distintos tipos de estructuras semánticas a la combinación de conceptos. En este modelo, como vimos en el capítulo 3, el uso de las distintas representaciones está organizado en etapas sucesivas que incorporan distintos mecanismos. Adoptar este modelo desde un enfoque pluralista permitiría evitar las críticas que señalé para el neo-empirismo. Éstas objetaban, por un lado, la apelación a partes de los conceptos que no está claro que todos los conceptos tengan y, por otro lado, la falta de vehículos simbólicos que pudieran garantizar la combinación productiva y sistemática. Al adoptar el mismo modelo pero desde un enfoque pluralista estos inconvenientes deberían disolverse, al mismo tiempo que se consigue aquello que Prinz pretendía: que la combinación estrictamente composicional sea un último recurso y no la estrategia general.

En caso de que no tengamos representaciones semánticamente estructuradas para los conceptos que se quiere combinar, el pluralismo extendido dispone de símbolos atómicos amodales. De este modo, para hacer frente a los famosos contraejemplos de Fodor a la composicionalidad de los prototipos, se puede recurrir a representaciones simbólicas. Por ejemplo, para pensar ABUELAS LA MAYORÍA DE CUYOS NIETOS SON DENTISTAS, no resulta plausible que dispongamos de ejemplares, ni prototipos, ni tampoco de conocimiento teórico. Podría haber raras excepciones de individuos que tengan de hecho uno o varios ejemplares de esa categoría y puedan abstraer un prototipo a partir de ellos. Pero, en general, en los casos como este, en los que no hay información de trasfondo

que sirva a la combinación, se puede recurrir a la composicionalidad de etiquetas simbólicas.

Adquisición

A diferencia de Weiskopf, considero plausible que cada tipo de estructura requiera un proceso distinto de adquisición (de manera que la generalización (iii) de §2, como sugerí antes, debería modificarse). Por ejemplo, un encuentro con una instancia de GATO puede resultar suficiente para adquirir un ejemplar, pero seguramente no resultará suficiente para formar un prototipo, para lo cual se requieren varias representaciones a partir de las cuales se puedan abstraer las características generales, así como mecanismos inductivos de generalización. Podría ser que los prototipos se adquieran por abstracción o generalización a partir de conjuntos de ejemplares almacenados. En este sentido, podría acordar con Weiskopf (2009b) en la posibilidad de que un tipo de concepto intervenga en la adquisición de otros tipos de conceptos. Esto podría contribuir a resolver el problema de Fodor respecto de la relación entre prototipos y etiquetas, ya que en lugar de tratarse de un pasaje de representaciones no-conceptuales a representaciones conceptuales, sería un pasaje entre tipos de conceptos.

Sin embargo, no es necesario comprometerse, como hace Fodor, con que la adquisición de los prototipos sea una fase inicial en la adquisición de conceptos, ya que, por un lado, adquirir un prototipo sería ya adquirir un concepto y, por otro lado, no veo la necesidad de adquirir un prototipo antes que un símbolo amodal. El mecanismo de acceso semántico al mundo puede estar dado por un conjunto de RP que no resulten suficientes para abstraer un prototipo, pero sí para abrir un archivo relacionado con una categoría y etiquetarlo con un símbolo. La adopción de un enfoque pluralista permite sostener que se ha adquirido un concepto cuando se abre un archivo para almacenar las representaciones de una categoría en particular y las primeras representaciones que se almacenan pueden ser conceptos de cualquiera de las estructuras que el pluralismo incorpora, siempre y cuando resulten suficientes para realizar procesos cognitivos superiores.⁶² Así, cualquier tipo de concepto sirve para abrir un archivo, en el sentido en que sirve como primera

⁶² Es posible que el desempeño para categorizar, por ejemplo, una instancia de gato bajo GATO, en ciertos contextos, sea peor si sólo se dispone de un ejemplar de gato que si el archivo incluye un prototipo y una teoría, pero esto no debería impedir la utilización de GATO_E en los procesos de categorización.

representación de una propiedad, aunque considero que la apertura de un archivo debe ir acompañada siempre de la asignación de una etiqueta al mismo, que es la que individualiza al archivo y fija su referencia.

