



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
FACULTAD DE FILOSOFÍA Y LETRAS
INSTITUTO DE INVESTIGACIONES FILOSÓFICAS
PROGRAMA DE MAESTRÍA Y DOCTORADO EN FILOSOFÍA

CONCEPTOS, CONTENIDO Y CATEGORIZACIÓN

TESIS PARA OPTAR AL GRADO DE DOCTOR EN FILOSOFÍA

MANUEL ERNESTO RODRÍGUEZ TUDOR

DIRECTORA
DRA. MAITE EZCURDIA O.



MÉXICO, D.F. OCTUBRE, 2010





Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

*Héctor Manuel Rodríguez Parra,
Irma Inés Parra de Rodríguez
y Pedro Rodríguez Parra*

A su memoria

INDICE

Agradecimientos	4
Introducción	5
1. Computación, contenido y composicionalidad	18
1.1. Computación	22
1.2. Contenido	26
1.3. Composicionalidad	31
1.3.1. Lenguaje y pensamiento	33
1.3.2. Composicionalidad y contenido mental	37
1.4. Conclusión	45
2. Conceptos Primitivos	47
2.1. Representación primitiva v/s contenido estructurado	49
2.2. Innatismo 1: poder representacional	54
2.3. Innatismo 2 : aprendizaje v/s adquisición	64
2.4. Contenido y Forma	72
2.5. Conclusión	77
3. Problemas: Contenido Cognitivo y Categorización	79
3.1. Contenido Cognitivo	83
3.2. Categorización	88
3.2.1. Teorías basadas en la categorización	94
3.3. Composicionalidad y teorías basadas en la categorización	99
3.4. Conclusión	109
4. Solución: Contenido y Forma	111
4.1. Contenido Cognitivo reconsiderado	114
4.1.1. Computación: Contenido y Sintaxis	116
4.1.2. Conceptos complejos	118
4.1.3. Conceptos primitivos	120
4.2. Categorización: conceptos y su uso	125
4.3. Contenido y forma re-unidos	130
4.4. Conclusión	134
Conclusión	138
Apéndice: Machery 2009	148
Referencias	155

AGRADECIMIENTOS

Si bien se suele asumir que realizar, y concluir, exitosamente una tesis de doctorado es un asunto más bien individual, el trabajo aquí presentado no hubiera sido posible sin el apoyo, colaboración y trabajo de muchas personas. No sólo a nivel académico, sino que en mi caso por el hecho de realizarlo fuera de mi país, los resultados aquí ofrecidos son el producto de muchas voluntades que se unen en estas páginas.

En primer lugar, quiero agradecer el trabajo generoso y dedicado de mi directora de tesis, Dra. Maite Ezcurdia O., quien supo guiar y colaborar activamente esta tesis. A las doctoras Ángeles Eraña L. y Laura Duhau G., miembros de mi comité tutorial, por cuya paciencia y colaboración efectiva les estoy totalmente agradecido. A la Dra. Salma Saab H. y el Dr. Axel Barceló A., revisores de mi tesis, por su trabajo y valiosos comentarios ya en la etapa final de este texto.

En lo académico quiero reconocer el aporte que significó durante todo este tiempo pertenecer al programa de Estudiantes Asociados del Instituto de Investigaciones Filosóficas de la UNAM. La participación en seminarios de proyectos de investigación, entre otras actividades, me permitió intercambiar puntos de vistas que resultaron sumamente valiosos a la hora de poner por escrito el trabajo realizado. Además de los investigadores ya mencionados, en particular quiero agradecer a los doctores Leonard Clapp y Mario Gómez-Torrente, y mi colega estudiante Raúl Rodríguez M.

En lo personal, fundamental ha sido el apoyo incondicional de mi esposa Andrea, quien lo dejó todo para estar a mi lado y acompañarme en esta aventura. A mi familia, en especial a mi madre, quién ha vivido con más intensidad que yo cada dificultad y éxito de este proceso. A todos los que me ayudaron a que mi estadía en México fuera lo que ya empieza a constituir uno de los mejores recuerdos que tengo. En especial quiero agradecer al Dr. Adalberto de Hoyos y Pilar Sánchez, quienes me recibieron al llegar y siempre han estado cerca.

Finalmente, quiero agradecer el apoyo económico del CONACYT de México y del Departamento de Filosofía de la Facultad de Filosofía y Humanidades de la Universidad de Chile. A CONACYT por la beca para estudios de posgrado, y a mi departamento y facultad por otorgarme la comisión de estudios que con esta tesis llega a su fin.

Coyoacán, México D.F.
Octubre de 2010.

INTRODUCCIÓN

“El concepto de concepto, además de resbaladizo, toma distintas formas no solo a través de las diversas disciplinas de la ciencia cognitiva, sino que también a través de distintas perspectivas dentro de cada disciplina¹.”

Barsalou 1993: 29

¹ Cada cita ha sido traducida personalmente a partir del texto original en inglés.

La disciplina, la teoría y la disputa

A fines del siglo pasado la discusión en ciencia cognitiva acerca de la naturaleza de los conceptos parecía llegar a un punto muerto. Una de las áreas más fértiles en cuanto a la naturaleza interdisciplinaria de la ciencia cognitiva, con importantes aportes de todas las ramas del conocimiento ligadas a esta joven ciencia, llegaba a una especie de encrucijada donde las distintas posiciones resultaban incompatibles.

Luego de despertar un enorme interés en las diversas áreas de la ciencia cognitiva, convirtiéndose en una verdadera área de intercambio disciplinario, donde podían hacer su aporte tanto psicólogos como filósofos, antropólogos cognitivos y científicos de la computación, poco a poco comenzaron a aparecer las diferencias entre los participantes del debate, no sólo entre disciplinas sino al interior de las mismas.

La teoría de conceptos debería proporcionar las bases para toda otra área de investigación de la mente, o al menos se suponía el uso de conceptos en las caracterizaciones de otros procesos cognitivos de orden superior². Por ejemplo Margolis y Laurence (2004: 190) identifican los conceptos como los constituyentes de los pensamientos; Solomon los llama los elementos básicos

² En particular me refiero a las distintas caracterizaciones de la forma en que se procesaba la información. Sin duda la investigación de los procesos cognitivos que pueden llamarse de orden inferior, por ejemplo aquellos ligados a la percepción, no suponen la posesión de conceptos, y en este sentido las teorías acerca de este tipo de representaciones mentales no constituyen la base para la explicación de toda la mente. Sin embargo, es al nivel de los conceptos dónde se supone que los procesos cognitivos podrían unificarse, poniendo en contacto los niveles perceptual, motor, y aquel llamado propiamente cognitivo.

a partir de los cuales se construyen los pensamientos (Solomon et al., 1999:99). Gran parte de la disputa se centró en los así llamados conceptos léxicos, aquellos que normalmente se expresan en los lenguajes naturales por medio de un ítem léxico, pues serían éstos los de más relevancia para los distintos estudios relativos a la mente computacional: los psicólogos del desarrollo mostraban su evidencia respecto de cuáles eran los primeros conceptos/palabras que los niños aprendían, y cuál era la estructura de los mismos revelada en los experimentos de tareas de categorización. Lingüistas y filósofos aportaban desde sus respectivos campos sus ideas respecto de cómo esos conceptos/palabras adquirían sus propiedades semánticas y como aparecía una sintaxis subyacente a los procesos asociados a la adquisición del lenguaje. Incluso los antropólogos cognitivos mostraban la evidencia de la poca o mucha variación que podía observarse a través de distintas culturas respecto de las capacidades que resaltaban psicólogos y lingüistas.

Así, aparecieron teorías acerca de los conceptos léxicos que decían que éstos eran una lista de rasgos que constituían condiciones separadamente necesarias y conjuntamente suficientes (definicionistas) para la aplicación del concepto, otros decían que en tal lista los rasgos sólo representaban condiciones estadísticamente relevantes para su aplicación (prototipos). Luego están los que sostienen que no hay tal lista de rasgos, sino que es más bien una especie de recuerdos de objetos completos contra los cuales comparamos los nuevos objetos lo que determina si se aplica o no la representación a ese nuevo objeto (ejemplares). También surgieron quienes defienden la idea de que los conceptos léxicos están inmersos en ciertos cuerpos de conocimiento o

mini-teorías y que es su papel en tales cuerpos de conocimiento lo que caracteriza su contenido y, por tanto, sus condiciones de aplicación (teoría-teoría). Finalmente, también hay lugar para una teoría que dice que tales conceptos no tienen estructura, se relacionan directamente con su referencia, y, por lo tanto, estos conceptos son innatos (atomismo conceptual)³.

Estas diferencias aparecieron como irreconciliables entre cada teoría de conceptos ofrecida por cada autor o grupo de autores. Si la teoría de la definiciones implicaba predicciones que la psicología cognitiva no confirmaba, entonces habría que renunciar a al menos una de las tareas explicativas que debía servir una teoría de conceptos: o bien se privilegiaba la explicación de la semántica conceptual, basada en el principio de composicionalidad, o bien se ponía el énfasis en lo que revelaban los experimentos de categorización. Ahora bien, si se optaba por esta segunda opción, todavía quedaba por dirimir el tipo de procesamiento y la naturaleza de las representaciones involucradas según cuál fuera el modelo de categorización escogido: si prototipos, ejemplares o teoría-teoría. Por su parte el atomismo conceptual proponía una escisión radical entre la naturaleza del contenido de la representación mental y los procesos en los cuales participaba, fueran éstos de constitución del contenido (definicionistas) o de procesos ligados a competencias epistémicas reveladas en la categorización. En este sentido lo que se pretende es desligar los aspectos psicológicos de los conceptos de aquellos propiamente semánticos.

³ Obviamente ésta es sólo una presentación simplista de la situación. En los capítulos siguientes se ofrecerá una caracterización más detallada de cada una de las posturas en juego. Por ahora sólo interesa caracterizar el estado de la cuestión y darle una idea al lector de cuál era la preocupación principal para los científicos cognitivos que se dedicaban a las teorías de conceptos. Además, si bien puede ser que las posturas aquí presentadas no agoten el espacio lógico de maneras de pensar los conceptos, si es exhaustiva respecto de las teorías ofrecidas dentro de la ciencia cognitiva.

Esta distinción se puede rastrear a las disputas entre teorías del contenido conceptual que intentan caracterizarlo en términos puramente externistas o referencialistas y aquellas que privilegian una noción puramente internista. Otra denominación que puede encontrarse en la literatura de esta diferencia es entre contenido amplio (referencia) y contenido estrecho (papel conceptual)⁴. El atomismo pretende que todo lo que tradicionalmente se pensó como contenido estrecho debe ser relegado a funciones psicológicas o epistémicas que no tienen ninguna relación con la determinación del contenido de las representaciones mentales. De seguir esta línea se pierde toda opción de explicar, por medio de una teoría de conceptos, las características más relevantes que se revelarían en la posesión de conceptos por parte de los agentes cognitivos, y así la teoría misma pierde atractivo para los científicos cognitivos interesados en las tareas psicológicas para las cuales se postulaban los conceptos.

Como en toda ciencia joven, el gran desacuerdo surgido entre los científicos cognitivos despertó el interés de defensores de marcos de investigación alternativos a lo que usualmente se conoce como el marco de investigación hegemónico en ciencia cognitiva, también llamada la “buena y anticuada inteligencia artificial”. Si dados los supuestos del marco hegemónico no era posible llegar a una visión coherente de la naturaleza de los conceptos, entonces parecería razonable poner en cuestión dicho marco. En este espíritu revisionario apareció el libro de Jesse Prinz, *Furnishing the Mind* (2002), donde se pone en cuestión la arquitectura cognitiva asumida en las investigaciones

⁴ Debe notarse que las nociones no son necesariamente equivalentes. Por ejemplo Segal (2002) rechaza las teorías del papel conceptual, sin embargo defiende la noción de contenido estrecho para una psicología científica.

respecto de los conceptos, y se propone que sólo en el marco de una arquitectura conexionista es posible satisfacer las necesidades explicativas para las cuales se postula una teoría de conceptos. De hecho Prinz toma la disputa entre las teorías del contenido conceptual antes expuesta, denominando a ambas posiciones “contenido intencional” y “contenido cognitivo” para referirse a las posturas externistas e internistas respectivamente, y usa la diferencia irreconciliable mencionada como motivación para su propuesta⁵.

Por otro lado, el trabajo en psicología cognitiva ha continuado en este campo, pero la tensión entre estas investigaciones y los diagnósticos de los filósofos sobre los graves errores cometidos por los psicólogos han dado pie para zanjar la cuestión de un modo distinto. Así, por ejemplo, Edouard Machery (2009) recientemente ha publicado un trabajo donde sostiene que la noción de concepto que manejan los psicólogos simplemente no es la misma que aquella sobre la que disputan los filósofos⁶.

La Propuesta

Pese a los problemas ligados a la formulación de una teoría de conceptos en ciencia cognitiva descritos en el apartado anterior, gran parte de

⁵ Sacrificando obviamente el principio de composicionalidad, al relegarlo a una estrategia alternativa que sólo se utilizaría por defecto cuando no está disponible la opción privilegiada por su teoría. Desarrollo esta objeción en el capítulo 3.

⁶ Sobre la propuesta de Machery me pronuncio en el apéndice, pues su alcance excede los lineamientos de esta tesis.

los investigadores en ciencia cognitiva, tanto filósofos como psicólogos, consideran que los conceptos constituyen la base de las propuestas para dar cuenta de la mente cognitiva. Por lo tanto en el presente trabajo no seguiré ninguna de las dos opciones descritas anteriormente, esto es, ni el rechazo del marco hegemónico (Prinz, 2002), ni la separación tajante entre el trabajo de las disciplinas que constituyen la ciencia cognitiva (Machery, 2009). Por el contrario, lo que aquí intento es una solución dentro del marco hegemónico en ciencia cognitiva, que sea capaz de dar cuenta tanto de los aspectos psicológicos de los conceptos como de sus características semánticas, principal preocupación de los filósofos. En este sentido, una parte sustancial de la propuesta consiste en mostrar una forma de reconciliar los aspectos ligados a la determinación del contenido, y específicamente con relación a la composicionalidad del mismo, con aquellos aspectos que se han privilegiado en los procesos relacionados con la posesión de conceptos, específicamente con relación a las tareas de categorización.

Sin embargo, conviene dejar claro desde el principio que no llevaré a cabo la labor de presentar una teoría completa de los conceptos, la que debiera incluir una clasificación exhaustiva de los conceptos primitivos, por una parte, y por otra una explicación detallada de todos los mecanismos en los que participan los conceptos. Lo que aquí pretendo más bien consiste en mostrar cómo a partir de ciertas consideraciones generales acerca de la arquitectura cognitiva clásica y las hipótesis básicas que guían la investigación en ciencia cognitiva, hay ciertas restricciones sobre lo que puede ser un concepto y los papeles que puede jugar dentro de la economía cognitiva. Eso en primer lugar,

y en segundo, que tal caracterización de los conceptos es posible de ser satisfecha, al menos en principio, sin tener que renunciar a ninguno de los elementos básicos del marco de investigación en cuestión. En este sentido, la tesis no es una defensa de la arquitectura clásica sino que más bien pretende señalar que las diferencias mencionadas al interior de este marco no son por completo irreconciliables, y por lo tanto que no constituyen en sí mismas una motivación para abandonar el marco hegemónico de investigación⁷.

De acuerdo a lo anterior, la tesis central a defender es que el contenido de los conceptos, tal y como se entienden en el marco de investigación denominado hegemónico de la ciencia cognitiva, puede ser caracterizado en términos puramente referenciales, al menos en su mayor parte⁸, y que esta opción permite dar cuenta de los aspectos psicológicos ligados al contenido cognitivo si se considera seriamente el papel que podría jugar la forma de los conceptos en su individuación. El hecho de plantear tal caracterización del contenido sólo como una posibilidad se debe a que no me parece que el externismo sea un constreñimiento de esta arquitectura, pero si a que lo considero un *desideratum* de la misma (capítulo 1).

A diferencia de otras defensas de esta posición, mi discusión no gira en torno a los conceptos léxicos, relativa tanto a la evidencia psicológica como a

⁷ Así, en el capítulo 4 específicamente, mi argumento se centra en desarticular la objeción de Prinz (2002) antes mencionada.

⁸ No puedo descartar totalmente el que existan algunos conceptos cuyo contenido esté determinado únicamente por su papel conceptual, es decir por sus relaciones con otros conceptos. Específicamente, si hay conceptos innatos complejos cuyo contenido dependa de sus relaciones con otros conceptos del mismo tipo (innatos y complejos), que determinen ciertas características propias de los individuos de una especie, por ejemplo, entonces la teoría del contenido conceptual no puede ser puramente referencial. Sobre esta posibilidad me pronuncio en el capítulo 2 al analizar las posibilidades de caracterización de la base conceptual primitiva y su relación con el innatismo.

las consideraciones filosóficas y lingüísticas respecto a este tipo de conceptos. Mi defensa se sostiene con base en los constreñimientos y un *desideratum* derivados de la arquitectura cognitiva asumida (Hipótesis del sistema de símbolos físicos de Newell y Simon, 1976) que requiere de primitivos puramente referenciales, más una noción de composicionalidad fuerte para los símbolos complejos que se obtienen a partir de los primitivos. No obstante lo anterior, parte importante del trabajo consiste en defender esta posición de las críticas provenientes de teorías de inspiración psicológica (y según Prinz 2002, rastreables a los casos de Frege) respecto de la impotencia explicativa de una teoría del contenido conceptual puramente referencial, específicamente respecto de la categorización. La solución propuesta consiste en mostrar una forma de concebir la relación entre el contenido y la forma o sintaxis de los conceptos que sea capaz de dar cuenta de la categorización y los casos de Frege sin tener que apelar a otra categoría semántica distinta a la referencia. La forma de lograr esto consistirá en concebir la relación entre forma y contenido para los conceptos primitivos de manera no-arbitraria, basándose en consideraciones relativas a las capacidades o recursos no-conceptuales que determinan las posibilidades de tales conceptos. De acuerdo a esto último la base conceptual primitiva no podrá ser tan amplia como para cubrir la mayoría de los conceptos léxicos, contra la postura de Fodor (1998), sino que más bien se verá reducida a aquellos conceptos que puedan ser rastreables a alguna capacidad innata, ya sea conceptual ella misma o no-conceptual, pero que juegue un papel determinante en las condiciones de posesión de un concepto primitivo (Caps. 2 y 4)

Para lograr lo anterior, en el capítulo 1 desarrollo un argumento que parte de la arquitectura computacional para llegar a una teoría externista del contenido por medio de la noción de concepto primitivo. En primer lugar describo los supuestos básicos del marco hegemónico, tomando como la formulación más clara del mismo la hipótesis del sistema de símbolos físicos de Newell y Simon en conjunto con la hipótesis del lenguaje del pensamiento de Fodor (1975). A partir de esta caracterización relativa a la implementación física en el marco computacional derivo un *desideratum* respecto del contenido de los símbolos primitivos, a saber una concepción externista del mismo, y un constreñimiento para los conceptos complejos, el principio de composicionalidad.

En el capítulo 2 señalo la diferencia de mi propuesta con lo que tradicionalmente se encuentra en las teorías de conceptos, en cuanto a la importancia que ha tenido la discusión relativa a los conceptos léxicos por sobre los conceptos primitivos. Dada la caracterización hecha en el capítulo 1 la relevancia de la naturaleza de los conceptos primitivos queda manifiesta. En primer lugar me dedico a establecer claramente la noción de concepto primitivo, en cuanto a que la noción de primitivo se refiere al contenido conceptual y no a sus condiciones de individuación de forma exhaustiva, pues aun en los conceptos primitivos se debe reconocer otro componente además del contenido, al menos en términos de una diferencia entre contenido y forma. Considero diversas propuestas acerca de la naturaleza representacional (contenido) de la base conceptual primitiva, e intento mostrar que, dados el *desideratum* y el constreñimiento alcanzados en el capítulo 1, la versión del

contenido de los primitivos debiera aceptar al menos algunos conceptos primitivos no-innatos⁹. En tanto la propuesta más desarrollada respecto de los conceptos primitivos, expongo la teoría de Fodor (1998)¹⁰, y muestro cómo ésta intenta lograr una noción de concepto primitivo no-innato mediante la distinción entre “adquisición” y “aprendizaje”¹¹, la que resultará fundamental para la presente propuesta. Si bien expongo aquí la propuesta fodoriana para evitar que todo concepto primitivo sea innato, las condiciones que requiero para lograrlo no están metafísicamente tan cargadas y, por lo mismo, no son capaces de cubrir a la mayoría de los conceptos léxicos sino que reducen drásticamente lo que puede aparecer en la base conceptual primitiva¹².