Alcance

Este requisito no debería resultar problemático en absoluto para el pluralismo. El enfoque que propongo no exige que para cada categoría dispongamos de un prototipo, un conjunto de ejemplares y una teoría, aunque sí una etiqueta -representación simbólica- que marca la apertura de cada archivo y fija su referencia. De este modo, los conceptos abstractos y teóricos de los cuales no se puede dar cuenta apelando a ejemplares ni a RP, no acarrearán dificultades, puesto que pueden ser representados por medio de representaciones simbólicas.

5. Conclusiones

En este capítulo, retomé las conclusiones de los capítulos anteriores que indicaban que ni el atomismo informacional ni el neo-empirismo podían dar cuenta de todos los requisitos fundamentales. Esto sugiere que adherir a las tesis de la uniformidad y la singularidad -i.e. postular un único tipo de entidad conceptual y una única RM para cada categoría- puede ser un obstáculo para el desarrollo de una teoría adecuada de conceptos. Asimismo, intenté mostrar que los estudios empíricos realizados en torno a distintos procesos cognitivos superiores sugieren que podría haber una multiplicidad de representaciones de distintos tipos, responsables de los distintos fenómenos cognitivos del pensamiento. Estas consideraciones llevaron al esbozo de propuesta de un pluralismo semántico y de vehículos.

La propuesta del pluralismo extendido parece heredar las ventajas explicativas tanto de los distintos tipos de estructuras semánticas como de los distintos tipos de vehículos. Además, he ofrecido sugerencias plausibles de cómo estos distintos tipos de estructuras podrían articularse para dar cuenta de los requisitos fundamentales de una teoría de conceptos. Una propuesta de este tipo acarrea, sin embargo, un riesgo: podría terminar siendo una teoría muy vacua y poco explicativa. Este sería el caso, si no resultara plausible que tengamos todos estos tipos de representaciones. Sin embargo, esta es una cuestión

empírica, de modo que le corresponde a los psicólogos responder, en última instancia, a la cuestión de si disponemos efectivamente de todos estos tipos de representaciones y de qué depende la utilización de uno u otro, más allá de estas sugerencias especulativas.

Conclusiones

El primer objetivo de esta tesis fue reconstruir el debate entre el atomismo informacional y el neo-empirismo de conceptos tanto en torno a la metafísica de los conceptos -sus vehículos y contenido- como a la naturaleza y características de los procesos mentales, de modo de poder evaluar las ventajas y limitaciones de ambas teorías en relación a cada uno de los requisitos fundamentales que una teoría adecuada de conceptos debería satisfacer -i.e., intencionalidad, alcance, composicionalidad, adquisición y categorización. El segundo objetivo fue esbozar una propuesta alternativa que permitiera superar las limitaciones de las teorías analizadas. Así, en los primeros cuatro capítulos intenté mostrar que el atomismo informacional y el neo-empirismo no consiguen acomodar los requisitos fundamentales. En el capítulo 5, sostuve que la adopción de un enfoque pluralista extendido permitiría conservar las ventajas de estas teorías al tiempo que superar sus limitaciones. A continuación reconstruiré el recorrido por los primeros cuatro capítulos, señalando las ventajas y desventajas de cada teoría.

El atomismo informacional, como vimos en el capítulo 1, sostiene que los conceptos son símbolos atómicos de un LDP que adquieren su contenido en virtud de leyes nomológicas que conectan instancias de conceptos con instancias de propiedades en el mundo. Así, nuestro sistema conceptual estaría constituido por símbolos que funcionan como etiquetas asignadas a propiedades y permiten, de este modo, representar el mundo. Argumenté que, dadas estas características, el atomismo informacional resulta especialmente adecuado para dar cuenta de los requisitos de alcance y composicionalidad. En relación al primero, en principio, no habría limitaciones respecto del tipo de conceptos que se pueden representar por medio de un sistema simbólico. Tanto el tipo de vehículo como la semántica que propone el atomismo informacional permiten representar conceptos de una gran variedad, i.e. abstractos o concretos, de clase natural o artefactos.⁶³ En relación a la composicionalidad, este sistema conceptual admite tanto la composicionalidad semántica como la sintáctica, de modo que los contenidos de las representaciones mentales

⁶³ Señalé sin embargo, que el contenido de los conceptos lógicos y matemáticos debe explicarse de un modo distinto al que ofrece la semántica informacional, recurriendo a definiciones. De modo que, para ampliar el alcance de la teoría, la semántica informacional debe complementarse con otra explicación para el contenido de ese tipo de conceptos.

complejas quedan exhaustivamente determinados por los contenidos de los conceptos atómicos constituyentes y las reglas de combinación sintáctica. Respecto de la intencionalidad, en cambio, sostuve que si bien la semántica informacional resulta adecuada para dar cuenta en general del contenido, no es suficiente para explicar la individuación de conceptos coextensivos. Para resolver esto, Fodor apela al formato de las RM, esto es, a su vehículo. Pero, como argumenté, esta solución no resulta suficiente, puesto que, por un lado, la sintaxis no recoge las diferencias epistémicas que permitirían explicar los llamados casos Frege, y, por otro lado, si se intentara apelar a los MAS para explicar estos casos, se estaría estableciendo un vínculo más fuerte entre los MAS y los conceptos que el de la mera mediación. De modo que este tipo de casos constituyen un límite para la teoría.

En el capítulo 2 relacioné la teoría metafísica de los conceptos que propone Fodor con su teoría de los procesos mentales, de modo de poder abordar los requisitos de adquisición y categorización. Según la TRCM los procesos de pensamiento son computacionales, i.e. consisten en transformaciones causales de símbolos gobernadas únicamente por la sintaxis de las RM, es decir por sus propiedades locales. Sostuve que de este modo, la TRCM resulta adecuada para explicar los procesos deductivos de inferencia, pero no puede dar cuenta de los procesos de aprendizaje y categorización que dependen de propiedades globales. Fodor ofrece una explicación causal de la adquisición de conceptos que recurre a los MAS como mediadores entre la mente y el mundo. Sostuve que esta explicación resulta un tanto vaga y deja interrogantes abiertos, en especial, respecto de la relación entre los conceptos y los MAS. Además, sugerí que se podría articular la metafísica fodoriana de los conceptos con explicaciones de la categorización, apelando a procesos de tipo distinto a los que ofrece la TRCM. Para esto sería necesario esclarecer la relación entre los conceptos atómicos y otros tipos de estructuras de información, tales como los prototipos, cuya mediación es necesaria para la adquisición y presumiblemente también para la categorización.

Respecto del neo-empirismo de conceptos, en el capítulo 3 reconstruí las tesis de esta teoría respecto de la metafísica de los conceptos, según la cual estos son proxitipos, cuyos vehículos consisten en copias o combinaciones de copias de RP. En relación a las propiedades semánticas, los proxitipos transportan dos tipos distintos de contenido: un

contenido real, que refiere a propiedades esenciales, y un contenido cognitivo o nominal que consiste en los rasgos por medio de los cuales detectamos categorías en el mundo. Prinz se concentra en la función que cumplen los conceptos para detectar categorías y postula que los proxitipos son construcciones variables y que el contexto determina en cada caso qué conjunto de RP se activará para detectar una determinada categoría. Sostuve que dada esta caracterización de los conceptos, el neo-empirismo tiene dificultades para dar cuenta de los requisitos de intencionalidad, alcance y composicionalidad.