En el capítulo 3 expongo las objeciones a una teoría puramente referencialista del contenido conceptual que surgen desde la psicología cognitiva a partir de lo que revelan los experimentos en tareas de categorización. En este capítulo sigo principalmente la exposición de Prinz por cuanto sus objeciones se dirigen directamente a la teoría de Fodor, expuesta en el capítulo 2, y que se aplicarían a cualquier teoría que intentara individuar el contenido en términos puramente externistas. A partir estas objeciones

⁹ Mi noción de innato (y no-innato) se apega al uso de Carey (2009), en tanto requiere que los conceptos primitivos innatos no utilicen recursos representacionistas previos. Así, si para la posesión de un concepto se utilizan recursos representacionistas no conceptuales, entonces tal concepto no contaría como innato.

¹⁰ Carey (2009) también se hace cargo de una caracterización de los primitivos. No la considero para fines de la exposición pues ella asume una semántica de doble aspecto y niega la tesis del continuo, dos supuestos incompatibles con mi propuesta.

¹¹ Aquí utilizo “adquisición” en un sentido más amplio que Carey (2009) donde todo concepto no aprendido o “adquirido” resulta ser innato (y sólo requieren de maduración o de ser “disparados” por experiencias adecuadas). En mi uso, un concepto primitivo puede ser adquirido si su posesión depende de recursos o mecanismos no conceptuales previos, probablemente innatos. Es justamente la idea de que algunos primitivos no son innatos la que pretendo rescatar al utilizar el término “adquisición”, y la distinción de “aprendizaje” en cuanto al uso de recursos conceptuales previos.

¹² Como se podrá observar desde ya y a lo largo de la tesis, si bien tomo muchas de las ideas de Fodor, creo que este es un asunto crucial donde mi propuesta se diferencia de la suya. Como se presentará en la conclusión, esta diferencia coloca a mi propuesta más cerca de las teorías definicionistas, con las cualificaciones que se harán en la mencionada conclusión.

expongo las teorías inspiradas en las tareas de categorización y muestro cómo es que no son capaces de cumplir con el constreñimiento de la composicionalidad, por lo que finalizó descartándolas para la propuesta¹³.

En el capítulo 4 intento mostrar cómo, a partir de las distinciones entre contenido y forma ya alcanzadas, es posible responder a las objeciones de Prinz sin abandonar una semántica de los conceptos puramente referencialista. Si bien no detallo alguna teoría externista en particular, presento los recursos a los cuales éstas pueden acudir. En particular, una adecuada caracterización de lo que se puede entender por la relación entre contenido y “forma” o “sintaxis” permite lidiar con los casos inspirados en los problemas de Frege así como en lo que revelan las tareas de categorización. Importante resulta en este capítulo la posibilidad de dar cabida a algunas inferencias dentro de la individuación de los conceptos complejos, específicamente inferencias analíticas¹⁴, pero que juegan el papel de forma del mismo y no de su contenido.

En la conclusión reúno lo logrado en el desarrollo de los capítulos a fin de mostrar qué tipo de teoría de conceptos es la que se sigue de las consideraciones hechas en la tesis. Esto me permite distinguir mi propuesta de otras teorías definicionistas (para los conceptos complejos), llamadas teorías híbridas. Finalmente, a partir de la concepción que Machery (2009) llama fundacionalista respecto de la relación entre las nociones de “concepto” de

¹³ Obviamente esto no constituye una objeción a la teoría de Prinz. Mi objetivo sólo consiste en bloquear la motivación que él presenta para desechar la arquitectura clásica que ha dominado la discusión respecto de los conceptos.

¹⁴ Aclararé en el capítulo 4 en qué sentido las llamo inferencias analíticas, pues en este caso es la constitución del concepto complejo, por medio de su forma, lo que hace que las inferencias sean analíticas.

psicólogos y filósofos, propongo una posible respuesta al rechazo de la tesis del continuo esgrimida por Carey (2009).

En último término, y en un apéndice a la tesis, enfrento la posición de Machery respecto a la distinción radical entre las nociones de “concepto” de psicólogos y filósofos, mostrando que, dados los logros alcanzados en la tesis, la distinción es relevante pero no implica necesariamente disociar de forma absoluta ambas líneas de investigación.

CAPÍTULO 1

Computación, Contenido y Composicionalidad

“Así, supongamos que una criatura ha evolucionado estados internos orgánicos que representan asuntos de importancia adaptativa, tales como el tamaño, el número y la velocidad de aproximación de los depredadores. Si ese proceso evolutivo da como resultado un sistema físico cuyas transiciones de estados causales, bajo esta interpretación, tienen sentido semántico (por ejemplo, si menos de dos depredadores detectados causan una instancia interna de "levantarse y luchar" que lleva a una salida de conducta agresiva, mientras que tres o más a una respuesta de "correr y esconderse"), entonces la naturaleza, en esta concepción, ha evolucionado una pequeña computadora. El cerebro, si las conjeturas exploradas anteriormente se confirman, es sólo un ordenador natural del tipo descrito, incorporando estados internos que representan acontecimientos externos (tales como la presencia de los depredadores) y explotando las rutinas de las transiciones de estado que hacen un uso sensible a la información así codificada.” (Clark, A. 2001:18).

Las conjeturas de las que habla Clark pueden ser vistas como definiendo los supuestos centrales de lo que ha sido el marco de investigación predominante en ciencia cognitiva, a saber, el computacionismo y el representacionismo¹⁵. De acuerdo a tal marco¹⁶, la mente es vista como un procesador de información encargado de negociar nuestra interacción con el ambiente dadas las metas y propósitos del sistema. En términos menos abstractos, o más cercanos a nuestra psicología popular, si se concibe la mente como un procesador de información, entonces tenemos dos fenómenos que explicar: los estados mentales y los procesos mentales. Los primeros responden al supuesto representacional, donde el carácter intencional de tales estados es lo que permite explicar la coordinación entre los estímulos y la acción por medio de su contenido, o, en términos más intuitivos, el hecho de que nuestras acciones sobre el ambiente parecen ser sensibles a las condiciones del entorno. Los procesos mentales corresponden a las computaciones (cálculos, combinaciones) que somos capaces de hacer sobre las representaciones, y que respetan su contenido. Así, si el hecho de que mi creencia en que va a llover, junto con mi deseo de no resfriarme, me lleva a tomar un paraguas al salir a la calle, se explica tanto porque mis estados

¹⁵ Sobre el aspecto evolutivo y sus implicaciones no me detendré. Lo importante que deseo destacar es que tales conjeturas dan pie para pensar la mente como un sistema descrito funcionalmente en términos de representaciones y transiciones de estado que operan sobre ellas. La idea de que tal sistema pueda haber evolucionado naturalmente sólo me compromete con el naturalismo acerca de lo mental (en cuanto a adoptar una postura ontológica acorde con algún tipo de fisicalismo), en contra de una interpretación dualista inaceptable dentro de este marco de investigación, y no con la evolución de todos y cada uno de los componentes del sistema, por ejemplo con la afirmación de que todas las representaciones primitivas sean innatas (Ver capítulo 2).

¹⁶ Dados los supuestos de la presente investigación no consideraré marcos alternativos surgidos en los últimos años que ponen en entredicho el supuesto representacionista, como la teoría de sistemas dinámicos defendida por Thelen y Smith (1994) o la aproximación de la robótica encarnada de Brooks (1991).

mentales son acerca de la lluvia, mi salud, y la función del paraguas, como por la forma en que combino todas esas representaciones¹⁷.

Si bien hay distintas formas de entender la computación, una idea básica que subyace a este supuesto es el carácter funcional que debe tomar la descripción de los estados y procesos cognitivos. En este sentido, la ciencia cognitiva está comprometida con la no reducibilidad de lo mental, por ejemplo en términos del vocabulario de la neurociencia, o alguna ciencia de nivel más básico, puesto que un sistema computacionalmente descrito sólo consiste de una especificación, en términos formales, de las funciones que es capaz de llevar a cabo el sistema, cómo lo hace y qué requiere para cumplir su tarea¹⁸. Por otro lado, el supuesto representacionista tiene un compromiso con el vocabulario mentalista propio de la psicología, en tanto el sistema requiere de elementos no observables en la conducta que conecten sus acciones sobre el exterior con datos que puedan ser considerados estímulos o insumos provenientes también del exterior. Así, a grandes rasgos la ciencia cognitiva puede verse como un rescate de la psicología mentalista, pero con las credenciales científicas que esgrimía el conductismo en contra de ésta¹⁹.

¹⁷ Esto sólo constituye un ejemplo para facilitar la comprensión. Nada de lo aquí expresado requiere que seamos conscientes de los estados y procesos mentales, como podría sugerirlo el ejemplo. Más bien el supuesto del que se parte es que la mayoría de los procesos mentales, computacionalmente descritos, ocurren a un nivel inconsciente.

¹⁸ Comúnmente se habla de la “realizabilidad múltiple” de un sistema computacionalmente descrito, en cuanto el material en el que se puede implementar admite distintas configuraciones posibles, para justificar este compromiso con la no-reducibilidad. No obstante, cabe mencionar que la múltiple realizabilidad tiene restricciones, en cuanto a que el material que puede servir de soporte físico al programa debe tener una cierta complejidad.

¹⁹ En tanto sólo pretendo dar aquí un panorama general del marco de investigación predominante en ciencia cognitiva, no insistiré en estas distinciones. En tanto no-dualista, la ciencia cognitiva goza de las credenciales científicas mencionadas (y, por lo mismo, la investigación podría demostrar su falsedad). En tanto no hace uso de la introspección como única fuente de evidencia para sus hipótesis, se aleja de la psicología mentalista criticada por los conductistas.

Dentro de este marco representacional/computacional resulta relevante considerar que la mente es vista como un procesador de información. La información es tanto aquello que viene de nuestro exterior como la información acerca del propio estado del sistema, en términos de lo que es capaz de hacer y lo que desea/necesita hacer.

En lo que sigue de este capítulo expondré una forma de entender el sistema computacional cognitivo basándome en las hipótesis del Sistema de Símbolos Físicos y la del Lenguaje del Pensamiento (1.1). Mostraré que a partir de esta idea de entender la mente como un programa computacional, realizado en el caso de los humanos en el cerebro, dentro de este marco de investigación, es posible derivar un *desiderátum* sobre lo que puede ser una representación, específicamente sobre la naturaleza de su contenido. Daré tres argumentos por los cuales me parece que el contenido de los conceptos primitivos debería caracterizarse en términos puramente referenciales (1.2). A su vez, tal noción de contenido, en el marco computacional descrito, conlleva restricciones para la operación del sistema mismo, en cuanto a sus capacidades para producir nuevos estados mentales, y que se encuentran íntimamente ligadas a la noción de composicionalidad, la cual considero no como un *desiderátum* sino como un constreñimiento de la arquitectura cognitiva clásica (1.3).

1.1. Computación

La idea de pensar la mente desde la perspectiva computacional/representacionista, toma cuerpo al considerar la noción de sistema de símbolos físicos. Un sistema de símbolos físicos, como lo definen Newell y Simon (1976, pp. 85-88)²⁰, está constituido por:

- a) un conjunto de elementos interpretables y combinables, y
- b) un conjunto de procesos capaces de operar sobre esos elementos.

Ahora bien, en tanto tal sistema debe ser implementado e implementable en un dispositivo físico se requiere que ambos constituyentes sean finitos en número. Es sólo a partir de esta base finita que el sistema puede exhibir y ejecutar sus capacidades en principio ilimitadas. De acuerdo a lo anterior, suele considerarse que los elementos interpretables y combinables no son ellos mismos el resultado de algún proceso operado sobre otros elementos, y en este sentido se les reconoce como los elementos básicos o primitivos del sistema. El problema relativo a la individuación de estos elementos constituye en gran medida lo que desarrollo en el capítulo 2 y que sustenta la hipótesis central de la tesis en cuanto al contenido individuado de forma puramente externista. Especialmente relevante resulta que la individuación de su poder representacional requiera de una semántica puramente referencial, como se

²⁰ Asumo aquí la hipótesis de Newell y Simon como hipótesis de trabajo, por lo que no me dedicaré a defenderla, en cuanto a que sea una condición necesaria y suficiente para la inteligencia.

verá más adelante. Del mismo modo ocurre en el caso de los procesos, éstos entendidos como las reglas o computaciones que operan sobre los elementos, en cuanto que este conjunto finito corresponde a reglas que no surgen a partir de la combinación de elementos a partir de otras reglas más básicas.

Al asumir que la mente es equiparable a un sistema de símbolos físicos nos comprometemos con la idea de que debe haber un número finito de símbolos primitivos y un número finito de reglas de procesamiento que operan sobre esos símbolos, y que pueden operar recursivamente²¹ sobre sus productos. Como lo dije en el párrafo anterior, esto es un requisito de la arquitectura del sistema puesto que, por un lado, necesariamente un sistema físico tiene recursos finitos, y por otro, que al intentar especificar las capacidades mentales éstas deben dar lugar a posibilidades indeterminadas de estados posibles (infinitas).

Si tomamos como modelo la implementación computacional del juego del ajedrez, entonces claramente los elementos y las operaciones finitas que describen el programa determinan una cantidad finita de estados posibles del juego (dados el estado inicial y las reglas del juego). Si en cambio tomamos como modelo la implementación computacional de un lenguaje posible, con operaciones recursivas, entonces el vocabulario/léxico inicial más la

²¹ Es común en la literatura acerca de los conceptos en ciencia cognitiva que el paralelo que suele establecerse entre el lenguaje y el pensamiento permee a los términos técnicos que se utilizan. En este caso la idea de recursividad proviene de los estudios concernientes al lenguaje natural y se refiere a las reglas que permiten formar expresiones más complejas a partir de expresiones más simples. También se puede rastrear este tipo de términos a los estudios de los lógicos y sus propuestas de lenguajes formales. Si bien bastaría con caracterizar el uso de estas reglas como “reiterativas” considero que el paralelo con los desarrollos de lenguajes formales, que sirvieron de base a las ciencias de la computación, justifican su uso en este contexto.

gramática²² nos dan una cantidad indeterminada de oraciones bien formadas posibles²³. En tanto nuestra mente estaría destinada a negociar información con el ambiente a fin de ejecutar los mejores cursos de acción dadas nuestras metas, por lo que su capacidad de representar estados posibles del ambiente debiera ser en principio ilimitada, entonces parece razonable optar por el modelo del lenguaje antes que por el del juego del ajedrez²⁴.

De hecho, las relaciones que suelen establecerse entre el pensamiento²⁵ y el lenguaje suelen ir más lejos. Fodor (1975) ha propuesto la hipótesis del lenguaje del pensamiento, según la cual una forma de concebir el sistema de símbolos físicos es considerar al conjunto de elementos interpretados como el vocabulario básico del sistema, esto es su léxico inicial, y al conjunto de procesos o reglas de transformación como la sintaxis del sistema²⁶.

²² De forma similar a lo dicho en la nota 19, aquí se utiliza “gramática” para nombrar el conjunto de reglas de transformación que operan sobre los elementos léxicos del lenguaje a modelar, el cual es tomado en un punto del tiempo y se considera fijo para su modelación. En este sentido no debe confundirse con una gramática de un lenguaje natural dado, asumiendo que las reglas que aparecen en esas gramáticas pueden cambiar con el tiempo si el uso así lo impone. En lo que sigue hablaré de reglas de transformación o sintaxis cuando hablé del lenguaje del pensamiento a fin de evitar confusiones.

²³ Importante es hacer notar que la cantidad finita de representaciones iniciales no nos compromete con un número específico. Aún más, si pensamos que el sistema de símbolos pueda adquirir nuevos símbolos de una forma distinta a la mera combinación de símbolos ya existentes, entonces esta cantidad perfectamente puede no ser fija, puede crecer por ejemplo. Lo anterior requiere de una distinción entre símbolos aprendidos a partir de otros símbolos y símbolos adquiridos de una forma que no involucre la posesión previa de otros símbolos. Más sobre esto en el capítulo siguiente sobre conceptos primitivos.

²⁴ Además, el juego del ajedrez es un sistema cerrado, en cuanto a que no representa ningún estado externo al juego mismo. Si entendemos la mente como un procesador de información dedicado a la negociación del organismo con el ambiente se requiere de alguna conexión, en términos representacionistas, con algo externo al sistema. Más en 1.2.

²⁵ Por “pensamiento” entiendo aquí el procesamiento de representaciones con contenido que lleva a nuevas representaciones interpretadas.

²⁶ No asumiré en esta investigación tesis más fuertes sobre la relación o paralelismo que puede establecer entre el lenguaje y el pensamiento. Por ejemplo, no asumiré que para cada oración en el lenguaje natural debe haber una representación mental compleja que sea el significado de tal oración. Lo que me interesa rescatar es la idea de símbolo físico (con contenido y forma) por una parte, y por otra la idea de sintaxis en cuanto constituyentes exclusivos, junto con el contenido, del contenido y sintaxis de las representaciones complejas.

Podemos reformular la constitución del sistema de símbolos físicos en términos de:

- a) Un conjunto finito de representaciones, y
- b) Un conjunto finito de reglas que operan sobre esas representaciones (y recursivamente sobre sus propios productos).

Si asumimos esta noción del sistema de símbolos físicos, entonces las tareas propias de la ciencia cognitiva son, al menos, dos:

- i) Caracterizar adecuadamente la naturaleza de las representaciones básicas que componen el sistema, y
- ii) Postular el conjunto de reglas que determinan el funcionamiento del sistema, dadas las representaciones básicas²⁷.

De acuerdo a lo anterior, y al propósito de la investigación aquí presentada, pasaré a revisar las restricciones que se siguen para la caracterización de las representaciones, específicamente respecto de su contenido.

²⁷ Esta caracterización de las reglas resulta insuficiente si se pretende capturar todo el funcionamiento de la mente. Específicamente, dicha caracterización sólo cubre las operaciones sobre representaciones, por lo que requiere además de reglas que conecten tal procesamiento con las entradas del sistema (comúnmente descritas en términos preceptuales) por un lado, y por otra con las salidas que resultan en comportamiento. Además, es posible pensar que dentro del sistema que procesa representaciones existan distintos cuerpos de reglas, de propósito específico por ejemplo. Como mi investigación se enfoca en el contenido de las representaciones mentales no entraré en estos temas.

1.2. Contenido

Una característica relevante del sistema de símbolos físicos es que éste debe estar dentro de una red más amplia de elementos y sucesos del mundo real. La forma de asegurar esto es que los símbolos representen tales elementos y sucesos. En esto radica el carácter propiamente representacional/intencional de los símbolos, y por medio de ellos del sistema. Newell y Simon llaman a esta característica la capacidad de designación del símbolo²⁸.

La manera tradicional, en ciencia cognitiva, de entender el contenido de los símbolos es apelando a las propiedades locales del sistema, como en el caso del ajedrez. Esta forma internista de individuación del contenido, que en la introducción señalé como teorías que privilegian el contenido estrecho o cognitivo, aparecen en primera instancia como la manera natural de identificar el contenido de las representaciones mentales. La forma en que una expresión afecta o puede afectar a otros objetos del sistema, y por ende al sistema como un todo, depende de las relaciones entre las expresiones. Así, las semánticas del papel conceptual dominan casi sin excepción los desarrollos en teorías del contenido, en parte por la herencia de las ciencias de la computación, en parte

²⁸ “Una expresión designa un objeto si, dada la expresión, el sistema puede o bien afectar al objeto mismo o bien comportarse de formas que dependen del objeto.” Newell & Simon, 1976: 86.

por las objeciones que surgen en contra de una teoría puramente externista del contenido²⁹.

No obstante lo anterior, creo que hay argumentos fuertes para sostener que la semántica³⁰ de las representaciones mentales debe ser externista:

Primero, la propia noción de designación de Newell y Simon parece sugerir una determinación externista del contenido, contrario a lo que podría parecer a primera vista. En contra de teorías puramente internistas del contenido, la caracterización de la relación de designación ofrecida por los autores parece requerir de una conexión más fuerte entre el objeto representado y la representación. Sin embargo, el ser afectado y afectar al objeto podría darse aun si la semántica de tales representaciones no dependiera del objeto. Sin embargo, tomar esta postura desde la perspectiva computacional conlleva el problema de hacer vacuo el modelo computacional/funcional. Boccardi (2009) ha argumentado que si no se asume una semántica externista se corre el riesgo de trivializar la noción de computación como un modelo materialista para entender la mente. En su exposición señala que la única forma en la que el modelo computacional/funcional de la mente conserve algún poder representacional sobre algo externo a la mente misma es apelando a una semántica externista, específicamente de naturaleza causal. Las consideraciones a partir de las cuales surge esta necesidad radican en que el

²⁹ En lo que sigue de este capítulo ofrezco tres argumentos para rechazar la individuación internista. En los capítulos 3 y 4 expondré las objeciones a la individuación externista y la forma en que mi propuesta intenta resolverlas, respectivamente.

³⁰ Utilizo aquí “semántica” en tanto se refiere al contenido de las representaciones mentales. Si bien “semántica” es el término utilizado por lingüistas y filósofos del lenguaje para referirse al significado de expresiones y oraciones del lenguaje, esto no implica conexiones más fuertes de las suscritas entre el pensamiento y el lenguaje.

modelo computacional de la mente es funcional, esto es que en la arquitectura propuesta no figuran especificaciones acerca del material en el cual se implemente el modelo. De acuerdo con esto, en principio cualquier sistema con la suficiente complejidad material podría instanciar o implementar una mente, lo que torna vacuo al modelo³¹. Si se intenta reducir la capacidad de implementación apelando a características del material donde se instancia el modelo, entonces la misma noción de arquitectura funcional pierde su motivación principal. No obstante, una posible forma de enfrentar el problema sin renunciar a una caracterización puramente funcional es apelando a los mecanismos que dotan de contenido a las representaciones del sistema. Si tales mecanismos no se reducen a las relaciones computacionales entre símbolos y más bien son de un tipo que media causalmente la relación entre el símbolo y su referencia³², entonces no cualquier sistema material podrá implementar una mente, sino sólo aquellos que exhiban el tipo de mecanismos adecuados. En este tipo de mecanismos se pueden considerar aquellos ligados a la transacción directa del organismo con su ambiente, de orden perceptual por ejemplo, cuyo funcionamiento dependa de características no cognitivas ellas mismas, pero cuyos productos puedan ser tomados por el sistema cognitivo como insumos o constreñimientos para su funcionamiento.

Segundo, Laurence y Margolis (1999) han presentado un problema común a todas las teorías de conceptos que asumen la semántica de los mismos

³¹ Consideraciones similares se encuentran en Block, N. (1978).

³² No se requiere que todas las relaciones entre símbolos sean de orden causal entre el símbolo y su referencia, sino que al menos algunas lo sean a fin de evitar el problema de la vacuidad. Podrían haber conceptos innatos cuyo contenido se determine por sus relaciones constitutivas con otros conceptos del mismo tipo, y que en ellos no haya una referencia o determinación externista del contenido, sino que más bien responda a cómo están organizadas nuestras mentes. Más de esto en el capítulo 2, sobre conceptos innatos.

como constituida por las relaciones inferenciales/computacionales en las que participa, que se siguen de objeciones externistas en filosofía del lenguaje a las semánticas internistas³³. El problema, denominado “la ignorancia y el error”, radica en que la pura fijación del contenido en términos de propiedades locales del sistema no es capaz de determinar la referencia de la representación. La intuición tras el problema es que nuestras creencias, entendidas en términos de las inferencias que podemos realizar gracias a la posesión de un concepto, o bien pueden estar equivocadas respecto de la referencia del concepto, o bien podemos ignorar estas inferencias y aun así poseer el concepto en cuestión. Por ejemplo, se deben enfrentar los problemas que surgen al considerar los casos de la tierra gemela (Putnam, 1975, Burge, 1979, 1986) donde dos individuos tienen exactamente las mismas disposiciones a actuar, realizan las mismas inferencias, etc., pero cuya representación se refiere a cosas distintas, por meras contingencias ambientales o sociales. En estos casos nuestras intuiciones nos llevan a pensar que aun cuando en términos puramente internistas las representaciones de los dos sujetos sean idénticas, no obstante son representaciones distintas a causa de la diferencia en referencia, y por lo tanto que tal referencia es constitutiva de los conceptos de cada uno. El que tal referencia sea constitutiva del contenido de las representaciones no depende de las creencias del individuo, sino más bien de algún tipo de relación que puede depender de factores puramente causales, sean naturales o sociales, o ambos. Los casos de conceptos que refieren a entidades definidas dentro de teorías científicas, como GALAXIA, o a clases naturales, como AGUA, o en las cuales ambas se conjuntan, son claros ejemplos donde las creencias a nivel

³³ Si bien las semánticas del papel conceptual no agotan las formas de caracterizar el contenido estrecho, para los fines de la presente investigación debe considerarse que todas las propuestas que serán revisadas asumen este tipo de individuación del contenido.

individual, o la ausencia de tales creencias, no determinan la referencia del concepto.

Tercero, y el más importante para lo que sigue, si el sistema requiere de una cantidad de símbolos básicos, entonces éstos no pueden obtener su contenido de sus relaciones con otros símbolos, puesto que de esa forma no se obtiene una noción de computación manejable³⁴, por un lado, y por otro se renuncia al carácter representacional/intencional del sistema. La noción de computación ofrecida para caracterizar los procesos que operan sobre las representaciones requiere que éstos no afecten el contenido de las mismas. Así, las reglas que permiten relacionar símbolos y crear otros nuevos operan sobre la sintaxis/forma del símbolo, preservando su contenido (Haugeland, 1981). Por tanto, los símbolos básicos no pueden estar definidos por operaciones y otros símbolos. Así, un problema que surge al caracterizar el contenido de los símbolos en términos de las computaciones en las cuales participan es que hace inmanejable la noción de computación tal y como se ha entendido. Ahora bien, si de todos modos se asume que los símbolos se caracterizan por sus relaciones con otros símbolos, como en el caso del juego de ajedrez, entonces volvemos a los problemas que se siguen de apelar a las propiedades locales del sistema, esto es, el problema de la vacuidad del modelo funcional/computacional expuesto anteriormente³⁵.

³⁴ Esta objeción a las semánticas del papel conceptual se puede encontrar en autores con motivaciones tan diversas como Cummins (1989), Fodor (1998) y Segal (2002).

³⁵ Sin duda la pura caracterización externista del contenido también debe enfrentar dificultades, las que serán abordadas en los capítulos 3 y 4. Lo importante aquí es señalar que desde la perspectiva computacional/representacionista de la mente la semántica externista parece ser la adecuada.

En suma, parece ser que si queremos seguir sosteniendo la imagen de la mente como un procesador de información que le permite al organismo negociar con el ambiente de acuerdo a sus metas y propósitos, y se asume el sistema de símbolos físicos más la hipótesis del lenguaje del pensamiento como marco para esa imagen, entonces resulta teóricamente preferible asumir una determinación externista del contenido de los conceptos/símbolos primitivos³⁶.

1.3. Composicionalidad

En una primera aproximación el principio de composicionalidad, a nivel semántico para un sistema representacionista, establece que el significado de las representaciones complejas de un sistema se hereda del significado de sus partes constituyentes más las reglas que dan lugar a tal composición. Tal principio se postula para explicar, al menos, dos capacidades que parecen exhibir los sistemas representacionales, a saber, su productividad y sistematicidad. Por productividad se entiende la capacidad de formar nuevas representaciones a partir de las ya existentes y las reglas de composición, esto es, la capacidad/propiedad de un sistema representacionista de producir una

³⁶ Por conceptos primitivos me refiero a esta base de representaciones. El hecho de que sean llamados primitivos sólo se refiere a que no son analizables en términos de otras representaciones y sus relaciones con ellas. En sentido estricto no hay términos primitivos pues incluso las representaciones primitivas deben admitir la distinción entre contenido y forma, esto es, exhibirían complejidad interna. Sin embargo, se puede hablar tanto del contenido de los primitivos como siendo atómico, por una parte, como de una sintaxis primitiva en cuanto a no analizable en términos de relaciones sintácticas externas al símbolo. Una discusión importante, que se detallará en el capítulo siguiente, concierne a si estos conceptos primitivos se caracterizan sólo por propiedades intrínsecas, como parecería sugerirlo el aspecto formal de los mismos, o bien si aun cuando sean primitivos requieren en su caracterización de propiedades extrínsecas, por motivos del contenido intencional de los mismos.

cantidad ilimitada de símbolos semántica y sintácticamente distintos. Así, aplicando una regla de conjunción podemos unir dos representaciones ya existentes como BLANCO y CUADRADO, y formar una tercera, CUADRADO BLANCO, cuyo contenido es la unión de los contenidos de sus constituyentes³⁷. Por otra parte, la sistematicidad de un sistema representacional permite explicar cómo es que el uso de las mismas representaciones en representaciones complejas distintas están relacionados semántica y/o sintácticamente, esto es, la capacidad/propiedad de un sistema representacionista que tiene familias de símbolos distintos pero semántica y sintácticamente relacionados. Por ejemplo, cómo el hecho de que si se entiende la oración “Pancho vive en México” y “Juan vive en Chile”, entonces, por el puro hecho de entender ambas expresiones complejas, es posible reemplazar los constituyentes y se está capacitado para entender las siguientes oraciones: “Pancho vive en Chile” y “Juan vive en México”. En tanto los sistemas representacionistas cuentan con recursos finitos, por una parte, y su capacidad de expresión es ilimitada, por otra, parece necesario asumir que se respeta este principio para poder explicar este fenómeno.

No obstante lo anterior, la noción o principio de composicionalidad es una fuente de debates en filosofía del lenguaje y ciencia cognitiva. Por ejemplo, parece ser que cualquier intento de formalización de los lenguajes naturales requiere que se respete dicho principio. En el debate en torno a los conceptos también ha sido esgrimido como una forma de rechazar ciertas aproximaciones, principalmente originadas desde la psicología cognitiva. Lo

³⁷ En este caso se puede entender tal contenido como la intersección de las extensiones de CUADRADO y BLANCO.

que se pretende en este apartado es clarificar una noción de composicionalidad con base en los requerimientos computacionales y representacionales del sistema de símbolos físicos. De este modo pretendo formular ciertas restricciones sobre la noción de contenido que no se basan en el rendimiento de los individuos en ciertas tareas cognitivas, así como tampoco en ideas acerca del procesamiento (comprensión/producción) del lenguaje natural. Si bien ambas capacidades debieran poder ser explicadas por una teoría de lo mental, prefiero adoptar la perspectiva del sistema computacional para luego ver si, dadas las restricciones que se siguen de su propia caracterización, es posible que por medio de tal sistema se puedan explicar las capacidades mencionadas³⁸.

1.3.1. Lenguaje y pensamiento

Hablar del principio de composicionalidad en una tesis sobre el contenido de las representaciones mentales pareciera asumir una relación entre el lenguaje y el pensamiento que trae consigo los problemas propios de la filosofía del lenguaje. En el centro del debate acerca de la posibilidad de dar una formalización de los lenguajes naturales, por ejemplo para dar cuenta de cierta uniformidad entre estos lenguajes y la capacidad de aprendizaje de los mismos que exhiben sus hablantes competentes, está el principio de composicionalidad. A este respecto Szabó (2002) ha argumentado que dicho

³⁸ Específicamente respecto de la categorización me pronunciaré en los capítulos 3 y 4. Respecto de la capacidad cognitiva relacionada con el procesamiento del lenguaje, solo me limitaré a mostrar algunas consecuencias que podrían seguirse de mi investigación en las conclusiones de la misma.

principio puede entenderse de distintas maneras, según cuáles sean los objetivos y alcances que una teoría del significado para los lenguajes naturales se requiere que cumpla. Tal principio sostiene que el significado de las expresiones complejas depende del significado de las expresiones más simples que las constituyen, más las reglas de composición, de acuerdo a la versión tradicional del mismo. Sin embargo este principio ha sido puesto en entredicho por algunos filósofos del lenguaje (Travis 1997, Perry 1986). Ahora bien, si se asume que el funcionamiento de la mente es similar al del lenguaje, entonces resulta tentador al menos pensar que tales críticas también se aplican al lenguaje del pensamiento.

De hecho, gran parte de los argumentos a favor de una cierta teoría del contenido provienen de esgrimir la composicionalidad del lenguaje como evidencia para la composicionalidad del pensamiento. Por ejemplo, Fodor y Lepore (1991, 1996) han argumentado de esa manera, y por tanto un ataque a la composicionalidad del lenguaje pareciera resultar fatal para la teoría del contenido mental que se basaba en tales consideraciones. De un modo similar, Fodor (1998) ofrece como el “mejor argumento para la composicionalidad” del pensamiento el hecho de que usamos y entendemos descripciones definidas con fines comunicativos cuando no contamos con un medio de demostrar la referencia de nuestras emisiones. Así, cualquier hablante competente del español puede entender la descripción “El abogado vietnamita, defensor de causas perdidas, que pasa la mayor parte de su tiempo en la iglesia, cuyos vecinos detestan y a quien nadie respeta” sólo basándose en el conocimiento del significado de las partes de la expresión, y sin tener ningún conocimiento

del individuo en cuestión. Tras estos argumentos se encuentra la idea de un fuerte paralelismo entre el lenguaje y el pensamiento, específicamente que para cada expresión compleja hay una representación mental compleja que la primera expresa. Así, a nivel semántico, y también a nivel sintáctico, parece haber un compromiso con un isomorfismo entre lenguaje y pensamiento, por lo que lo que se diga acerca del pensamiento depende de lo que ocurra con las investigaciones acerca del lenguaje.

Sin embargo, este compromiso no es obligatorio, a menos que uno se comprometa con la hipótesis de que el lenguaje del pensamiento es el lenguaje natural que usan los agentes cognitivos en su comunicación con otros agentes (el inglés para el caso de los hablantes del inglés, el español para los hispanoparlantes, etc.)³⁹. Una opción consiste justamente en rechazar la composicionalidad de los lenguajes naturales para así sostener que la composicionalidad del pensamiento no es derivada de aquellos (Fodor, 2001b). Uno puede no aceptar, por ejemplo, que cada oración, en cuanto expresión lingüística de un pensamiento, deba hacer explícita la estructura de tal pensamiento, aunque de hecho pueda hacerlo (Carston, 2008: 339).

No obstante lo anterior, creo que el distanciarse del paralelismo fuerte entre lenguaje y pensamiento no basta para asumir la composicionalidad del pensamiento, y ciertamente los argumentos de Travis pueden ser usados al menos para debilitar las consideraciones tradicionales en su favor. Así alguien podría asumir que, del mismo modo como en la expresión “las hojas son

³⁹ Tal vez si se asume una postura minimista en filosofía del lenguaje también debiera haber un compromiso fuerte con el isomorfismo antes señalado. En tanto éste no es el tema de mi tesis no ahondaré en ello.

verdes” no hay condiciones de verdad independientes de la emisión de la expresión en el contexto⁴⁰, en el caso del pensamiento LAS HOJAS SON VERDES podría ocurrir lo mismo.

En lo que sigue intentaré permanecer neutral respecto de esta disputa, y pretendo mostrar que hay otro argumento, basado en la arquitectura cognitiva asumida, que puede justificar el asumir la composicionalidad del pensamiento de manera independiente de la relación que éste tenga con el lenguaje. Por ejemplo, uno de los problemas que no intenta solucionar la presente investigación es qué ocurre en casos como la expresión “vino blanco”. En lo que hasta aquí considero como composicionalidad la semántica de esta expresión debiera remitir a la referencia de los términos constituyentes, por lo que la referencia de “vino blanco” estaría constituida por todo aquel objeto que instancia tanto la propiedad de ser vino como la de ser blanco. Claramente esto es un contraejemplo al principio de composicionalidad así entendido, sin embargo no es el lenguaje natural el objeto de investigación aquí. Con lo que si estoy comprometido es que si hay un concepto complejo VINO BLANCO entonces su semántica debe ser la anteriormente expresada. La única conclusión que podría seguirse es acerca de la relación entre lenguaje y pensamiento, en el sentido de que el término complejo “vino blanco” no está expresando el concepto VINO BLANCO. Otra forma de concebir una solución al contraejemplo sin renunciar a la composicionalidad como conjunción de propiedades, es asumir que “blanco” expresa una propiedad que requiere de

⁴⁰ No estoy asumiendo aquí que Travis esté en lo correcto respecto de la semántica de los lenguajes naturales. La postura del pragmatismo radical está lejos de ser ampliamente aceptada en filosofía del lenguaje. Sólo me interesa aclarar que mi posición acerca del principio de composicionalidad para las representaciones mentales complejas no depende de este tipo de argumentos relativos a su relación con las expresiones complejas de los lenguajes naturales.

algún parámetro a ser llenado por el contexto de uso (Pietroski, 2005), por ejemplo una clase de comparación. Así, para el caso de la representación mental BLANCO, ésta expresa siempre la misma propiedad, pero en el caso del lenguaje del pensamiento no es el contexto de uso sino que su instanciación en conjunto con otro concepto, en una representación mental compleja, la que determina la referencia del concepto complejo del que es parte. Así, a lo que obliga el principio de composicionalidad en este caso es a adoptar una ontología de propiedades relacionales, donde ya no hay la propiedad de *ser blanco*, a secas por decirlo de una manera, sino más bien la propiedad de *ser blanco para x*, donde el lugar del argumento debe ser llenado por una propiedad que acepte la propiedad relacional como un modificador de su extensión. Casos similares ocurrirían con los adjetivos⁴¹ que se usan de forma comparativa, como GRANDE, ALTO, SUAVE⁴².

1.3.2. Composicionalidad y contenido mental

La mayoría de los teóricos de la representación mental en ciencia cognitiva están de acuerdo en que una condición que debe cumplirse es que se respete el principio de composicionalidad. La propia hipótesis del lenguaje del pensamiento parece implicar este principio en su versión de conjunción de

⁴¹ Entiéndase aquí la representación mental, en cuanto a su contenido, que expresamos por medio del uso de tales adjetivos en el español.

⁴² En este apartado sobre la relación entre el lenguaje y el pensamiento sólo he querido mostrar que mi propuesta no depende, ni está comprometida, con alguna posición específica respecto de los intensos debates en filosofía del lenguaje. Así, no estoy dando por resuelto ningún problema semántico pues soy consciente de que tanto el caso de los adjetivos, los términos de color, y el principio de composicionalidad siguen siendo fuente de discusiones en ese ámbito.

propiedades. Como ya se mencionó, de acuerdo a esta hipótesis los estados mentales como creencias, deseos y otras actitudes proposicionales son representaciones. Como otras representaciones, retratos u oraciones por ejemplo, los estados mentales consisten de un vehículo que tiene propiedades semánticas, que podrían ser el tener referencia, un sentido o condiciones de verdad, que constituyen el contenido de la representación. El vehículo de la representación es la entidad física que tiene esas propiedades semánticas. Además, la hipótesis del lenguaje del pensamiento sostiene que hay dos tipos de representaciones mentales. Algunas representaciones mentales son simples o primitivas, mientras que otras son complejas, esto es, están constituidas a partir de representaciones más simples, y en última instancia de primitivas, de acuerdo a reglas de composición, comúnmente llamadas la “gramática” o “sintaxis” del lenguaje del pensamiento. Así, el contenido de una representación compleja es una función del contenido de las representaciones más simples que la constituyen más las reglas de composición relevantes. Como Van Gelder (1990) ha sostenido, en cualquier sistema representacional que distingue representaciones simples y complejas, una operación específica sobre los vehículos debe corresponder a una regla de composición. Aún más, las operaciones sobre los vehículos correspondientes a las reglas de composición son relaciones estructurales entre los vehículos. Si las instancias de un estado mental son estados del cerebro, por ejemplo, entonces, de acuerdo a la hipótesis del lenguaje del pensamiento, un estado cerebral que realiza una representación mental compleja consiste de varios estados cerebrales, los cuales realizan las representaciones simples y que están en algún tipo de relación estructural. Del mismo modo, Fodor y Pylyshyn (1998:13)

asumen que las estructuras simbólicas en un modelo clásico corresponden a estructuras físicas reales en el cerebro, y que la estructura combinatorial de una representación tienen una contraparte en las relaciones estructurales entre las propiedades físicas del cerebro.