Con respecto a la intencionalidad, esta teoría no puede explicar la individuación de conceptos sin recurrir a un tipo de RM distinta de las RP. En primer lugar, Prinz recurre a los enlaces predicativos para individuar redes de memoria a largo plazo, sin embargo, las distintas redes pueden superponerse y no está claro que este tipo de enlace sea suficiente para individuarlas. En segundo lugar, los enlaces predicativos no parecen tener una base puramente perceptiva. Por último, aun si fuera posible individuar redes, Prinz admite la posibilidad de que haya dos conceptos distintos con el mismo contenido cognitivo. De manera que lo único que permitiría distinguir los conceptos en un caso como éste sería una etiqueta simbólica. Por otro lado, esta teoría también tiene dificultades para articular los dos tipos de contenido que propone, ya que al identificar a los conceptos con proxitipos, los conceptos quedarían reducidos al contenido cognitivo. En relación al alcance, mostré que las RP no resultan suficientes para representar conceptos abstractos y teóricos, ya que sólo permiten dar cuenta de su contenido cognitivo, pero no de su contenido real. En relación a la composicionalidad, sostuve que el neo-empirismo enfrenta numerosas dificultades. Éstas surgen tanto de la estructura atribuida a los proxitipos como de la propuesta de las RP como único vehículo conceptual. En primer lugar, gran parte de la explicación que ofrece Prinz de este requisito se basa en la composicionalidad de los prototipos, la cual, sin embargo, no consigue demostrar. En segundo lugar, aun si se aceptara la composicionalidad de los prototipos, Prinz no muestra que todo proxitipo tenga un prototipo, ni que el resto de las estructuras semánticas que pueden formar parte de un proxitipo (tales como ejemplares y teorías) exhiban composicionalidad. Y, por último, los vehículos perceptivos no admiten el tipo de combinación sintáctica que es necesaria para la composicionalidad del pensamiento, puesto que no admiten estructura en constituyentes canónicos.

En el capítulo 4, al igual que en el capítulo 2 respecto del atomismo informacional, relacioné la metafísica de los conceptos que postula Prinz con su teoría de los procesos mentales. Señalé que la TREP, según la cual pensar consiste en reactivar estados perceptivos, no resulta adecuada para dar cuenta de los procesos de razonamiento formales, puesto que las RP no admiten forma lógica. Sin embargo, señalé que el neo-empirismo ofrece explicaciones atractivas de los procesos de adquisición y categorización. En primer lugar, explica la adquisición de conceptos a partir de la experiencia, sin necesidad de apelar a representaciones conceptuales innatas. Y ofrece distintas sugerencias de cómo se podría conseguir la categorización por medio de funciones que comparan los rasgos de representaciones almacenadas con los rasgos de las representaciones de las instancias que se quiere categorizar. Sin embargo no queda claro que las características de las RP sean compatibles con estas funciones.

La evaluación de estas dos teorías actualmente en debate me permitió extraer, como conclusión principal, que ninguna consigue satisfacer la conjunción de los requisitos fundamentales, pero que ambas ofrecen elementos que resultan explicativos. Así señalé, en el capítulo 5, que esto sugiere que quizás estas teorías, en lugar de oponerse, podrían complementarse.⁶⁴ Las dificultades de estas teorías podrían atribuirse, en parte, a su adhesión a las tesis de uniformidad y singularidad, de modo que propuse abandonar estas tesis y optar por una propuesta pluralista de conceptos, que permita incorporar los elementos ventajosos de distintas teorías. Reconstruí entonces dos propuestas actuales que siguen esta línea de investigación, la propuesta eliminativista de Machery (2009) y la propuesta pluralista de Weiskopf (2009b). Respecto del eliminativismo de Machery señalé que ofrece buenos argumentos para la existencia de prototipos, ejemplares y teorías, pero que su argumento eliminativista no resulta decisivo. Una manera de conservar la noción de concepto sin retornar a las tesis de la homogeneidad y la singularidad, es, tal como propone Weiskopf, aceptar que los conceptos no constituyen una clase natural, sino una clase funcional y mostrar que se pueden establecer generalizaciones relevantes para la ciencia sobre todas las entidades que constituyen la clase concepto. Sin embargo, a diferencia de Weiskopf, propuse un pluralismo tanto de estructuras semánticas como de vehículos.

⁶⁴ Una conclusión similar podría extraerse respecto de teorías como la de prototipos o la teoría-teoría, ya que ninguna consigue acomodar todos los requisitos (para una buena síntesis de las ventajas y desventajas de cada una, véase Laurence & Margolis 1999).