En este caso se asume que este principio es semántico en cuanto remite a la constitución del contenido de las representaciones mentales complejas⁴³. Además, como se verá en el capítulo siguiente, en cuanto la individuación de los estados mentales procede funcionalmente y se asume la intencionalidad de los mismos, el problema del contenido de los símbolos resulta crucial en la discusión.

La motivación principal de adherir a este principio radica en la necesidad de dar cuenta de las capacidades ilimitadas de la mente, en cuanto a estados posibles, dadas los recursos finitos del sistema. Específicamente hay dos capacidades/propiedades que parecen exhibir los organismos con mente que podrían ser explicadas si el sistema de símbolos físicos obedece el principio, a saber:

- i) Productividad: es la capacidad/propiedad de un sistema representacional de producir una cantidad ilimitada de símbolos semántica y sintácticamente distintos;

⁴³ Si bien uno podría evitarse muchos problemas al señalar que los científicos cognitivos más bien están interesados en las capacidades combinatorias del sistema, la literatura al respecto insiste en hablar de composicionalidad. Espero que el apartado anterior deje claro qué compromisos no se están asumiendo en la presente investigación.

- ii) **Sistematicidad:** es la capacidad/propiedad de un sistema representacional que tiene familias de símbolos distintos pero semántica y sintácticamente relacionados.

Así, un sistema representacional será productivo cuando a partir de algunas expresiones y algunas reglas sea capaz de producir ilimitados símbolos nuevos, por ejemplo si tiene los símbolos p y q más la regla del condicional, es capaz de producir los símbolos nuevos “si p entonces q ”, “si p entonces (si p entonces q)”, etc. Un sistema será sistemático si por el sólo hecho de tener la capacidad de computar la representación “si p entonces q ” podemos atribuirle la capacidad de computar “si q entonces p ”⁴⁴.

Ahora bien, la primera consideración a realizar es que dicho principio normalmente es entendido respecto del contenido de las representaciones complejas. En este sentido el principio dice que el contenido de una representación mental compleja es una función del contenido de las representaciones que lo constituyen más las reglas que se aplicaron en su constitución⁴⁵. Si bien se hace alusión a las reglas de constitución, debe recordarse que tales reglas operan sobre la forma de las representaciones constituyentes y, por tanto, también sólo aportan esa forma a la representación compleja, preservando el contenido. En el ejemplo descrito “ p ” y “ q ” en tanto

⁴⁴ He preferido tomar ejemplos de un lenguaje formal a fin de distanciarme lo más posible de las analogías con el lenguaje natural.

⁴⁵ Para distintas formulaciones de este principio ver: Fodor (1998), Prinz (2002). Otra manera de apelar al principio es mediante el Constreñimiento de Generalidad de Evans (1982), sin embargo dejaré esta formulación para el capítulo siguiente donde tendré que apelar a la distinción entre lo propiamente conceptual y otras formas de “tener” información.

símbolos conservan su contenido independientemente de las operaciones en las que figuren.

Esta distinción entre forma y contenido, y la relación que se ha de establecer entre ambos aspectos de las representaciones mentales, resulta de vital importancia para la presente propuesta. Creo que uno de los principales problemas que debe enfrentar el intento de individuar el contenido de los conceptos de forma referencial es que deja sin explicar la no arbitrariedad de tal relación. Como ya lo expuse en la introducción, el internismo que predomina en las teorías de conceptos pone dentro del contenido las relaciones inferenciales que se establecen entre los conceptos, y el externismo intenta privilegiar la relación entre el símbolo y la referencia. En ambos casos la forma de las representaciones no juega ningún papel, más allá del compromiso materialista en cuanto a que es ésta la que posibilita las relaciones causales (mecánicamente instanciadas) entre las representaciones. Una aproximación a la forma levemente desarrollada consiste en identificar ésta con la sintaxis del concepto, concebida de forma abstracta, pero aún sin permitir relacionarla con el contenido del mismo. Teniendo presente esta distinción, y la falta de desarrollos explicativos respecto de la relación entre ambos aspectos de los símbolos mentales, me parece necesario insistir en el carácter puramente semántico del principio de composicionalidad aquí defendido.

A modo de ejemplo, considérese la objeción de Recanati (2003) a la propuesta de Fodor respecto del papel que cumple el principio de composicionalidad en cuanto a restringir las condiciones de posesión de un

concepto: el argumento de Fodor consiste en asumir que todo lo que importa para la individuación de un concepto complejo es aquello que hereda (composicionalmente) de aquellos conceptos que lo constituyen. En tanto cualquier capacidad epistémica no es composicional en este sentido, deben rechazarse tales capacidades como siendo constitutivas de los conceptos, y por tanto figurar en su individuación. En contra de esta opción Recanati muestra que en el argumento de Fodor hay dos afirmaciones que le parecen correctas, pero de las cuales no se sigue la conclusión que pretende. La primera establece que los conceptos son usados como constituyentes de conceptos más complejos, y la segunda que la referencia de un concepto complejo está determinada por las referencias de sus constituyentes (y la forma en que los constituyentes se combinan). Acepta además que las propiedades epistémicas que podríamos asociar a la posesión de un concepto no componen, en este sentido, pero que sin embargo no hay razones para rechazar el que sean constitutivas de la identidad de un concepto.

Lo que quiero rescatar de la discusión entre Recanati y Fodor, por una parte y para los fines de este capítulo, es que la noción de composicionalidad que utilizaré se refiere justamente a la que destaca Recanati, esto es la composicionalidad de la referencia. Así, si el contenido de todos los conceptos primitivos es individuado de forma puramente referencial, entonces por el principio de composicionalidad, el contenido de cualquier concepto complejo se determinará de la misma manera. Por el contrario, si aceptamos la posibilidad de algunos conceptos primitivos cuyo contenido se individúe por su papel funcional, o por éste más la referencia, entonces tendremos que admitir que los

conceptos complejos en cuya constitución participen estos primitivos tendrán un contenido que no es puramente referencial⁴⁶. Por otra parte, en la discusión queda de manifiesto la tensión que esta tesis trata de resolver (al menos dentro de la arquitectura clásica) entre las aproximaciones que privilegian el contenido referencial versus aquellas que ponen el énfasis en el contenido cognitivo⁴⁷. Esto último es lo que me llevará a intentar una caracterización funcional más desarrollada de la forma conceptual, y su posible relación no arbitraria con el contenido.

De manera provisional, entiendo aquí por forma de la representación mental lo que podría denominarse el vehículo del contenido. Por una parte, en cuanto hablamos de un sistema de símbolos físicos, tal forma es entendida como la configuración material que soporta el contenido mental⁴⁸. Dentro de una aproximación materialista, y para evitar el problema de un dualismo de sustancias, se asume que si los conceptos, en cuanto constituyentes de estados mentales, tienen poderes causales, ya sea sobre otros estados mentales o sobre el sistema motor (causando una cierta conducta, por ejemplo), estos deben en última instancia depender de las propiedades materiales que posean. Por otro lado, en cuanto la aproximación a lo mental en

⁴⁶ La posibilidad de este tipo de conceptos depende de discusiones relativas al innatismo de los conceptos primitivos (Cap. 2), y su papel en la constitución de conceptos complejos dependerá de la caracterización de su forma y contenido. Una posibilidad que no va contra mi propuesta es que existan algunos conceptos cuyo contenido esté dado por una regla de aplicación y la referencia, como podría ser el caso de conceptos que exhibieran algún tipo de indexicalidad (Cfr. Ezcurdia 2001). Lo que desde ya quiero hacer notar es que parece que ser que este tipo de conceptos, si los hubiera, no jugarían un papel relevante en la constitución de conceptos complejos como MANZANA ROJA, por ejemplo.

⁴⁷ Que en el caso de mi propuesta he preferido colocar en términos de la tensión entre una adecuada caracterización del contenido, por una parte, y la posibilidad de dar cuenta de las capacidades cognitivas asociadas a la posesión de conceptos que se revelan en la categorización. De ahí el título de la tesis.

⁴⁸ Carey (2009) habla de estados del sistema nervioso con contenido cuando se refiere a las representaciones mentales. Si hacemos el paralelo con los lenguajes naturales, las propiedades causales de los símbolos dependen de las propiedades sensoriales en las cuales se instancian.

el marco hegemónico de la ciencia cognitiva es funcionalista, entonces la forma se entiende como aquella propiedad del concepto en virtud de la cual tiene los poderes causales que exhibe. En cuanto gran parte de la propuesta ofrecida en esta tesis descansa sobre una caracterización más fina del papel de la forma en la individuación de los conceptos, a desarrollar principalmente en los capítulos 3 y 4, sólo señalo aquí esta manera provisional de caracterizar la forma del concepto.

Si nos quedamos en la aproximación a la composicionalidad que se centra en el contenido de las representaciones complejas, y asumimos lo ya dicho respecto de los símbolos primitivos, entonces tenemos un constreñimiento sobre el contenido conceptual de todos los símbolos del sistema⁴⁹: el contenido de cualquier representación mental debe ser heredado exhaustivamente de aquellos conceptos más básicos que lo constituyen⁵⁰. Dicho de otro modo, la única forma en que las representaciones complejas pueden adquirir el contenido que tiene es a través del contenido de sus constituyentes, y sólo a partir de ese contenido pues la sintaxis no lo afecta⁵¹.

⁴⁹ Todos los símbolos que juegan el papel de conceptos. Tal como lo digo en la nota anterior, en el capítulo siguiente distinguiré tipos de símbolos y se verá que lo dicho hasta ahora sólo es válido para aquellas representaciones mentales que están disponibles para ciertas computaciones, específicamente aquellas entendidas como inferencias que operan sobre símbolos semánticamente evaluables.

⁵⁰ Estoy asumiendo aquí lo que Carey (2009) llama la “Tesis de la uniformidad” respecto del aprendizaje de nuevos conceptos, según la cual la capacidad representacionista de un sistema no puede exceder la capacidad representacionista de sus primitivos más las reglas de combinación. Esta tesis es puesta en entredicho por Carey quien considera que es posible aprender nuevos conceptos cuyo contenido no es reducible a sus constituyentes, por medio de un tipo de aprendizaje que denomina *bootstrapping* quineano, similar a aquel proceso por el cual se forjan nuevos conceptos en las teorías científicas y que depende de inferencias no-demostrativas, pensamiento analógico, procesos abductivos, etc., además de estar íntimamente relacionados con la competencia en un lenguaje natural. Sobre la diferencia entre adquisición y aprendizaje de conceptos me pronunciaré en el capítulo siguiente.

⁵¹ Dado el ejemplo que he utilizado esto podría sonar más fuerte de lo que intento. Si se asume que la regla del condicional no es parte del conjunto de operaciones sino parte del conjunto de símbolos (Peacocke, 1992), entonces la sintaxis no jugaría ningún papel en la composición. Aun si ese fuera el caso, se debe reconocer que debe haber algún nivel de sintaxis/procesos que determinan el contenido de una expresión compleja. Lo importante es que el contenido de la expresión compleja está determinado por

1.4. Conclusión

En el presente capítulo he dado razones para preferir una teoría externista del contenido conceptual como un *desideratum* de la arquitectura computacional, entendida en el marco de un sistema de símbolos físicos. Además, he presentado la noción de composicionalidad como un constreñimiento ligado al supuesto computacional, apelando a las capacidades productivas y sistemáticas del sistema representacional. Finalmente, sostengo que si esa es la noción de composicionalidad requerida computacionalmente, entonces el contenido de las representaciones complejas debe ser heredado del contenido de las representaciones primitivas más las reglas de composición. Lo anterior nos lleva al requerimiento de una noción de contenido fuerte o “identidad de contenido”, en cuanto la participación de una representación en la composición de una representación compleja no altera su contenido original (y su aporte a la representación compleja, en cuanto al contenido, se limita a ese contenido original). Dicho de otro modo, el contenido de una representación debe ser invariable a través de las computaciones en las que participa. Así, en tanto el contenido de las representaciones simples es puramente referencial (externismo), también lo será el de las representaciones complejas⁵².

sus partes constituyentes, y que cada una de estas (representaciones y procesos) tiene ya un contenido determinado, y es ese contenido de las partes el que no cambia al estar integrado en la representación compleja.

⁵² A excepción del caso en que hubiera representaciones innatas cuyo contenido se determinara de manera puramente funcional. Si se aceptan este tipo de representaciones, y ellas participan en la formación de otras representaciones complejas, entonces no todas las representaciones tendrían un contenido puramente referencial. Exploraré esta opción en el capítulo 4.

En el siguiente capítulo examinaré la cuestión relativa a la caracterización de los conceptos primitivos: su forma y su contenido, su carácter innato o adquirido. Intentaré mostrar que es posible articular una noción que satisfaga los requerimientos que se siguen del desarrollo del presente capítulo: que su contenido sea caracterizado en términos externalistas y que su forma permita una semántica composicional como la descrita.

CAPÍTULO 2

Conceptos Primitivos

"Mucha gente tiene Sentimientos Muy Fuertes acerca de a qué conceptos se les permitiría ser innatos, por lo tanto, sobre cuán grande sería la base conceptual primitiva que podría reconocer una versión aceptable de la Teoría Representacionista de la Mente. Casi todo el mundo está dispuesto a permitir ROJO en dicha base, y muchos de los de mente liberal también permitirían CAUSA o AGENTE. Pero hay, en la actualidad, un fuerte consenso en contra de admitir en tal base a MANILLA (DOORKNOB) o CARBURADOR. No tengo ningún deseo de unirme a este juego, ya que, hasta donde yo sé, no tiene ninguna regla. Baste decir que sería bueno que una teoría de los conceptos proporcionara una explicación de principio de lo que está en la base conceptual primitiva, y sería bueno que los principios a los cuales apela establecieran la distinción de forma independiente en algún lugar que sea plausible." Fodor, 1998:28.

Si bien la discusión en teorías de conceptos es acerca del contenido de los conceptos léxicos, dado lo dicho en el capítulo previo resulta inevitable tomar una postura respecto de los conceptos primitivos. Primero aclararé dos formas en que puede concebirse la distinción primitivo/estructurado, para mostrar que es posible pensar los primitivos como teniendo componentes (2.1). Luego, en lo relativo a la discusión respecto de los primitivos y el innatismo, señalaré que aun cuando existan conceptos innatos (y en qué sentido se toman como innatos), si éstos agotaran la base primitiva y no tuviéramos que dar una teoría respecto de su contenido (representacional/intencional) tendríamos que explicar cómo surgen conceptos genuinamente representacionales a partir de ellos (2.2). Si, por otro lado, los conceptos innatos no agotan la base primitiva, entonces se requiere una explicación de los primitivos que no son innatos. En este punto resultará crucial la introducción del argumento de Fodor (1998, Cap. 5): Si hay conceptos primitivos no innatos, entonces se requiere de alguna forma de adquisición⁵³. Si la adquisición requiere de otros conceptos previos, entonces no hay primitivos de ese tipo. Si no hay primitivos entonces tampoco hay representaciones complejas y así, por tanto la teoría de fondo no es sostenible (2.3). Finalmente, dado que la adquisición no puede requerir de conceptos previos, resultará necesario apelar a mecanismos que no afecten al contenido (por lo dicho anteriormente) pero que medien la posesión (en este caso adquisición) de conceptos. Lo anterior llevará a considerar la distinción entre lo propiamente conceptual y ciertos mecanismos y/o representaciones,

⁵³ Utilizaré “adquisición” por cuanto la forma en que se constituyen estos conceptos dependerá, en última instancia, de la operación de mecanismos innatos, y por lo tanto no podrían ser explicados por un proceso de aprendizaje.

responsables en la adquisición de conceptos primitivos, que pueden ser caracterizados como no-conceptuales (2.4).

2.1. Representación Primitiva v/s Contenido estructurado

La mayor parte de la literatura dedicada a los conceptos se ha centrado en aquellos denominados “léxicos”⁵⁴ (aquellos que en los lenguajes naturales usualmente son expresados por sustantivos que funcionan como categorías que se predicán de individuos, p.e. PERRO, AGUA, PROTÓN, CARBURADOR), debido a que tales conceptos jugarían un papel prominente dentro de la vida mental, por lo que su explicación en términos de condiciones de individuación resulta necesaria para cualquier explicación de la naturaleza de la conducta inteligente⁵⁵.

Se pueden distinguir en primer lugar dos grandes grupos con relación a esta disputa: aquellas teorías que sostienen que los conceptos léxicos son estructurados y aquellas que niegan tal estructura. Dentro del primer grupo de teorías la disputa se centra en cómo se concibe tal estructura. Las teorías definicionistas plantean que existen relaciones necesarias entre los conceptos

⁵⁴ La excepción más notable es Peacocke (1992), cuyo interés se concentra en los conceptos que codificarían los operadores lógicos, específicamente las conectivas. A este respecto uno puede pensar tales conectivas como refiriéndose a la operaciones o reglas del sistema, o bien como a símbolos que codifican tales operaciones (Cf. Nota 49). En el primer caso la caracterización puramente funcional de las reglas de operación del sistema no amenaza una semántica puramente externista de los conceptos. En el segundo caso su compatibilidad con el *desideratum* externista que he propuesto dependerá tanto de si son tomados como símbolos primitivos o estructurados, como de si mi postura requiere que todo concepto sea individuado, en cuanto a su contenido, de forma externista.

⁵⁵ Dentro de una teoría representacionista/computacional de la mente como la que se supone en la presente investigación.

constituyentes y el concepto estructurado que constituyen. Por su parte, las teorías de prototipos y estereotipos sostienen que las relaciones de constitución son más bien en términos de inferencias estadísticas. Finalmente, la así llamada teoría-teoría sostiene que estas inferencias constitutivas son de orden no-demostrativo, similares a las relaciones constitutivas de los términos teóricos que figuran en las teorías científicas. A pesar de sus diferencias todas estas teorías parecen requerir de una base conceptual primitiva de conceptos no estructurados a partir de la cual se constituyan los conceptos léxicos⁵⁶. A la segunda postura, aquella que niega la estructura de los conceptos léxicos, se la conoce como atomismo conceptual, representada casi exclusivamente por Fodor⁵⁷. Otra forma de caracterizar el atomismo, no puramente negativa, es decir que éste supone una base conceptual primitiva que incluye a los conceptos léxicos, y por tanto es bastante grande el número de conceptos primitivos que esta teoría supone.

Debido a que las teorías que defienden el carácter estructurado de los conceptos léxicos no suelen pronunciarse respecto de los conceptos primitivos que suponen⁵⁸, en el presente capítulo tomaré como guía los desarrollos de Fodor (1998) por cuanto su posición explícitamente se hace cargo de la base conceptual primitiva. Cabe notar que no defenderé la tesis de que los conceptos léxicos son primitivos, sino que consideraré la caracterización de lo

⁵⁶ Una nueva versión de la teoría-teoría ha sido propuesta por S. Carey (2009) donde se aborda el problema de los conceptos primitivos y su relación con los conceptos complejos. En este caso ya no se habla de teorías (o mini-teorías) respecto de los conceptos primitivos, sino que más bien se asume que los conceptos a partir de los cuales se constituyen todos los otros conceptos son innatos, pero sin embargo su contenido se individúa tanto por algún tipo de relación causal con su referencia como con el papel inferencial asociado a su posesión.

⁵⁷ Laurence y Margolis se incluyen a sí mismos como exponentes de esta corriente, en su compilación de 1999.

⁵⁸ Exceptuando, como mencioné en la nota anterior, la propuesta de Carey (2009).

que es ser un concepto primitivo, centrándome en las restricciones que tal caracterización impone sobre cualquier teoría de las representaciones mentales dentro de lo que he llamado el marco hegemónico en ciencia cognitiva.

La primera aclaración pertinente a realizar es que, tal como ya lo hice notar en el capítulo anterior respecto de las discusiones relativas a la composicionalidad, al trazar la distinción entre estructurado y primitivo sólo se considera el contenido conceptual. Tanto Fodor como sus oponentes se limitan a caracterizar el contenido de los conceptos léxicos, argumentando ya sea a favor o en contra de que tal contenido sea estructurado, y en caso de serlo se discute acerca de la forma en que se realiza tal estructura. De acuerdo a esto, en sentido estricto el atomismo sólo dice que el contenido de los conceptos léxicos no es una función del contenido de otros conceptos (a partir de los cuales se estructuraría su contenido). De hecho, si vamos a considerar los conceptos como símbolos con contenido y forma, tal como se indicó en el capítulo precedente, entonces todo símbolo posee al menos dos elementos, pues al menos serían descomponibles en un elemento intencional (su contenido) y un elemento formal (su sintaxis)⁵⁹.

Por lo anterior, y considerando que lo que me interesa aquí es revisar las opciones para caracterizar los conceptos primitivos, debe tenerse presente que

⁵⁹ Si bien esta aclaración puede no parecer pertinente a la discusión, tenerlo en mente servirá para ver como una semántica puramente referencialista puede lidiar con los problemas de conceptos correferenciales distintos. Como se verá en el capítulo 4, uno de los argumentos de Prinz (2002) en contra del atomismo conceptual es que en tal teoría los conceptos carecen de la estructura suficiente como para explicar los papeles que debieran jugar en una teoría de las representaciones mentales, específicamente respecto de la categorización.

los conceptos primitivos son tales en cuanto su contenido y su sintaxis no son una función del contenido y sintaxis de otros conceptos. Es eso lo que los hace primitivos, y no el que carezcan por completo de partes distinguibles. (Dentro del espacio lógico es posible pensar conceptos innatos no-primitivos, en este sentido ver 2.2).

Sin embargo, pese a esta caracterización de los conceptos primitivos, relativa al contenido de los mismos, hay otra relación que suele asumirse cuando se teoriza acerca de ellos, a saber, que un concepto primitivo es innato. Si bien el uso de “innato” aquí sólo parece como contrario a “aprendido”, debe notarse que las discusiones acerca de los usos y significados de “innato” en ciencia cognitiva han dejado ver que no hay una formulación estándar. Cowie (1999) y Prinz (2002, Cap.8) identifican distintos sentidos de “innato” utilizados en la literatura respecto de capacidades cognitivas⁶⁰. La noción de innato que aquí quiero rescatar corresponde a aquella según la cual un concepto es no aprendido en el sentido descrito anteriormente, y que lleva a la identificación de “innato” con “primitivo” para los conceptos.

Conviene aclarar que la noción de “aprendido” tampoco resulta unívoca. La discusión tradicional respecto del innatismo entre empiristas y racionalistas era acerca de si había “ideas” innatas. Esta discusión se ha traspasado a la

⁶⁰ Así, está la diferencia entre “presente al nacimiento” v/s aquello que depende de la experiencia post-natal; otras veces por innato se pretende implicar que un rasgo está genéticamente determinado, que la norma de reacción de un rasgo es casi invariable a través de distintos ambientes; en la misma dirección relativa a destacar la prácticamente nula influencia del ambiente, algunos autores utilizan la noción biológica de “canalización”, en el sentido de que el desarrollo de un rasgo aparece como independiente de cambios en el ambiente; otra acepción es la de “genéticamente arraigado”, según la cual un cambio en un rasgo innato provocaría cambios mayores en el sistema al cual pertenece; por último, otra forma de concebir la distinción entre innato y aprendido es apelando a “mecanismos disparadores” del rasgo que se activan en presencia de las experiencias adecuadas.

ciencia cognitiva en la forma de si es necesario postular conceptos innatos. Para los empiristas todos los conceptos deben ser aprendidos, mientras que para los innatistas los conceptos primitivos dependen de un proceso llamado comúnmente adquisición. La adquisición no dice que los conceptos estén presentes desde el nacimiento, sino sólo que en su formación participan mecanismos que son ellos mismos innatos, que están genéticamente determinados y que dependen para su activación de una experiencia adecuada⁶¹.

No obstante, un empirista como Prinz adhiere a la postura de que todos los conceptos son aprendidos por medio de la experiencia. Sin embargo, acepta que algunos conceptos deben ser primitivos. Así, habría que distinguir al menos dos formas de aprendizaje: la ya descrita anteriormente que cuenta para la formación de representaciones complejas, por una parte; y por otra el aprendizaje de los primitivos a partir de la experiencia. Aquí la opción de Prinz es aceptar un innatismo débil, esto es aceptar que deben haber mecanismos o representaciones no conceptuales (perceptuales) innatas que sirven de base al aprendizaje de los conceptos primitivos. Pero esta distinción parece volver a confundir la discusión entre empiristas e innatistas: el “innatismo débil” del que habla Prinz es contrario al “innatismo” del que habla Carey (2009), pues según ella un concepto es innato si su adquisición no depende de la existencia de otras representaciones, sean estas conceptuales o no. Además, si aceptamos el innatismo débil entonces la distinción entre adquisición y aprendizaje resulta inútil, al menos para los conceptos primitivos.

⁶¹ Sólo enumero aquí las características relevantes para la discusión. Otra característica es tener un tiempo de maduración y un rango de tiempo para su desarrollo. En general, estas características se desprenden de los argumentos de pobreza de estímulo comunes en la literatura sobre innatismo.

En lo que sigue consideraré a un concepto como innato en el sentido que Carey (2009) propone, esto es que su adquisición no esté mediada por ningún tipo de representación.

En los apartados siguientes intentaré mostrar que la relación entre primitivo e innato no es constitutiva (ni conceptualmente ni de ninguna otra forma), y que además resulta necesario rechazarla, al menos a un cierto grado, pues se requiere de conceptos primitivos no innatos para dar cuenta del carácter genuinamente representacional de los conceptos (2.2). Para conseguir lo anterior será necesario caracterizar el innatismo como relativo al no-aprendizaje de los conceptos primitivos, en el sentido descrito, y así un concepto primitivo puede ser no innato si su adquisición no requiere de conceptos previos pero sí de la experiencia con instancias de su contenido y mecanismos adecuados para su formación.

2.2. Innatismo 1: poder representacional

Como señalé en el apartado anterior, la mayor parte de los desarrollos en teorías de conceptos tratan sobre conceptos léxicos, y el problema de los primitivos queda prácticamente sin considerar. Ya que el atomismo conceptual está comprometido con la caracterización de los léxicos como primitivos,

tomaré esta formulación como la aproximación más desarrollada de los mismos.

Uno de las principales objeciones que enfrenta el atomismo conceptual es que parece estar comprometido con un innatismo que a todas luces resulta insostenible, justamente por la relación que suele establecerse entre conceptos primitivos e innatismo. Lo que me interesa aquí no es la plausibilidad del atomismo como teoría de los conceptos léxicos, sino la inferencia que se hace desde el ser primitivo al ser innato, lo que Fodor llama el “argumento estándar”. Tal argumento supone que la única forma de aprender un concepto es combinando otros conceptos que ya se tienen. Así, por ejemplo, para aprender el concepto CUADRADO debo pensar que se aplica a cosas que tienen cuatro lados, que son figuras planas, que todos sus lados son iguales y que sus ángulos son rectángulos. Como resulta obvio, aprender el concepto CUADRADO requerirá de la posesión previa de los conceptos COSA, CUATRO, LADOS, FIGURA, etc. Ahora bien, si ésta es la forma de aprender conceptos, entonces un concepto primitivo que no es descomponible en otros conceptos tendrá necesariamente que ser no aprendido, es decir innato. Sobre este argumento volveré luego de revisar otras propuestas sobre la caracterización de la base conceptual primitiva.

La mayoría de los teóricos de los conceptos están de acuerdo en que tienen que haber conceptos innatos y/o primitivos, pero incluso dentro de este consenso existen serias discrepancias:

- a) Existen los que suponen que los primitivos no son innatos, principalmente dentro de la tradición empirista, con la restricción de que su posesión y adquisición debe explicarse sólo en términos de ciertas experiencias sensoriales cuyo contenido no es analizable en conceptos más básicos⁶².
- b) Existen los que suponen que los conceptos primitivos son innatos y corresponden a categorías abstractas que nos permiten categorizar el mundo: AGENTE, CAUSA, EVENTO y OBJETO por ejemplo son parte de la base conceptual primitiva admitida por estos teóricos⁶³.
- c) Es posible pensar que los conceptos innatos no sean primitivos, por ejemplo que se interdefinan. En este caso no habría conceptos primitivos propiamente tal, al menos en el sentido que he propuesto⁶⁴.
- d) Finalmente, otra opción sería pensar que los primitivos no necesariamente se aprenden a partir de otros conceptos, sino que se adquieren de otra forma que no requiera de la participación de

⁶² Si bien a excepción de Prinz (2002) nadie se declara abiertamente como empirista respecto de los conceptos, en esta postura se suele incluir a pensadores tan distantes como Locke, Quine y Piaget (Según Carey, 2009). Por otra parte, dentro de las discusiones en psicología cognitiva normalmente los defensores de la teoría-teoría acusan a los teóricos de los prototipos y ejemplares de ser empiristas en cuanto su análisis de los conceptos siempre remite a rasgos observables o experienciales (Murphy, 1998)

⁶³ En esta categoría se pueden incluir a teóricos definicionistas como Jackendoff, así como a defensores de la teoría-teoría como Carey. Un antecedente a estas teorías puede encontrarse en la formulación de las 12 categorías del entendimiento en Kant.

⁶⁴ Claramente en este grupo caen defensores tradicionales de la teoría-teoría como Gopnik y Meltzof, así como definicionista como Pinker. Los primeros piensan que incluso los conceptos más básicos se encuentran en una relación de interdefinición con otros conceptos básicos en forma de mini-teorías. Para Pinker, por otro lado, algunos conceptos son aprendidos a partir de otros pero su contenido no es analizable en términos de los últimos, pues habría un proceso de *bootstrapping* en el cual el nuevo concepto exhibe propiedades que no están presentes en sus constituyentes. El ejemplo favorito de Pinker es el *bootstrapping* semántico, donde se puede aprender el contenido de un concepto a partir de la sintaxis que acompaña a su expresión en el lenguaje.

conceptos. En este caso no se requiere, necesariamente al menos, de la restricción que opera para la posición descrita en a)⁶⁵.

En lo que sigue de este apartado quiero mostrar que, dadas las restricciones alcanzadas en el capítulo anterior, las tres primeras posibilidades no las satisfacen, y que por lo tanto debe haber al menos una cantidad de conceptos primitivos que no sean innatos. En el apartado siguiente revisaré el argumento estándar contra el atomismo conceptual y las consecuencias que se siguen de él para cualquier teoría de conceptos dentro del marco de una teoría representacional/computacional de la mente. Finalmente, la caracterización de la base conceptual primitiva conseguida, más las restricciones de la teoría de fondo sobre la naturaleza de los conceptos en general, permitirá ofrecer razones para preferir la posibilidad d), expresada más arriba, sobre la caracterización no primitiva c).⁶⁶

Una versión de la base conceptual primitiva suele ser ofrecida sobre bases empiristas. Según estas teorías no se requiere que los primitivos sean innatos, pues bastaría que nuestra mente estuviera constituida de tal forma que fuera capaz de adquirir conceptos por medio de mecanismos sensibles al entorno en el que se desenvuelve el individuo. Este papel lo podrían jugar nuestros sentidos entendidos como analizadores de insumos sensoriales, por lo cual no sería necesario contar con representaciones previas para poder

⁶⁵ No he incluido dentro de las opciones la versión de Fodor anterior a 1998, donde la base conceptual primitiva, que incluye a la mayoría de los conceptos léxicos, es a su vez innata. Tal innatismo radical es suavizado por Fodor en su exposición de 1998, y es esta versión la que presentaré en el apartado siguiente. Como ya lo he señalado, tomaré la versión de Fodor a fin de presentar una opción de conceptos primitivos no-innatos, si bien la postura que defiende termina siendo incompatible con el Atomismo respecto de la mayoría de los conceptos léxicos.

⁶⁶ Otra forma de concebir esta opción es como consistiendo de una mezcla de las tres anteriores.

llegar a formar conceptos. Lo anterior sin embargo nos obliga a asumir que el contenido de cualquier concepto debe ser en último término rastreable a relaciones entre nuestros sentidos y las experiencias adecuadas.

En este tipo de teorías resulta crucial la distinción entre dos formas de aprender un concepto. Por un lado está la forma tradicional de concebir el aprendizaje de conceptos complejos, esto es como un proceso de testeo de hipótesis acerca del contenido del nuevo concepto (basado en asociaciones regulares dadas en la experiencia, en el caso de los empiristas). Como ya se vio, esta forma pareciera requerir de conceptos innatos, contrario a la postura empirista tradicional: en este proceso uno debiera formular una hipótesis acerca de los rasgos constitutivos del concepto a aprender, ya sean definitorios o prototípicos, y luego ver si las instancias del concepto (es decir, los objetos a los cuales se aplica el concepto) poseen estos rasgos. Por ejemplo, si creo que algo para ser un cuadrado debe tener cuatro lados, o algo para ser categorizado como rojo debe tener ciertas cualidades o rasgos sensoriales, entonces 4 y LADO para CUADRADO, o las representaciones mentales de los rasgos sensoriales para ROJO, deben estar ya disponibles para poder formar la hipótesis correspondiente. Otra manera de decir lo mismo consiste en señalar que un concepto aprendido debe ser, necesariamente, complejo⁶⁷. Por lo mismo el aprendizaje de conceptos requiere de la posesión previa de otros conceptos. Así, el empirista requiere de alguna forma de aprender los conceptos primitivos que no requiere de representaciones previas, a fin de mantener su postura de que todos los conceptos son aprendidos.

⁶⁷ De allí la conexión entre primitivo e innato.

Una solución que ya fue indicada es la de Prinz y su “innatismo débil”. Sin embargo, en su disputa con los empiristas, Carey (2009) señala que en su concepción las representaciones innatas son aquellas que no son el resultado de procesos de aprendizaje, y estos últimos se definen como aquellos que construyen representaciones del mundo a partir de computaciones sobre insumos que son ellos mismos representacionales (sean éstos de naturaleza conceptual o no). De este modo, en la propuesta de Carey los conceptos que cuentan como innatos pueden no estar desde el nacimiento, ni ser totalmente independientes de tener la adecuada experiencia, y lo relevante para que un concepto cuente como innato es que la capacidad para construir la representación sea innata, en el sentido de que no requiera de otras representaciones para cumplir su tarea, y que funcione dada una adecuada maduración y/o experiencia⁶⁸.

Pese a las ventajas que pueda exhibir en cuanto al rechazo de las tesis empiristas, por ejemplo respecto de la imposibilidad de aprender el concepto OBJETO, la distinción de Carey entre aprendido e innato parece demasiado restrictiva: su propuesta requiere que los conceptos innatos o primitivos no provengan de ningún tipo de representación mental, sea ésta conceptual o de otro tipo⁶⁹. Dadas esas restricciones, resultaría difícil aislar suficientemente

⁶⁸ Esta caracterización de Carey respecto del innatismo parece más cercana a la noción tradicional de adquisición utilizada por los innatistas (Cfr. Chomsky, 1959).

⁶⁹ Carey (2009) distingue al menos tres tipos de representaciones mentales, a saber: sensoriales, perceptuales y conceptuales. La distinción, además de apelar al formato, remite a la capacidad de la representación de participar en inferencias, o en sus propios términos del papel inferencial asociado a tales representaciones. Creo que el negarle cualquier papel a otro tipo de representaciones en la formación de los conceptos se debe a su caracterización del aprendizaje, no distinguido de la adquisición en los términos que aquí se señalan. Más bien para ella el aprendizaje es lo que estoy llamando aquí adquisición de primitivos, y reserva el término adquisición para el proceso de crear nuevos conceptos que exceden el

algún concepto primitivo como para que contara como innato, pues la participación de representaciones perceptuales y/o sensoriales en la adquisición de tales conceptos parece ineludible.

Por mi parte, lo que estoy llamando aquí adquisición de conceptos primitivos sólo requiere de algún tipo de mecanismo y/o representación, no conceptual, encargado de responder selectivamente a información que proviene del entorno y a la cual accede el individuo por medio de sus sentidos. En este sentido mi posición puede ser vista como cercana al innatismo débil de Prinz, pero por las razones ya esgrimidas prefiero seguir reservando el término “innato” en el sentido de Carey.

Volviendo a las teorías que suponen una base conceptual primitiva no-innata, el candidato que a menudo se elige como concepto primitivo corresponde a algunas de las cualidades secundarias, siguiendo la terminología de Hume, o conceptos de apariencia, como ROJO. No es necesario postular que aprendemos ROJO, por ejemplo a partir del concepto previo COLOR más las experiencias que tenemos con objetos de color rojo, ni tampoco resulta necesario postular que nacemos con ese concepto. Lo único que se requiere es que nuestra mente esté diseñada de tal manera que haya ciertos mecanismos cognitivos, no conceptuales ellos mismos, encargados de hacer que ante experiencias de rojo de manera confiable nuestra mente concluya formando el concepto ROJO⁷⁰.

poder representacionista de sus constituyentes, por medio del proceso que ella denomina *bootstrapping* quineano.

⁷⁰ En inglés se suele hablar de un *trigger* o disparador para nombrar este tipo de mecanismo.

Así, la base conceptual primitiva estaría constituida por aquellos conceptos que se refieren a rasgos del entorno a los que puede acceder el individuo por medio de sus sentidos. Luego, todos los conceptos que podríamos formar a partir de dicha base tendrían que ser en principio rastreables a conceptos que refieren a experiencias sensoriales.

La objeción clásica a este tipo de teorías es que claramente hay conceptos complejos cuyo contenido no corresponde a la suma de puros conceptos sensoriales. Por ejemplo, conceptos como CAUSA, DEMOCRACIA, VOLUNTAD, DIOS, parecen en principio no corresponder a ningún conjunto de rasgos sensoriales que pudieran dar cuenta de su contenido. Ahora bien, dada la noción de composicionalidad para un sistema de símbolos físicos alcanzada en el capítulo previo, forzosamente cada concepto complejo (o no primitivo dada la noción aquí defendida de concepto complejo/estructurado), en cuanto a su contenido, debe ser descomponible en términos de la base conceptual primitiva, lo que resulta en un problema para la versión empirista de los conceptos primitivos⁷¹.

Una forma de enfrentar el problema que trae consigo postular una base conceptual primitiva de carácter puramente empirista consiste en sostener que existen conceptos innatos y primitivos que refieren a categorías abstractas. Ya que el problema del empirismo era dar cuenta de conceptos abstractos que no

⁷¹ Prinz (2002) ofrece una defensa de la teoría empirista de los conceptos que intenta soslayar este problema. Sin embargo, en cuanto creo que parte de lo que expone ahí puede ser tomado como una fuente de argumentos a favor de lo que defenderé en el capítulo 4, dejaré la revisión de su propuesta para ese momento. Por otro lado, cabe hacer notar desde ya que la teoría de los conceptos de Prinz supone una arquitectura mental distinta de la que sirve de marco para esta investigación, a saber el conexionismo, por lo que aun cuando tomo algunos de sus argumentos, no es posible considerarla en su totalidad dentro de la discusión aquí revisada.

parecen poder ser reducibles a abstracciones a partir de conceptos sensoriales, la alternativa de postular ciertos conceptos abstractos como primitivos resulta obvia. Algunos autores, Jackendoff por ejemplo, parecen sugerir que ciertas categorías como „suceso’, „agente’ y „causa’, no son definibles en términos de otros conceptos, y más bien se requieren para poder dar cuenta del poder representacional de la mente, y así poder cubrir los casos que el empirismo de los conceptos dejaba fuera.