Así, señalé las ventajas que ofrecería adoptar un pluralismo extendido. Este permitiría superar los límites explicativos de las distintas teorías, es decir, permitiría dar cuenta de cada uno de los requisitos fundamentales. Por un lado, la inclusión de prototipos, ejemplares y teorías permitiría dar cuenta de los fenómenos observados en la categorización. La inclusión de símbolos atómicos, por otro lado, garantizaría la composicionalidad en sentido estricto. La adopción de vehículos tanto perceptivos como simbólicos garantizaría a su vez un amplio espectro de representaciones, de modo que no habría dificultades con el alcance. Las representaciones perceptivas ofrecerían un nexo entre percepción y cognición y, a la vez, permitirían explicar los fenómenos que sugieren el uso de la imaginación para resolver, como vimos en el capítulo 5, tareas tales como la verificación de propiedades. Por último, respecto del contenido de los conceptos, señalé que una teoría pluralista deberá ofrecer una articulación de los distintos tipos de contenido que tiene cada concepto, apelando a alguna teoría semántica de doble factor, lo cual constituye un desafío. También señalé que un enfoque pluralista tendría que explicar por qué se activa una u otra representación en cada contexto particular, lo cual deberá responderse mediante investigación psicológica.

Ahora bien, querría concluir con una observación respecto de la posibilidad de formular una teoría unificada de conceptos. Ante la imposibilidad de las distintas teorías de satisfacer todos los requisitos, en el capítulo 5 señalé dos alternativas: continuar buscando una estructura única que permita finalmente explicarlos a todos o abandonar la tesis de la uniformidad y optar o bien por un pluralismo o bien por un eliminativismo de conceptos. Sin embargo, existe una tercera alternativa. Se podría sostener que lo que debe abandonarse son ciertos requisitos. Esta es la estrategia que adopta Fodor respecto del requisito de categorización, como vimos en el capítulo 2. La misma consiste en apelar a una división del trabajo científico y esperar que las distintas explicaciones obtenidas en cada dominio sean luego articulables tanto teórica como empíricamente. Sin embargo, así como una teoría demasiado abarcadora corre los riesgos de resultar vacua o trivial o carecer totalmente de adecuación empírica, una teoría demasiado específica corre el riesgo de ajustarse a sólo uno o algunos de los requisitos. Por ejemplo, así como los fenómenos de productividad y sistematicidad del pensamiento sugieren que contamos con un sistema de símbolos atómicos, los fenómenos observados en la categorización sugieren que contamos con

conceptos estructurados. Si optáramos por dividir los requisitos y “repartirlos” entre la filosofía y la psicología, esta división de requisitos iría en contra del espíritu interdisciplinario de las ciencias cognitivas, el cual no pretende negar la especificidad de cada ámbito de investigación, sino que tiene como meta la integración de ellos en una explicación, en el caso que nos ocupa, del pensamiento.

Por este motivo considero especialmente fructíferas las investigaciones conjuntas de los distintos ámbitos de estudio que las teorías de conceptos abarcan. El recorrido por las distintas tesis del atomismo informacional y el neo-empirismo, así como por los estudios psicológicos en torno a diversas tareas, puso en evidencia la diversidad de disciplinas que intervienen en los estudios sobre conceptos. Cada requisito exige incursionar en algún área de estudio científico o filosófico particular. Por ejemplo, el tratamiento del contenido exige estar familiarizado con teorías semánticas que, en muchos casos provienen de la filosofía del lenguaje, el análisis de los vehículos, en cambio, ofrece la posibilidad de establecer un nexo con las neurociencias, mientras que la categorización nos pone en contacto con estudios psicológicos. Por otra parte, también surgen vinculaciones con la psicología del desarrollo en relación al cambio conceptual y el aprendizaje, así como con la lingüística por argumentos que se construyen por analogía con el poder expresivo de los lenguajes naturales, tales como los que vimos en relación al requisito de composicionalidad.

La situación de las teorías de conceptos es quizás semejante a la de la medicina, donde las distintas especialidades cada vez son más específicas. Sin embargo, siempre hará falta una teoría que integre estos distintos conocimientos tan compartimentados. De lo contrario, se perdería, por ejemplo, la capacidad para derivar a un paciente al especialista correspondiente. No creo que las teorías de conceptos, entendidas en el marco del espíritu integrador mencionado, deban renunciar todavía a la posibilidad de formular una explicación de la metafísica de las representaciones mentales articulada con una teoría de los procesos mentales de pensamiento. Y considero que la adopción de una postura pluralista constituye un primer paso en esa dirección.