Sin embargo, si uno toma esta posición en su versión que podríamos denominar “pura”, esto es donde todos los conceptos primitivos son estas categorías abstractas, y son innatos, entonces pareciera que debiera renunciarse al compromiso con el contenido intencional de los estados mentales⁷². Como se expuso en el capítulo 1, la posibilidad de caracterizar el contenido de los conceptos en términos puramente de relaciones locales, internas al sistema representacional, no cumpliría el *desideratum* para el sistema de símbolos físicos como modelo de la mente. Además, una base conceptual primitiva innata del tipo postulado no tiene el poder representacional de una teoría externista del contenido. Como Dretske (1986) lo señaló, una de las características de la representación es su capacidad de representar erróneamente. Sin embargo, si la base conceptual primitiva se agota en estas representaciones de categorías abstractas, entonces no es capaz de entregar representaciones erradas, pues en este caso el contenido de las

⁷² Nótese que aquí menciono una versión pura de esta postura que apele en su caracterización del contenido sólo a los papeles inferenciales. Si por el contrario se considerará que la determinación del contenido de estas representaciones primitivas dependiera de la etiología de los mismos, de algún tipo de dependencia nómica respecto de sus causas, del cumplimiento de alguna función biológica seleccionada por evolución, o una mezcla de estas opciones, entonces tal postura sí podría cumplir con el *desideratum* externista.

representaciones complejas, al estar determinado por relaciones locales a partir de los contenidos de la base conceptual primitiva, determinaría la referencia de tales conceptos, cayendo así en lo que en el capítulo 1 señalé como el “problema de la ignorancia y el error”.

Del mismo modo, si se postula que la base conceptual primitiva está constituida por un número finito de conceptos innatos, correspondientes por ejemplo a las categorías abstractas anteriormente mencionadas, pero que sin embargo tales conceptos no son primitivos⁷³, pues se interdefinirían entre ellos, entonces también parece perderse el poder representacional/intencional que requiere de alguna forma de externismo respecto del contenido⁷⁴.

Podría objetarse que la postulación del innatismo de algunos conceptos no requiere necesariamente del compromiso con una semántica internista para los mismos. Por ejemplo, podría contarse una historia acerca de cómo la posesión de esos conceptos es un logro evolutivo y por tanto su contenido es sensible (refiere a) al entorno. Una opción para estas teorías sería comprometerse con el hecho de que todas estas representaciones primitivas son producto de la selección natural, asegurando así una semántica externista para las mismas, lo cual es perfectamente coherente con lo que aquí defiendo. Carey (2009) de hecho ofrece diversa evidencia experimental para mostrar que los conceptos OBJETO y AGENTE al menos cumplen con esta característica.

⁷³ Al igual que en la nota anterior, estoy caracterizando versiones “puras”. Tal vez sea posible, incluso necesario, admitir en la base conceptual primitiva tanto conceptos adquiridos que expresan propiedades sensoriales como conceptos abstractos innatos. Además, lo que me importa aquí es señalar las características que son impuestas por las teorías para cualquier concepto que va a ser categorizado como primitivo, antes que tomar una decisión respecto de si tal o cual concepto en particular es primitivo o no.

⁷⁴ Como señale anteriormente, es posible aceptar que algunos conceptos se individúen de esta manera. Sin embargo, si acepta el *desiderátum* externista del capítulo 1, entonces tales conceptos no pueden agotar la base primitiva.

Sin embargo, cuáles sean los conceptos que constituyen la base conceptual primitiva es un asunto empírico que va más allá del alcance de esta tesis, por lo que prefiero limitarme a considerar las posibilidades y restricciones⁷⁵ que se siguen de cada una de estas posturas. En tanto no hay un catálogo completo de la base conceptual primitiva ofrecido por alguna de las teorías revisadas, mis consideraciones se centran en mostrar sus limitaciones o incompatibilidades con el marco de fondo, cuando es posible hacerlo. No creo que haya un argumento definitivo en contra de una versión del innatismo conceptual de este tipo, más bien sería la investigación empírica la que podría revelar finalmente cuáles conceptos constituyen la base conceptual primitiva, y cuáles de esos conceptos son innatos en ese sentido. Lo que pienso es que si es posible no asumir que todo concepto primitivo es innato y fue seleccionado naturalmente a través de la evolución, entonces resulta mejor intentar elaborar esa opción.

2.3. Innatismo 2: Aprendizaje v/s Adquisición

Lo que antes llame la opción d) respecto de la base conceptual primitiva sostiene que los conceptos que la constituyen son primitivos, en cuanto a no requerir de otros conceptos para su posesión, y que no son innatos⁷⁶. La

⁷⁵ Posibilidades relativas al cumplimiento del *desiderátum* externista, por un lado, y restricciones relativas al constreñimiento del principio de composicionalidad. En ese sentido, en este capítulo específicamente me refiero a las posibilidades, y en el capítulo 3 me centraré en las restricciones al evaluar las teorías basadas en los experimentos de categorización.

⁷⁶ Más precisamente, que al menos parte de la base conceptual primitiva no es innata. Podría aceptarse la existencia de los conceptos abstractos de b), agregando además los conceptos no-innatos. Los argumentos

formulación de la opción ya parece paradójica, y un defensor del atomismo conceptual como Fodor debe hacerse cargo de la misma. La revisión de tal defensa me permitirá caracterizar de mejor forma la base conceptual primitiva y por medio de esto hacer explícita la diferencia que tiene con mi propuesta.

Brevemente, la teoría de Fodor se basa en la idea de que el contenido de un primitivo conceptual está determinado por la posición del concepto en una relación causal apropiada con las cosas en el mundo. Para Fodor, la relación causal es una conexión nómica entre tipos de conceptos y las propiedades que expresan sus símbolos (*tokens*). Por ejemplo, el contenido del concepto PÁJARO no está dado por su relación con conceptos como ANIMAL, ALAS, etcétera. En cambio, PÁJARO expresa la propiedad *pájaro*, en parte, porque existe una ley causal que asocia la propiedad de ser un pájaro con el concepto PÁJARO. Hasta esta parte de la teoría, la explicación de Fodor se ubica claramente en la tradición semántica basada en la información, de acuerdo a la cual el contenido mental es un tipo de contenido informacional (Dretske, 1981). La información es básicamente un asunto de correlaciones confiables. Una correlación confiable supone una conexión legaliforme entre dos sucesos, esto es que hay una cierta relación causal que gobierna la correlación. Cuando un tipo de suceso es una causa confiable de otro, se dice que el segundo incluye información respecto al primero. De este modo, para Fodor, el contenido mental requiere que un concepto contenga información respecto de la propiedad que expresa. Pero hay más respecto al contenido mental que la información. Como es ampliamente reconocido, existe una variedad de casos

aquí desarrollados son de carácter filosófico, sobre las posibilidades conceptuales de un marco de investigación.

donde un concepto es un efecto confiable de cosas que no están en su extensión. El caso estándar de este tipo es una situación en que una aplicación errónea de un concepto es, por alguna razón, confiable. Considérese, por ejemplo, una situación en que las condiciones visuales son precarias. Es una noche oscura, quizás con un poco de niebla, y uno piensa que ha visto una vaca en el campo al otro lado del camino. En otras palabras, uno aplica el concepto VACA a la entidad de más allá, y uno lo hace porque tiene buenas razones para hacerlo: parece una vaca. Sin embargo, es un caballo: uno ha aplicado incorrectamente su concepto. Esta situación es esperable en condiciones como ésta, puesto que bajo las condiciones que estamos imaginando, el caballo efectivamente parece una vaca. Si aceptamos que nuestra instanciación del concepto VACA puede ser el resultado de dos sucesos en el mundo, y que esta correlación puede resultar confiable, como parece sugerirlo el ejemplo, el resultado es que nuestro concepto VACA es el efecto confiable de al menos dos causas: vacas y caballos. Sin embargo, si no existe más que la información para el contenido, no tendríamos un caso de error en absoluto, sino que una aplicación verídica de un concepto que expresa la propiedad disyuntiva *vaca o caballo*. En los círculos filosóficos, este tema ha sido denominado El Problema de la Disyunción, y es un problema que cualquier teoría (externista y) causal del contenido debe enfrentar.

Los semánticos que se basan en la información han explorado una gran cantidad de maneras para superar el Problema de la Disyunción. La solución de Fodor es afirmar que ciertas relaciones informacionales son más básicas que otras y que es ésta la diferencia que cuenta. Su teoría tiene dos partes:

- (1) Un concepto (VACA, por ejemplo) aparece en una relación legal, L , respecto a la propiedad que expresa, a saber, *vaca*.
- (2) Otras relaciones legales que involucran VACA, $L1-Ln$, son dependientes asimétricamente según la relación legal entre VACA y *vaca*. En otras palabras, $L1-Ln$ no incluirían sino lo que L incluye, y no al revés.

Por lo tanto, la diferencia crítica entre la ley *vaca* / VACA y la ley *caballo* / VACA es que, aunque ambas son confiables, la primera es más fundamental: se puede obtener incluso si la dependencia *caballo* / VACA no se pudiera obtener, mientras que la dependencia *caballo* / VACA no se obtendría sin la dependencia *vaca* / VACA. Es por eso que VACA expresa la propiedad *vaca* y no, como podría darse, *vaca o caballo*.

Además de la fuerte carga ontológica que supone la teoría informacional de Fodor, al ponerla en conjunción con su atomismo respecto de los conceptos léxicos, la objeción que a primera vista salta es la siguiente: como los conceptos primitivos no son aprendidos, entonces deben ser innatos. Pero, ¿cómo, en el caso del atomismo conceptual, podría CARBURADOR ser innato⁷⁷? Lo anterior constituye lo que se ha llamado el innatismo radical de Fodor.

⁷⁷ Sigo aquí la argumentación de Fodor 1998, por ello aparece CARBURADOR, pues para Fodor tal concepto es primitivo. Para mis propósitos no es necesario asumir tal postura, sino sólo revisar lo que dice Fodor acerca de la posibilidad de los primitivos y su relación con el innatismo. De hecho, finalmente el modelo que tomará Fodor es el de los conceptos de apariencia (p.e. ROJO), lo que está más en consonancia con lo defendido en mi investigación.

Sin duda que este problema debe ser solucionado, pero antes dejemos en claro, y téngase presente en lo que sigue, que cualquier teoría sobre conceptos postula una base conceptual primitiva puesto que es un requisito de la arquitectura asumida⁷⁸, por lo que la solución a este problema implica consecuencias para cualquiera de tales teorías, sentando las reglas del juego que se echaban de menos en la cita que inicia este capítulo. De hecho, el problema del innatismo surge para cualquier teoría de los conceptos: si el contenido de un concepto es fijado recurriendo al contenido de sus constituyentes, entonces el contenido de los conceptos primitivos no puede ser fijado respecto de otros (contenidos de) conceptos, por ende, deben ser innatos⁷⁹.

El atomismo informacional⁸⁰ tiene una alternativa al innatismo radical. En primer término, esta teoría es claramente no-cognitivista respecto de la posesión de conceptos (respecto al contenido). Para esta postura tener un concepto no es conocer algo sino estar en una cierta relación mente-mundo. Tal relación es justamente aquella en virtud de la cual el concepto tiene el contenido que tiene. Así, en tanto el atomismo informacional es no-cognitivista respecto de la posesión de conceptos, esto es que poseer un concepto no involucra estar dispuesto a realizar ciertas inferencias, o a saber ciertas cosas acerca del referente del concepto, del mismo modo la noción de adquisición de

⁷⁸ Incluso una postura como la de Prinz, que finalmente rechaza la arquitectura computacional-serial en favor de una conexionista, requiere de la base conceptual primitiva, aunque en su caso ésta se remite a representaciones perceptuales cuyo contenido está limitado a las apariencias de los objetos que caen bajo tales representaciones.

⁷⁹ Dada la idea de aprendizaje que subyace a estas discusiones.

⁸⁰ Atomismo Informacional es la denominación que el propio Fodor (1998) da a su teoría: el atomismo respecto de los conceptos léxicos combinado con una teoría informacional respecto del contenido. Esta parte de la teoría de Fodor es la que explica su versión no-cognitivista de la adquisición de conceptos, o en los términos de esta tesis, la distinción entre adquisición y aprendizaje que ya se ha mencionado.

conceptos es no-cognitvista (Cf. Fodor 1998:124). La propuesta de Fodor, para evitar el innatismo, es que adquirir un concepto es estar “nomológicamente vinculado (*locked*⁸¹)” con la propiedad que ese concepto expresa, por tanto, la posesión de un concepto es algún tipo de vinculación. Si bien la noción de “vínculo” puede ser caracterizada en términos simples como en el caso de la teoría semántica de la referencia directa, en Fodor, al igual que en su teoría semántica de la dependencia asimétrica, tiene fuertes compromisos ontológicos. Así como uno podría decir que dos átomos de elementos químicos quedan vinculados en ciertas circunstancias de acuerdo a las leyes de la química, del mismo modo Fodor nos invita a pensar que el vínculo entre una propiedad y el contenido del concepto que la expresa surge de acuerdo a ciertas leyes del mundo, específicamente de leyes acerca de nuestras mentes. No hay capacidades epistémicas involucradas por parte de los individuos que tienen mente, ni tampoco hay una historia evolutiva que asegure ese vínculo por parte del mundo. Fodor lo considera un hecho básico acerca de nuestras mentes, y es un compromiso ontológico que acepta dada la teoría que defiende⁸². Así puestas las cosas, la pregunta a responder es si acaso el atomismo informacional requiere, de manera esencial, del innatismo radical.

⁸¹ En una formulación breve, la noción del *locking* expresa una conexión causal privilegiada entre la propiedad y el concepto, por ser la primera el contenido del segundo.

⁸² Esta es la forma de caracterizar la noción en el capítulo donde la introduce, y la utiliza para dar cuenta de propiedades “dependientes de la mente”, como sería el caso de *espejo*. Por otro lado, no es menos cierto que en el capítulo que sigue a su introducción de la noción de “vínculo” (1998, Capítulos 6 y 7 respectivamente), Fodor parece volver a una noción más simple, y a la vez inexplicada, de “vinculo”, cuando al hablar de los conceptos de clase natural “como conceptos de clase natural”, esto es, no dependientes de nuestras mentes, considera que no puede estar “vinculado” a las propiedades que expresan estos conceptos de diversas formas, por medio de aprender una teoría, usando un microscopio, o simplemente escuchando a un experto hablar acerca de ella. Sin embargo, pese a lo anterior, Fodor señala que el “vínculo” en cuestión es independiente de la forma en que se realice, y que el hecho de que se dé es una cuestión nomológica.

El modelo del vínculo permite negar, entonces, que el adquirir conceptos involucra la adquisición de creencias. El error que quiere destacarse es que el inductivismo⁸³ acerca de la psicología de la adquisición de conceptos está motivado en primer lugar por la metafísica asumida respecto de la posesión de conceptos, esto es, el cognitivismo asociado a ésta. De acuerdo a este modelo es posible afirmar la existencia de conceptos que sean primitivos pero no innatos.

Ahora bien, si se acepta que al menos parte de los conceptos primitivos no son innatos, entonces debe haber una historia acerca de su adquisición que no implique aprendizaje en el sentido expuesto anteriormente (2.2). En particular, debe haber una historia que sea capaz de dar cuenta de otro problema que surge una vez que se acepta la posibilidad de que los primitivos no sean innatos sino adquiridos, a saber, que sea capaz de dar cuenta de la relación evidencial que parece mediar tal adquisición. Si bien pareciera que la propuesta de Fodor puede lidiar contra el argumento del innatismo radical o argumento estándar, aún queda otro problema por explicar, a saber, que aun cuando se tenga una concepción no cognitivista de la posesión de conceptos y no se presuponga una tesis inductivista general acerca del rol del testeo de hipótesis en la adquisición de conceptos, todavía queda el asunto más simple relativo a que lo que nos lleva a adquirir un concepto típicamente consiste en tener el tipo adecuado de experiencias. Para ponerlo de una vez en palabras más simples, aún hay que explicar cómo es que el concepto TAL⁸⁴ es adquirido

⁸³ “Inductivismo” es el término de Fodor para caracterizar la forma de aprendizaje que lleva a la postulación del innatismo de los primitivos.

⁸⁴ Prefiero ya en esta sección dejar de lado el ejemplo de Fodor, en este caso MANILLA, para alejarme definitivamente del atomismo respecto de los conceptos léxicos.

justamente en virtud de experiencias con Tales y no con Cuales. A este fenómeno Fodor lo denomina el "efecto t/T". Todo parece indicar que el modelo inductivo de aprendizaje de conceptos vuelve a aparecer. De acuerdo con este modelo, la relación entre el contenido de los conceptos que uno adquiere y el contenido de las experiencias que actúan en la adquisición de ellos es de carácter evidencial, particularmente es mediada por las relaciones de contenido entre una hipótesis y las experiencias que sirven para confirmarla.

La forma tradicional de explicar el efecto t/T era apelar al inductivismo respecto de la adquisición, incluidos los casos de los conceptos primitivos. Esto resulta un problema para cualquier teoría de conceptos dentro del marco de la teoría representacional de la mente, por lo siguiente: el argumento estándar del innatismo dice que los conceptos primitivos no pueden ser aprendidos inductivamente, pero, por otro lado, cualesquiera que sean los conceptos primitivos que se escojan, todos ellos exhiben el efecto t/T en cuanto a su adquisición, en lo relativo a la relación entre el contenido del concepto y el contenido de la experiencia que ocasiona su adquisición.

Frente a lo antes expuesto el atomismo informacional tiene sólo una posibilidad: si uno quiere asumir que en la adquisición de conceptos la relación entre la experiencia y el concepto adquirido no es evidencial (de acuerdo al modelo inductivista de testeo de hipótesis), entonces sólo se puede apelar a una causalidad subyacente que dé cuenta del efecto t/T, i.e. de la relación entre los "tales" del mundo y el concepto TAL.

Aquí Fodor ve un problema fundamental: si el efecto t/T muestra que los conceptos primitivos deben ser aprendidos inductivamente, y el argumento estándar muestra que los primitivos no pueden ser aprendidos inductivamente, la conclusión es que no hay conceptos primitivos. Pero si no hay conceptos primitivos, entonces no hay conceptos en absoluto, pues los conceptos complejos los son en virtud de estar constituidos a partir de primitivos, en última instancia. Y, si no hay conceptos, entonces no hay teoría representacional de la mente.

2.4. Contenido y forma

Una solución que plantea el propio Fodor consistiría en asumir que los primitivos son conceptos de apariencia como ROJO, cuya relación con lo que los causa parece ser arbitraria, en cuanto lo que causa ROJO es el ser bombardeado por fotones, y tanto ROJO como FOTÓN no son similares en cuanto a contenido. De esta forma, los conceptos primitivos no son aprendidos inductivamente, y pueden ser innatos. Sin embargo, para Fodor esta solución no funciona, pues la relación entre la causa y el contenido del concepto deja de ser arbitraria cuando la primera se describe intencionalmente, esto es, cuando se considera que lo que causa que adquiramos ROJO son encuentros con cosas que experimentamos *como rojas*, sean fotones u otra cosa. Así, apelar al *sensorium* por ejemplo para salir del problema, en el sentido de postular mecanismos innatos que arbitrariamente conectan ciertas experiencias

sensoriales con ciertos contenidos conceptuales, no es de utilidad si se toman los estímulos para esa adquisición bajo la descripción intencional.

Sin embargo, el rechazo de Fodor a esta solución sólo se comprende por su teoría semántica: el contenido de un concepto está determinado por una relación nomológica entre la propiedad del mundo y la representación que refiere a esa propiedad. Fodor requiere la postulación de propiedades irreduciblemente intencionales. En este caso la propiedad de *ser rojo* tendría que individuarse como aquella propiedad que tienen ciertas cosas en virtud de la cual mentes como las nuestras instancian el concepto ROJO.

Creo que es posible salvar el externismo sin hacer tales concesiones. Parece perfectamente plausible que exista algo acerca de nuestras mentes que sea el responsable de que formemos/adquiramos el concepto ROJO en virtud de encuentros con instancias prototípicas de lo que a un individuo normal, en un entorno como el nuestro, se le aparece como rojo. En este caso habría que apelar a mecanismos causales subyacentes a la adquisición que no sean ellos mismos conceptuales, y el aparato sensorio conectado con la mente cognitiva/conceptual puede perfectamente jugar este papel⁸⁵. Estos mecanismos claramente podrían cumplir la condición de ser lo suficientemente confiables como para admitir la posibilidad del error, aunque no de forma sistemática. Aún más, el funcionamiento de tales mecanismos subyacentes

⁸⁵ Otras formas de concebir la relación, de manera no-arbitraria, entre contenido y forma para los conceptos primitivos serán presentadas al final del capítulo 4.

puede afectar la forma de la representación, determinando de ese modo el papel causal del concepto adquirido⁸⁶.

La objeción obvia a esta postura es que parece ser que cualquier mecanismo que medie tal adquisición debe ser considerado en la individuación del contenido conceptual. En las teorías que Fodor está enfrentando, a saber las del papel inferencial, se identifica el contenido conceptual con el papel que juega el concepto en el pensamiento, identificando este papel ya sea con disposiciones a reaccionar ante un estímulo realizando ciertos juicios de categorización, a realizar ciertas inferencias, etc., o bien con las inferencias o computaciones en las que de hecho participa. En estas teorías tales disposiciones y/o inferencias se conciben tanto como el papel causal del concepto así como su contenido⁸⁷. De hecho, esta es otra manera de describir el proceso inductivista de aprendizaje de conceptos, sólo que en su versión de aplicación de los conceptos⁸⁸. Por lo mismo, si esto es lo que se tiene en mente, entonces no se ha logrado una versión que combine el ser primitivo con el ser adquirido para una representación.

Sin embargo, me parece que en la objeción mencionada se pasa por alto la distinción que se hizo al comienzo del presente capítulo. Una cosa es que el contenido y la forma de un concepto primitivo no dependa de la posesión de otros conceptos, y otra muy distinta es que un concepto primitivo no tenga

⁸⁶ Podría objetarse que sólo estoy desplazando el problema del innatismo desde los conceptos hacia los mecanismos subyacentes. En tanto tales mecanismos no sean descritos al nivel intencional/conceptual, no veo el problema en aceptar mecanismos innatos que medien la adquisición de conceptos primitivos. Estoy dispuesto a aceptar el innatismo de nuestro mecanismo de visión, siempre y cuando los conceptos que puedo llegar a adquirir a través del funcionamiento de ese mecanismo no sean ellos mismos innatos.

⁸⁷ La exposición y análisis de estas teorías es el contenido del capítulo 3.

⁸⁸ Como se verá en el capítulo siguiente respecto de la categorización, esto es justamente lo que tienen en mente los psicólogos cognitivos cuando hablan del contenido conceptual.

partes constitutivas. De acuerdo a la hipótesis de símbolos físicos es la forma la responsable de los procesos en los que participa un concepto/representación, sin alterar su contenido. La forma de un concepto puede caracterizarse funcionalmente de acuerdo al papel causal de tal concepto, sin que por esto el contenido del concepto se vea afectado. De ahí que la inferencia que se hace desde el papel causal al papel inferencial no esté autojustificada, sino que más bien queda excluida por razones de la arquitectura asumida.

Por otro lado, el contenido de un concepto primitivo también puede estar mediado por otras representaciones, siempre y cuando esas representaciones no sean ellas mismas conceptuales. Esto es, pueden haber más componentes con contenido en el sistema informacional además de los conceptos. Un mal funcionamiento puede representar un problema en el flujo de información, incluso la falta de información puede representar información acerca del funcionamiento de un mecanismo.

Por ejemplo, Carey (2009) caracteriza las representaciones conceptuales como aquellas cuyo formato probablemente sea amodal, y cuyo papel en el pensamiento consista de un número mucho más rico de computaciones o inferencias respecto de las representaciones sensoriales o perceptuales. Por otro lado, para Evans (1982), el contenido conceptual es aquel que está disponible para ser utilizado en distintas inferencias (procesos computacionales, pensamiento, en el ámbito de esta tesis), por ejemplo al recordar, al planear acciones futuras, al categorizar un objeto inmediatamente

presente y dado a la percepción, etc., de forma tal que el concepto no se ve alterado por su participación en tales procesos. El Constreñimiento de Generalidad (Evans, G. 1982) nos dice que si alguien es capaz de pensar acerca de un objeto a que tiene la propiedad F, y de otro objeto b que tienen la propiedad de ser G, entonces el individuo tiene todo lo necesario para pensar que a es G y que b es F. El contenido conceptual en este caso tiene la propiedad de ser el mismo contenido en todas las combinaciones donde aparezca. Es en este sentido que digo que, de acuerdo a la caracterización de Evans, el contenido conceptual está disponible para ser utilizado en todas las inferencias que ocurren a nivel del procesamiento de información ligado a la posesión de conceptos⁸⁹. Creo que esto es otra forma de poner de manifiesto hasta qué punto el principio de composicionalidad, tal como se presentó en el capítulo 1, está ligado a la forma de entender los conceptos y su procesamiento.

Así, el sistema informacional puede poseer variadas representaciones, pero sólo algunas cumplirán con el requisito de ser propiamente conceptuales. Por ejemplo, puedo tener información visual representada de alguna forma en mi sistema informacional, sin que sea capaz de ocupar esa información para otros procesos cognitivos, tal vez debido a que no soy capaz de categorizar/conceptualizar tal información. Es decir, hay información/contenido y hay representación, pero ésta no alcanza el nivel conceptual requerido por el

⁸⁹ Si bien el texto de Evans trata acerca de términos singulares y aunque rechaza la existencia de los símbolos físicos a favor de una ontología de disposiciones respecto de lo mental, su Constreñimiento de Generalidad para los conceptos se aplica perfectamente en el marco de investigación aquí asumido. De hecho, tal constreñimiento puede verse como otra forma de presentar el principio de composicionalidad expuesto en el capítulo 1. Además, la idea de un lenguaje del pensamiento, para las actividades cognitivas superiores al menos, se encuentra en línea con el constreñimiento de generalidad, en tanto sólo un formato amodal de representación parecería ser capaz de cumplirlo en cuanto a que la misma representación estuviera disponible para diversos procesos inferenciales.

Constreñimiento de Generalidad. De acuerdo a esto, es posible pensar que hayan representaciones no-conceptuales que medien la adquisición de conceptos, tanto en su contenido como en su forma, sin que esto signifique que los conceptos primitivos sean innatos, o que no existan en tanto primitivos.

2.5. Conclusión

En este capítulo he caracterizado una forma de concebir al menos algunos conceptos primitivos que no responde a las teorías estándar. En particular, he llegado a la noción de concepto primitivo no-innato. Los motivos para tal caracterización responden a las restricciones alcanzadas en el capítulo 1:

- a) En cuanto a mantener una semántica externista, la adquisición de los conceptos primitivos permite asegurar la conexión entre el mundo y la mente, salvando así el carácter intencional de la última;
- b) En cuanto a satisfacer el principio de composicionalidad, puesto que la distinción entre lo conceptual y lo no-conceptual permite sostener una noción de contenido robusta para los conceptos, sin recurrir a procesos inferenciales o disposiciones; y

- c) En cuanto a hacer sostenible una teoría representacional de la mente frente al argumento que se sigue de la relación evidencial que parece sustentar cualquier forma de adquisición conceptual que no esté mediada por la posesión de conceptos previos.

Ahora bien, la caracterización del contenido conceptual de los primitivos, y por tanto por composicionalidad de cualquier concepto, debe enfrentar serias objeciones cuando se ve enfrentada a cierta evidencia. Específicamente, los casos de Frege relativos a la importancia cognitiva y la evidencia empírica aportada por las tareas de categorización imponen restricciones sobre la teoría de fondo en general, y sobre la individuación conceptual en particular. En los siguientes dos capítulos expondré estas objeciones y la forma en que la teoría hasta ahora expuesta es capaz de lidiar con ellas.

CAPÍTULO 3

Problemas: Contenido Cognitivo y Categorización

En el capítulo anterior hemos visto cómo es posible dar cuenta de distintos aspectos en los conceptos primitivos, en términos de contenido y forma, y por ende de los conceptos complejos, sin tener que renunciar a una semántica externista para el contenido intencional de las representaciones mentales. La necesidad de encontrar distintos elementos en los conceptos radica en que la pura referencia parece ser insuficiente para explicar las capacidades cognitivas asociadas a su posesión. La idea es que la forma o el vehículo del contenido podría proporcionar tal estructura para el caso de los primitivos. Para el caso de las representaciones complejas debiera valer una teoría definicionista de las mismas, según la cual el contenido de tales representaciones se hereda del contenido de sus constituyentes de forma estricta, esto es que los constituyentes son condiciones individualmente necesarias y conjuntamente suficientes para la posesión del concepto⁹⁰. Así, si el contenido de los primitivos se individua de forma referencial, y de acuerdo al principio de composicionalidad que esta teoría respeta, entonces el contenido de los conceptos complejos también debiera ser individuado de manera externista.

Sin embargo, una vez que se introduce la forma en la individuación conceptual, resulta tentador querer ver en ella un aspecto del contenido conceptual. En cuanto a la posesión de conceptos parece ser la mejor guía

⁹⁰ Como se verá en lo que sigue del capítulo, esta teoría se distingue de aquellas inspiradas en las tareas de categorización en la forma de constitución de las representaciones complejas. Las teorías psicológicas tienden a ver tal constitución como un asunto de grados de similitud que pueden rastrearse desde los constituyentes hacia los conceptos complejos, donde ningún constituyente en último término constituye una condición necesaria para la posesión del concepto complejo. Por lo tanto el contenido de una representación compleja podría no heredar el contenido de uno de sus constituyentes, violando el principio de composicionalidad tal como se ha presentado.

para la individuación de tal contenido, y en tanto la introducción de la forma del concepto sirve para dar cuenta de cómo es que usamos los mismos, parecería natural asumir que estamos hablando del contenido conceptual, y no de la forma. En particular, hay dos condiciones asociadas a la posesión de conceptos que motivan incluir dentro del contenido conceptual aquello que juegue el papel que hasta aquí se le ha asignado a la forma, a saber, las consideraciones inspiradas en Frege respecto del contenido cognitivo, por una parte, y por otra consideraciones basadas en los trabajos en psicología cognitiva relativos a las tareas de categorización. Así como Frege vio la necesidad de introducir una categoría semántica adicional a la de la referencia, esto es el sentido fregeano, distintas aproximaciones al contenido conceptual intentan rescatar esta noción en términos del papel conceptual o inferencial que es exhibido en el uso de los conceptos⁹¹, tanto en casos paralelos a los que motivan a Frege como en las tareas psicológicas de categorización en las que se revelarían rasgos constitutivos de los conceptos tal como los usamos y aprendemos⁹².

⁹¹ Claramente identificar el sentido fregeano con el papel conceptual no es la única opción (Cf. Ezcurdia 1995), sin embargo, las teorías aquí revisadas asumen tal identidad.

⁹² Claramente aquí las teorías son acerca de los conceptos léxicos y se asume su carácter estructurado en términos de contenido. Dada la forma de concebir la estructura por parte de estas teorías, tal caracterización no está disponible para los conceptos primitivos (si los hubiera). Cómo mi tesis parte de una caracterización de los conceptos primitivos y la caracterización del contenido, más el principio de composicionalidad tal como fue esbozado en el capítulo 1, no me pronunciaré acerca de la disputa sobre los conceptos léxicos. Más bien, lo que interesa en este capítulo es mostrar que la forma de concebir la semántica basada en el papel inferencial es incompatible con la caracterización de los conceptos complejos en términos del principio de composicionalidad. Por lo mismo, y dado que la semántica externista y el principio de composicionalidad obligan a una teoría definicionista de los conceptos complejos, en este capítulo sólo me centraré en los problemas que presentan a este respecto las teorías de conceptos que asumen que la estructura del contenido de los conceptos complejos se establece en términos no de relaciones definicionales sino de relaciones estadísticas, a saber las teorías de los prototipos y la de los estereotipos. Del mismo modo, en cuanto la teoría-teoría también pareciera rechazar la teoría definicionista para los conceptos complejos, mostraré que tal aproximación es insostenible dentro del marco supuesto en la investigación.

Además, las teorías de conceptos, a excepción de la propuesta de Fodor⁹³, suelen asumir que el contenido caracterizado en términos puramente externistas no es suficiente para dar cuenta de algunas capacidades cognitivas comúnmente asociadas a la posesión de conceptos. Brevemente, la pura referencia no permite hacer distinciones más finas al nivel de los procesos cognitivos que involucran la posesión de conceptos, por lo que se requiere enriquecer el contenido de forma tal que sea posible explicar tales fenómenos.

Esta objeción vale para una teoría que asuma que los conceptos léxicos son atómicos. Una teoría como la definicionista no se enfrenta a este problema. Sin embargo, como hemos visto toda teoría de las representaciones mentales debiera dar cuenta de los conceptos primitivos a partir de los cuales se estructuran los complejos. En este sentido las consideraciones que se siguen del contenido cognitivo ligadas a las tareas de categorización ponen en cuestión no sólo al atomismo sino que principalmente a la teoría definicionista.

A continuación expondré los dos problemas que hacen surgir la necesidad de un aspecto del contenido conceptual distinto al puramente referencial, a saber el del contenido cognitivo y el de la categorización⁹⁴. Luego presentaré las caracterizaciones del contenido conceptual basadas en lo que revelan los experimentos de categorización. Finalmente, mostraré que tal caracterización, basada en el papel inferencial asociado a la posesión de

⁹³ Si bien Fodor reconoce la insuficiencia de la pura referencia para distinguir conceptos correferenciales distintos, intenta explicar tal diferencia en términos de la forma o papel causal de los conceptos, sin introducir una nueva categoría semántica. En el capítulo siguiente intentaré dar cuerpo a esta postura, pues me parece que tal y como la formula Fodor resulta insuficiente para enfrentar las objeciones de Prinz (2002).

⁹⁴ En esta sección seguiré la presentación que hace Prinz (2002) con relación a los *desiderata* que debe satisfacer una teoría de conceptos.

conceptos, no satisface la noción de composicionalidad esbozada en el capítulo 1. En el capítulo siguiente intentaré mostrar que es posible dar cuenta de los problemas aquí presentados sin tener que renunciar a la noción de contenido ya alcanzada en el capítulo precedente, basándose en una caracterización más detallada de lo que hasta aquí he denominado la forma del concepto.

3.1. Contenido Cognitivo

Además del contenido referencial, caracterizado en términos puramente externistas, parece haber otro aspecto del contenido conceptual que debe jugar algún rol en su individuación. Específicamente se ha pensado que la posesión de conceptos es lo que debiera explicar ciertas características cognitivas. Así por ejemplo, Prinz (2002:6) sostiene que hay razones bien conocidas por las cuales sería sensato pensar que los conceptos no pueden ser individuados en términos puramente referenciales o del contenido intencional. En particular se alega que hay dos argumentos derivados de la filosofía del lenguaje de Frege⁹⁵. Los dos argumentos a los que alude Prinz corresponden a la informatividad de juicios verdaderos de identidad, por una parte, y por otra a la falla en la substitución de términos correferenciales en contextos de atribución de actitudes proposicionales. En ambos casos la igualdad de contenido

⁹⁵ Prinz denomina “contenido intencional” a lo que hasta aquí he llamado “contenido referencial”, privilegiando un uso de “intencional” donde lo relevante es la relación entre la representación y lo representado, o la relación mente-mundo en el caso de los conceptos. Este uso de Prinz de “contenido intencional” le sirve para distinguirlo del contenido atribuido al papel inferencial que denomina “contenido cognitivo”.

referencial de los términos involucrados no resulta suficiente para sostener la igualdad de los términos, en cuanto a su contenido. Más bien lo que sugieren estos casos es que los términos/conceptos involucrados deben distinguirse, y como el contenido referencial no puede realizar esa labor, surge la necesidad de postular otro tipo de contenido que dé cuenta de la diferencia observada.

Según Prinz, argumentos paralelos a los de Frege pueden construirse sin mencionar el lenguaje, y aplicarse a la noción de contenido referencial para las representaciones mentales. De acuerdo a esta postura, así como Frege postula la necesidad de otra categoría semántica además de la referencia, a saber la del sentido, respecto del contenido conceptual necesitamos además del contenido intencional lo que se podría llamar contenido cognitivo. En una primera aproximación, el contenido cognitivo es lo que permite que dos representaciones correferenciales parezcan semánticamente distintas a un agente cognitivo. Tal como lo presentó el propio Frege, la importancia cognitiva de dos oraciones será distinta si es posible que el mismo individuo tome actitudes epistémicas distintas hacia ellas, a pesar de que su referencia (contenido intencional) sea la misma. Por ejemplo, es perfectamente posible que alguien crea que es verdadera la oración "Cicerón fue un gran orador romano" al mismo tiempo que crea falsa la oración "Tulio fue un gran orador romano", a pesar de que ambas oraciones expresan la misma proposición o tienen las mismas condiciones de verdad, esto es, tienen el mismo contenido intencional. Del mismo modo, en el caso de que dos conceptos no correferenciales le parezcan a un agente cognitivo como semánticamente similares, ha de ser el contenido cognitivo de tales representaciones lo que

explique este fenómeno. En este caso se sugiere que un agente racional puede asociar la misma información con dos términos cuyo contenido difiere, y por tanto comportarse de la misma forma ante su presencia, pese a la diferencia en contenido intencional. Un ejemplo típico es el del uso de demostrativos, como en el caso de dos individuos que tienen la creencia que sería expresada por las emisiones de “Yo estoy en peligro”, si bien el contenido intencional es distinto para ambos individuos, no es menos cierto que hay algo en común a ambos que explica su supuesta conducta similar ante la creencia. Otro ejemplo ya clásico es el de Oscar y su gemelo de la otra tierra y el concepto/término “agua” (Putnam, 1975). En este caso parece posible describir a ambos individuos como similares en cuanto a su conducta dadas sus creencias acerca del agua, expresadas en oraciones que contienen la emisión del término “agua”, pero que desde la perspectiva del contenido referencial tales similitudes resultan irrelevantes a la hora de individuar sus estados mentales.

Es importante notar aquí que Prinz insiste en que lo relevante debe ser considerado como semántico, pues al depender la diferencia de actitudes epistémicas que puedan tomar los sujetos, éstas deben basarse en última instancia en el contenido de las representaciones. Las motivaciones para considerar que la forma no es capaz de cumplir ese papel provienen básicamente de dos tipos de objeciones:

- a) Si consideramos el paralelo entre el lenguaje natural y el lenguaje del pensamiento seriamente, entonces parece claro que en el primer caso la forma de las palabras u oraciones no es suficiente para dar cuenta del

contenido cognitivo. Por ejemplo, que dos palabras difieran en forma no elimina la posibilidad que tengan el mismo contenido cognitivo: considérese el caso de “water” y “agua” para un hablante competente bilingüe. Por otro lado, oraciones o palabras con la misma forma pueden tener distintos contenidos cognitivos, como lo ilustran los casos de Kripke (“Paderewski”)⁹⁶.

- b) Por otra parte, si uno no admite el paralelo entre el lenguaje natural y las expresiones del lenguaje del pensamiento, a fin de evitar la objeción anterior, entonces no parece haber una caracterización clara de las propiedades formales de las representaciones mentales. Según Prinz, apelar a la individuación material del concepto, en cuanto símbolo físico, no es sino postergar el problema o renunciar a la posibilidad de explicarlo. Además, si uno se queda con la pura la individuación funcional de las representaciones mentales, como lo hace Fodor al decir que lo que sea que hace que dos conceptos correferenciales sean distintos es su forma, entonces esto nos conduce a decir que sus propiedades formales deben asimilarse a sus papeles causales. Sin embargo, al nivel de la individuación funcional los papeles causales suelen identificarse con los papeles inferenciales que cumple la representación mental en sus interacciones cognitivas con otras

⁹⁶ En estos casos la intuición es que es perfectamente posible pensar que Paderewski no es Paderewski si las circunstancias son tales que uno asocia el nombre con dos individuos distintos, por ejemplo porque tuvo noticia del individuo en dos situaciones distintas y no hay nada que lo haga pensar que se le hablaba de la misma persona. Aquí sólo presento el problema tal y como lo esgrime Prinz. En el siguiente capítulo me ocuparé de mostrar como mi propuesta es capaz de lidiar con ellos.

representaciones mentales, y de acuerdo a Prinz esto finalmente nos lleva a considerar tales papeles como esencialmente semánticos⁹⁷.