Referencias

- Adams, F. & Aizawa, K., (en prensa) "Causal Theories of Mental Content", *The Stanford Encyclopedia of Philosophy* (Spring 2010 Edition), Edward N. Zalta (ed.), URL = <<http://plato.stanford.edu/archives/spr2010/entries/content-causal/>>.
- Aydede, M. (1998) "Fodor on concepts and Frege puzzles", *Pacific Philosophical Quarterly*, 79 (4), pp. 289-294.
- Beck, J. "The Mixed View of Concepts: A Defense" (manuscrito).
- Barsalou, L. (1987) "The Instability of Graded Structure: Implications for the Nature of Concepts" en U. Neisser (ed.), *Concepts and Conceptual Development: Ecological and Intellectual Factors in Categorization*, Cambridge, Cambridge University Press.
- Barsalou, L. (1999) "Perceptual Symbol Systems", *Behavioral & Brain Sciences*, 22, pp. 577-660.
- Barsalou, L., Solomon, K. & Wu, L. (1999) "Perceptual simulation in conceptual tasks", en M. Hiraga, C. Sinha, & S. Wilcox (Eds.), *Cultural, typological, and psychological perspectives in cognitive linguistics: The proceedings of the 4th conference of the International Cognitive Linguistics Association, Vol. 3*, Amsterdam, John Benjamins (pp. 209-228).
- Barsalou, L., Pecher, D. & Zeelenberg, R. (2003) "Verifying different-modality properties for concepts produces switching costs", *Psychological Science*, 14 (2), pp. 119-124.
- Barsalou, L., Pecher, D. & Zeelenberg, R. (2004) "Sensorimotor simulations underlie conceptual representations: Modality-specific effects of prior activation", *Psychonomic Bulletin & Review*, 11 (1), pp. 164-167.
- Block, N. (1986) "Advertisement for a Semantics for Psychology", en P. French, T. Uehling, & H. Wettstein (Eds.) (1986) *Midwest Studies in Philosophy*, vol. X, *Studies in the Philosophy of Mind*. Minneapolis, University of Minnesota Press.
- Boyd, R. (1991) "Realism, anti-foundationalism and the enthusiasm for natural kinds", *Philosophical Studies*, 61, 127-148.
- Carey, S. (1991) "Knowledge Acquisition: Enrichment or Conceptual Change?", en E. Margolis & S. Laurence (eds.) (1999) *Concepts: Core readings*, Cambridge, MA, MIT Press.

- Connolly, A., Fodor, J., Gleitman, L. & Gleitman, H. (2007) "Why stereotypes don't even make good defaults", *Cognition* 103 (1), pp. 1-22.
- Damasio, A. (1989) "Time-locked multiregional retroactivation: A systems-level proposal for the neural substrates of recall and recognition", en P. Eimas & A. Galaburda (eds.) (1990) *Neurobiology of cognition*, Cambridge, MA, MIT Press.
- De Rosa, R. (2005) "Prinz's Problematic Proxytypes", *The Philosophical Quarterly*, 55 (221), 594-606.
- Dretske, F. (1981) *Knowledge and the flow of information*, Cambridge, MA, MIT Press.
- Fodor, J. (1981) "The present status of the innateness controversy" en *Representations: Philosophical Essays on the Foundations of Cognitive Science*, Cambridge, MA, MIT Press.
- Fodor, J. (1983) *Modularity of Mind*, Cambridge, MA, MIT Press.
- Fodor, J. (1984) *El lenguaje del pensamiento*, Madrid, Alianza.
- Fodor, J. (1987) *Psychosemantics*, Cambridge, MA, MIT Press. [*Psicosemántica*, Tecnos, Madrid, 1994].
- Fodor, J. (1990a) "A theory of Content, I: The Problem" en J. Fodor *A Theory of Content and Other Essays*, Cambridge, MA, MIT Press.
- Fodor, J. (1990b) "A Theory of Content, II: The Theory" en J. Fodor *A Theory of Content and Other Essays*, Cambridge, MA, MIT Press.
- Fodor, J. (1990c) "Information and Representation" en E. Margolis & S. Laurence (eds.) (1999) *Concepts: Core readings*, Cambridge, MA, MIT Press.
- Fodor, J. (1994) *The Elm and the Expert: Mentalese and its semantics*, Cambridge, MA, MIT Press.
- Fodor, J. (1999) *Conceptos*, Barcelona, Gedisa.
- Fodor, J. (2000) "Replies to Critics", *Mind & Language*, 15 (2&3), 350-374.
- Fodor, J. (2001) "Doing without what's within: Fiona Cowie's critique of nativism", *Mind*, 110, pp. 99-148.
- Fodor, J. (2003) *La mente no funciona así*, Madrid, Siglo XXI.
- Fodor, J. (2004) "Reply to Commentators", *Mind & Language* 19 (1), 99-112.
- Fodor, J. (2007) "Revenge of the given", en B. McLaughlin & J. Cohen (eds.) *Contemporary Debates in Philosophy of Mind*, Oxford, Blackwell.

- Fodor, J. (2008) *LOT 2*, Oxford, Clarendon Press.
- Fodor, J. & Lepore, E. (1996) "The Red Herring and the Pet Fish: Why Concepts Still Can't Be Prototypes", *Cognition* 58, pp. 253-270.
- Fodor, J. & Pylyshyn, Z., (1988) "Connectionism and Cognitive Architecture: A Critical Analysis", en C. Macdonald & G. Macdonald (eds.) (1995) *Connectionism: Debates on Psychological Explanation*, Vol. II, Oxford, Basil Blackwell.
- Frege, G. (1892) "Sobre el sentido y la denotación" en T. Moro Simpson (ed.) (1969) *Semántica Filosófica: problemas y discusiones*, Buenos Aires, Siglo XXI.
- Gopnik, A. & Meltzoff, A. (1997) *Words, Thoughts, and Theories*, Cambridge, MA, MIT Press.
- Gopnik, A. & Schulz, L. (eds.) (2007) *Causal learning: Psychology, philosophy, and computation*, New York, Oxford University Press.
- Haimovici, S. (2009a) "Un problema para la teoría fodoriana de la individuación de conceptos" en *Filosofía teórica, Metafilosofía o Postfilosofía*, Córdoba, Universidad Nacional de Córdoba, pp. 223-230.
- Haimovici, S. (2009b) "Cómo organizar las representaciones perceptivas en una mente empirista: Un análisis de la individuación de conceptos en la teoría de Prinz" en D. Letzen & P. Lodeyro (eds.) (2009) *Epistemología e Historia de la Ciencia: Selección de trabajos de las XIX Jornadas*, Córdoba, Universidad Nacional de Córdoba, pp. 261-265.
- Hampton, J. (1997) "Associative and similarity-based processes in categorization decisions", *Memory & Cognition*, 25, pp. 625-640.
- Hampton, J. (2006) "Concepts as prototypes" en Ross B. (ed.) *The psychology of learning and motivation: Advances in research and theory*, New York, Academic Press, 46, pp. 79-113.
- Hampton, J. & Jönsson, M. (2008) "On prototypes as defaults (Comment on Connolly, Fodor, Gleitman and Gleitman, 2007)", *Cognition*, 106 (2), pp. 913-923.
- Keil, F. (1989) *Concepts, kinds and cognitive development*, Cambridge, MIT Press.
- Landau, B. (2000) "Concepts, the Lexicon and Acquisition: Fodor's New Challenge" *Mind & Language*, 15 (2&3), pp. 319-326.
- Lakoff, G. & Johnson, M. (1980) *Metaphors We Live By*, Chicago, University of Chicago Press [*Metáforas de la vida cotidiana*, Madrid, Cátedra, 1986]

- Laurence, S. & Margolis, E. (1999) "Concepts and cognitive science", en E. Margolis & S. Laurence (eds.) (1999) *Concepts: Core readings*, Cambridge, MA, MIT Press.
- Machery, E. (2005) "Concepts are not a natural kind", *Philosophy of Science*, 72, pp. 444-467.
- Machery, E. (2006) "Two Dogmas of Neo-Empiricism", *Philosophy Compass* 1 (4), pp. 398-412.
- Machery, E. (2007) "Concept empiricism: A methodological critique", *Cognition* 104, pp. 19-46.
- Machery, E. (2009) *Doing without concepts*, Oxford, University Press.
- Machery, E. (2010) "Philosophy of Psychology", en F. Allhoff (ed.) *Philosophies of the sciences. A guide*, Oxford, Wiley-Blackwell.
- Margolis, E. & Laurence, S. (2007) "The Ontology of Concepts- Abstract Objects or Mental Representations?", *Noûs* 41 (4), 561-593.
- Margolis, E. & Laurence, S. (2008) "Concepts", *The Stanford Encyclopedia of Philosophy (Fall 2008 Edition)*, Edward N. Zalta (ed.), URL = <http://plato.stanford.edu/archives/fall2008/entries/concepts/>.
- Medin, D. & Schaffer, M. (1978) "A Context Theory of Classification Learning", *Psychological Review*, 85 (3), pp. 207-238.
- Murphy, G. (2002) *The Big Book of Concepts*, Cambridge, MA, MIT Press.
- Osherson, D. & Smith, E. (1981) "On the Adequacy of Prototype Theory as a Theory of Concepts", en E. Margolis & S. Laurence (eds.) (1999) *Concepts: Core readings*, Cambridge, MA, MIT Press.
- Pitt, D. (1999) "In defense of definitions", *Philosophical Psychology*, 12 (2), pp. 139-156.
- Prinz, J. (2002) *Furnishing the Mind: Concepts and their Perceptual Basis*. Cambridge, MA, MIT Press.
- Prinz, J. (2005) "The Return of Concept Empiricism". En H. Cohen, & C. Leferbvre (eds.), *Categorization and Cognitive Science*. Elsevier.
- Prinz, J. (en prensa) "Regaining Composure: A Defense of Prototype Compositionality", en M. Werning, W. Hinzen & E. Machery (eds.) (en prensa), *The Oxford Handbook of Compositionality*, Oxford University Press).

- Putnam, H. (1975) "The meaning of 'meaning'", en *Philosophical Papers, Vol. 2*, Cambridge, Cambridge University Press. ["El significado de 'significado'", en L. Valdés (Ed.) (1999), Madrid, Tecnos.]
- Rehder, B. (2003) "Categorization as causal reasoning", *Cognitive Science*, 27, pp. 709-748.
- Rehder, B. & Burnett, R. (2005) "Feature inference and the causal structure of categories", *Cognitive Psychology*, 50, pp. 264-314.
- Rosch, E. (1975) "Cognitive representations of semantic categories", *Journal of Experimental Psychology: General*, 104 (3), pp. 192-233.
- Rosch, E. & Mervis, C. (1975) "Family resemblances: Studies in the Internal Structure of Categories", *Cognitive Psychology*, 7, pp. 573-605.
- Sarnecki, J. (2004) "The multimedia mind: an analysis of Prinz on concepts", *Philosophical Psychology*, 17 (3), pp. 403-418.
- Schneider, S. (2009a) "The Nature of Symbols in the Language of Thought", *Mind and Language*, 24 (4), pp. 523-553.
- Schneider, S. (2009b) "LOT, CTM, and the Elephant in the Room", *Synthese*, 170 (2), pp. 235-250.
- Skidelsky, L. (en prensa) *Representaciones mentales: ¿Un eslabón entre el individuo y el mundo?*, Bs. As., Catálogos.
- Smith, E. & Medin, D. (1981) "The Exemplar View", en E. Margolis & S. Laurence (eds.) (1999) *Concepts: Core readings*, Cambridge, MA, MIT Press.
- Smith, E., Osherson, D., Rips, L. & Keane, M. (1988), "Combining Prototypes: A Selective Modification Model", en E. Margolis & S. Laurence (eds.) (1999) *Concepts: Core readings*, Cambridge, MA, MIT Press.
- Sloman, S. (2005) *Causal Models*, Oxford, UK, Oxford University Press.
- Weinberg, J., Yarlett, D., Ramscar, M., Ryder, D., Prinz, J. (2003) "Review Symposium: Making Sense of Empiricism?", *Metascience*, 12, pp. 279-303.
- Weiskopf, D. (2007) "Concept Empiricism and the Vehicles of Thought", *The Journal of Consciousness Studies*, 14, pp. 156-183.
- Weiskopf, D. (2009a) "Atomism, pluralism, and conceptual content", *Philosophical and Phenomenological Research* (79), 130-162.

Weiskopf, D. (2009b) "The plurality of concepts", *Synthese* (169), 145-173.