Finalmente, lo que parece seguirse de las objeciones antes mencionadas, junto con el requerimiento de dar cuenta del contenido cognitivo, es que tal contenido no puede equipararse con la forma. Las diferencias o similitudes en contenido cognitivo parecen responder a diferencias en capacidades o actitudes epistémicas, tal como lo señaló Frege, y por lo tanto deben ser tratadas a un nivel semántico. Tal como Prinz destaca:

“Si las diferencias y similitudes epistémicos juegan un papel central en la verificación de nuestras atribuciones de contenido cognitivo y las diferencias formales no, entonces los contenidos cognitivos deben ser epistémicos en su naturaleza. Estas diferencias deben derivar de las características semánticamente interpretables de los conceptos, en lugar de las propiedades puramente formales de los conceptos.” (2002, 96)⁹⁸.

⁹⁷ En tanto este contenido cognitivo es el que explica las inferencias que aquel que posea el concepto es capaz de realizar, Prinz considera que debe ser semántico. Aquí sólo sigo el argumento de Prinz, a fin de poder responderlo en el capítulo siguiente.

⁹⁸ Tal como lo dije en la nota anterior, la conclusión a la que llega Prinz no es obvia. Primero, Prinz se olvida de que la relación entre pensamiento y lenguaje ha sido declarada neutral por él mismo. Que ciertas diferencias y/o similitudes epistémicas estén involucradas en nuestra práctica de atribuir y verificar actitudes proposicionales no basta para concluir nada acerca del contenido mental. Nos podrían decir algo acerca de la semántica de tales prácticas lingüísticas, pero no necesariamente del contenido de los conceptos. Además, del hecho de que tales capacidades sean epistémicas no se sigue, al menos necesariamente, que deban derivarse de rasgos semánticos. En el siguiente capítulo desarrollo un argumento para responder a esto. Además, el reciente trabajo de Machery (2009) también muestra que gran parte del trabajo desarrollado por los psicólogos cognitivos, basado en las tareas de categorización, puede ser interpretado como no siendo acerca de los conceptos, sino más bien acerca de los procesos que la mente puede hacer en virtud de poseerlos. Sobre el trabajo de Machery hablaré en la conclusión.

Aún más, justamente esta noción de contenido cognitivo es la que juega el papel preponderante en una forma de entender la categorización, de orden psicológico, y que debiera contar como un constreñimiento para una teoría de los conceptos (Prinz, 2002; Rosch, 1978), de la que me ocupo en el apartado siguiente.

3.2. Categorización

“La referencia es una relación semántica, pero también tiene una contrapartida epistémica. Además de la serie de cosas a las que un determinado concepto refiere, hay también un conjunto de cosas a las que el concepto se *piensa que refiere*. Tenemos mecanismos para la formación de creencias acerca de qué cosas caen bajo nuestros conceptos, los mecanismos de categorización. Los conceptos son a menudo identificados con esos mecanismos.” (Prinz, 2002: 9)

Como se verá en este apartado, gran parte de la discusión respecto del contenido conceptual surge a raíz de esta otra forma de entender qué son los conceptos, ligada al papel inferencial. En particular, el debate gira en torno a la noción de contenido, y cuál de las dos opciones, o si las dos, debe ser considerada como el candidato a ocupar ese papel.

Las tareas de categorización han guiado la investigación sobre conceptos en psicología cognitiva. En tales tareas parece revelarse el cómo, de

hecho, utilizamos los conceptos que poseemos, y por ende nos dicen o son evidencia acerca de, cómo son los conceptos. Hay al menos dos sentidos de categorización que deben considerarse: la identificación de un objeto como perteneciendo a una categoría por una parte; y por otra la capacidad de producir una categoría, en el sentido de decidir y/o identificar los atributos que deben poseer los miembros de una categoría para pertenecer a ella, y ser capaces de hacer inferencias desde la pertenencia a una categoría por parte de un objeto a los atributos que debiera o podría poseer tal objeto.

En cuanto a la producción de categorías, una forma de poner a prueba esta capacidad consiste en exponer a los participantes a una serie de objetos no conocidos, señalarles que todos ellos pertenecen a una misma categoría, y luego presentarles una serie de objetos nuevos a fin de que decidan cuáles pertenecen a la categoría recién formada. Esto proporcionaría evidencia acerca de qué factores, rasgos o aspectos de la muestra inicial son considerados relevantes para la pertenencia a la categoría, y por tanto cuál es la forma en que producimos la categoría.

Word	Concept	Pack I	Pack II	Pack III	Pack IV	Pack V	Pack VI	Pack VII	Pack VIII	Pack IX	Pack X	Pack XI	Pack XII
Series A	oo	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎
Series B	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎	𠄎

Figura 1. Hull (1920)

Uno de los primeros experimentos relativos a la categorización fue realizado por Hull en 1920. En primer lugar se les presentan a los individuos dos objetos desconocidos para ellos, se les informa que pertenecen a categorías distintas y finalmente tienen que decidir sobre nuevos objetos respecto de su pertenencia a una de las dos categorías iniciales. En la Figura 1 aparecen los resultados de categorización sobre dos objetos desconocidos, esto es las dos categorías, bajo la columna "*concept*" asociados a palabras "oo" y "yer".

En este experimento pareció demostrarse que la formación de conceptos seguía el patrón que predecía la teoría definicionista. Los sujetos tomaron algunos rasgos de la categoría como condiciones necesarias y suficientes para la pertenencia, y así decidieron sobre los casos nuevos.

Sin embargo, nuevos experimentos mostraron que no siempre los individuos siguen esta estrategia. Aun cuando no hayan rasgos definitorios de la pertenencia a una categoría, los individuos sometidos a las tareas de categorización son capaces de responder a las pruebas, por ejemplo basándose en patrones de similitud.

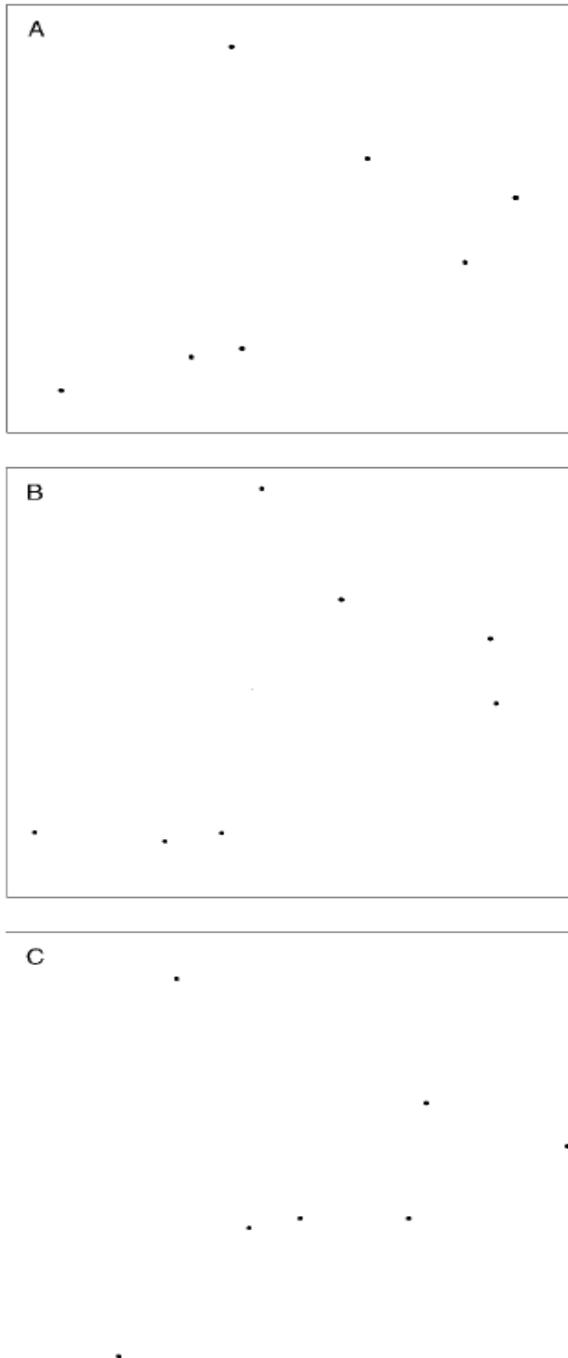


Figura 2. Posner y Keele (1968, 1970)

En la figura 2 se muestran patrones de puntos distribuidos al azar. El cuadro A representa el estado inicial. Luego de mostrarles a los participantes varios cuadros similares con pequeñas distorsiones respecto de la configuración del patrón de puntos, se les pide que categoricen B y C, en

cuanto a si son del mismo tipo o categoría que A. En este caso los individuos respondieron a la tarea, señalando que el patrón de B es del mismo tipo que A, a diferencia del patrón C que no fue categorizado como siendo del mismo tipo. En este caso los sujetos no consideraron rasgos necesarios y conjuntamente suficientes, pues cualquiera sea el rasgo que se tome aparece en ambos patrones (B y C). La medida que pareció guiar a los individuos fue más bien el grado de similitud que puede observarse entre los patrones de puntos.

Frente a estos resultados la teoría definicionista no parece tener una explicación satisfactoria. Si las representaciones que sirven para realizar la categorización están constituidas de elementos que cuentan como condiciones individualmente necesarias y conjuntamente suficientes, entonces la pertenencia a una categoría debiera ser una cuestión de "sí o no". Sin embargo, si es posible realizar tales tareas apelando a similitudes, entonces tal pertenencia pasa a ser una cuestión de grado, y los límites de la categoría misma se vuelven difusos.

Finalmente, los trabajos en psicología relativos a la categorización han mostrado efectos, llamados de *tipicidad*, que son contrarios a las predicciones de la teoría definicionista. Así, en tales experimentos se han encontrado casos dudosos, donde el individuo no está seguro si el objeto cae o no bajo el concepto, también el hecho de que ciertos objetos parecen ser mejores instancias del concepto que otros, en términos de la psicología hay casos más típicos, menos típicos y casos atípicos de la categoría, entre otros de estos efectos de tipicidad. Un ejemplo tomado de Rosch (1973), figura 3, muestra

que al ser consultados los participantes sobre cuán típica es una subcategoría de una categoría mayor, en este caso tipos de frutas sobre la categoría más general “fruta”, se observa que los individuos son capaces de responder la pregunta, lo que parecería imposible si ellos tuvieran codificada una definición de rasgos separadamente necesarios y conjuntamente suficientes para la pertenencia a la categoría, y aún más las respuestas son consistentes entre los participantes. En la misma línea de investigación, se observó que la rapidez con que un objeto es categorizado depende de la presencia de ciertos rasgos que parecen ser distintivos de la misma. Los rasgos no son constitutivos pues un objeto puede carecer de dicho rasgo y aun así ser categorizado como perteneciendo a la categoría, lo que afecta la posesión o ausencia de tales rasgos es la rapidez con que se efectúa el juicio de categorización. En la figura 4 se da una muestra de los rasgos prototípicos de la categoría pájaro, lo que predice, acertadamente, que los petirrojos serán categorizados como pájaros más rápido que los pollos.

Fruit	Typicality rating on a scale of 1–7 (with 1 being highest)
Apple	1.3
Plum	2.3
Pineapple	2.3
Strawberry	2.3
Fig	4.7
Olive	6.2

Figura 3. Rosch (1973)

Feature	Bird	Robin	Chicken
Flies	yes	yes	no
Sings	yes	yes	no
Lays eggs	yes	yes	yes
Is small	yes	yes	no
Nests in trees	yes	yes	no
Eats insects	yes	yes	no

Figura 4. Rosch y Mervis (1975).

Así, las teorías de los conceptos inspiradas en las tareas psicológicas de categorización surgen como una alternativa a la teoría definicionista de los conceptos léxicos. Además, tomando en cuenta las objeciones de Wittgenstein al análisis conceptual predominante en filosofía en su época, recogen la noción de “parecido de familia” como aquella que caracterizaría adecuadamente la relación inferencial entre conceptos, en cuanto a ser una relación que recoge los rasgos estadísticamente relevantes para la pertenencia a una categoría, y por tanto que revela la estructuración de los conceptos léxicos de una forma radicalmente distinta a lo que la teoría definicionista propone, esto es, en términos de condiciones individualmente necesarias y conjuntamente suficientes.

3.2.1. Teorías basadas en la categorización

Como respuesta a los efectos de tipicidad, los psicólogos cognitivos han planteado sus propias teorías. Principalmente se pueden reconocer tres teorías

dentro de esta familia, a saber, la teoría de prototipos, la de ejemplares y la teoría-teoría. Además, la propia teoría de Prinz toma estos desarrollos para plantear una alternativa tanto a la teoría definicionista como al atomismo de Fodor⁹⁹.

La idea sobre cómo se individúan los conceptos en la teoría de prototipos nos remite a lo que puede revelarse de su estructura en los experimentos de categorización. En este sentido parece plausible asumir que la evidencia sobre la posesión de conceptos que nos entregan tales experimentos resulta una buena guía para su individuación. Por un lado, tiene sentido atribuir un concepto a un sujeto con base en lo que él mismo demuestra saber acerca de ese concepto, por ejemplo cuando se le pide que diga cuáles son las características que definen o son relevantes para pertenecer a la categoría¹⁰⁰. Por otro, en otros experimentos no se le pida al sujeto que diga nada acerca de lo que él cree, sino que más bien se ve cómo utiliza el concepto en tareas de categorización, por lo que esta manera de individuarlos parece psicológicamente plausible¹⁰¹. Así, si queremos saber cómo es el concepto de un individuo cualquiera podemos recurrir a los estudios de categorización realizados en su población, o una población de muestra lo suficientemente

⁹⁹ Aun cuando la teoría de los prototipos de Prinz supone una arquitectura cognitiva distinta a la que sirve de base a esta investigación, la considero aquí como otro exponente de teorías que toman los datos de la categorización como relevantes para la individuación conceptual. Además, las críticas basadas en el principio de composicionalidad (3.3) también se aplican a la teoría de los prototipos.

¹⁰⁰ Podría objetarse que una de las principales ventajas de la ciencia cognitiva sobre la psicología especulativa es que no depende de la introspección como método privilegiado de acceso a lo mental. Además, también se podría apelar a consideraciones pragmáticas por las cuales los sujetos no nombran condiciones necesarias, por ejemplo la de “ser vivo” para pájaro.

¹⁰¹ Uno podría distinguir a nivel cognitivo entre mecanismos de identificación de la categoría y el contenido de la representación mental. Los primeros podrían ser pensados como capacidades epistémicas que de hecho dependen de la posesión de los conceptos, pero no como parte del concepto. Parte de la respuesta referencialista que ofrezco en el siguiente capítulo abordará este problema, así como consideraciones generales acerca de la relación entre estas capacidades epistémicas y los conceptos que incluyo en la conclusión de la tesis.

cercana al individuo, y predecir con una alta probabilidad cuál será su concepto e incluso qué tipo de acciones tomará dadas ciertas circunstancias.

La teoría de prototipos fue la primera teoría de conceptos de inspiración psicológica basada en las tareas de categorización. En ella se propone la representación mental como un resumen o lista con los rasgos estadísticamente relevantes para la pertenencia a la categoría. Además los rasgos en cuestión son independientes unos de otros. Por su parte, la teoría de los ejemplares, en competencia con la de prototipos, señala que más bien es un conjunto de „memorias’ de instancias de la categoría las que están guardadas en la representación mental, cada una como un objeto coherente y sin distinciones entre rasgos relevantes e irrelevantes.

Recurriendo a la caracterización de Murphy (2002) de ambas teorías, si uno ha aprendido el concepto PÁJARO a partir de dos instancias, digamos una paloma caminando sobre el pasto y un gorrión volando, entonces al encontrarse con una paloma volando las teorías harían predicciones distintas: según la teoría de prototipos no sería difícil clasificar la nueva instancia como perteneciendo a la categoría pájaro, en tanto comparte los rasgos relevantes con las dos instancias de la etapa de aprendizaje, pero la teoría de ejemplares diría que tal clasificación no sería tan fácil, pues la instancia nueva no es similar a ninguna de las dos antes aprendidas.

Respecto de la categorización en particular hay diferencias entre ambas teorías, específicamente en lo que Murphy (2002) llama los “efectos del

ejemplar". En ciertos experimentos de categorización se ha mostrado la influencia que tienen las instancias ya aprendidas (entendidas como objetos coherentes, lo cual implica tanto rasgos relevantes como irrelevantes) al categorizar nuevas instancias. Hay evidencia de que la gente compara nuevas instancias con ejemplos anteriores antes que separar los rasgos relevantes y hacer un chequeo de cada uno de ellos en la nueva instancia.

Obviamente, respecto de los "efectos del ejemplar" la teoría de ejemplares resulta explicativa y predictivamente superior a la de prototipos. Sin embargo, las categorías utilizadas en tales experimentos distan mucho de las categorías naturales, en cuanto a la pobreza de su estructura, la cantidad de instancias y la dificultad en su aprendizaje, por lo que sus resultados son al menos cuestionables.

La teoría-teoría surge más bien como una respuesta a las teorías del prototipo y de los ejemplares, aunque dentro de la misma tradición psicológica que basa sus especulaciones acerca de la estructura conceptual en lo que revelan los experimentos de categorización. Respecto de las otras dos teorías basadas en la evidencia que arrojan los experimentos de categorización, la teoría-teoría se distancia de ellas en cuanto considera que no es el parecido de familia o la similitud con otras instancias de la categoría lo que es determinante para la categorización de nuevas instancias, al menos no siempre. Lo que la teoría-teoría propone es que más bien la información codificada en la representación mental hace referencia a patrones inferenciales de carácter explicativo que pueden remitir a cualidades funcionales para categorías

artificiales, rasgos inobservables (de tipo esencialista) en el caso de categorías naturales, u otros que no requieren ni se manifiestan en parecidos o similitudes a nivel perceptual. Así, por ejemplo, lo que permite inferir que una nueva instancia como un pingüino de pocos días cae bajo el concepto PINGÜINO no es su similitud con los ejemplares característicos de tal categoría o el que comparta rasgos observables relevantes del prototipo, lo que de hecho no se da, sino el que camine detrás de un pingüino adulto, como la mayoría de las crías lo hacen respecto de sus progenitores. Además de privilegiar patrones inferenciales explicativos en la estructuración del contenido conceptual, la teoría-teoría señala que tales patrones constituyen toda una red de inferencias que en su conjunto constituyen una teoría donde cada concepto se individua por el rol que juega dentro de esa teoría. Es en ese sentido que el conocimiento previo influye sobre la categorización.

Finalmente, está la teoría de los proxitipos de J. Prinz. A diferencia de la teoría de prototipos, que asume que la representación mental de la categoría, es decir el concepto, está constituida por una lista de rasgos, cada uno de los cuales es a su vez otro concepto, para Prinz los proxitipos están constituidos por un conjunto de perceptos traídos desde la memoria en relación directa con la tarea para la cual son requeridos. Una primera diferencia entre estas teorías está en el carácter temporal de los proxitipos, pues éstos no corresponden a estructuras fijas y disponibles almacenadas en la mente. Por el contrario, lo que está almacenado son los perceptos y es un conjunto de ellos los que se instancian en cada caso según la tarea cognitiva. Otra diferencia es que los proxitipos admiten todo tipo de información en sus constituyentes, esto

es, rasgos irrelevantes para la pertenencia a la categoría pero que son contextualmente buenos indicadores, así como rasgos de naturaleza teórica en algunos casos. Como los perceptos se forman a partir de situaciones de experiencia concreta, y no hay un proceso donde tal información es convertida a un código amodal, todo lo que incluye la situación en la categorización queda almacenado de forma relacionada. Sólo en la situación donde no hay un contexto, natural o teórico que dirija, en algún sentido, la formación del proxitipo, resultará que los rasgos que aparecen en él serán los que los teóricos de los prototipos seleccionan como los constitutivos de los conceptos.

Además, respecto del contenido de las representaciones, los teóricos de los prototipos asumen que el contenido cognitivo es equivalente al contenido intencional, en los términos del propio Prinz. Por su parte, Prinz asume que una semántica informacional es la responsable del contenido intencional. Esta posibilidad no está disponible para los adherentes a la teoría de los prototipos, en tanto asumen que hay un solo tipo de contenido, y éste es de carácter inferencial.

3.3. Composicionalidad y Teorías basadas en la Categorización

Más allá de las diferencias entre las tres teorías antes presentadas, hay un factor común que comparten en su rechazo de la teoría definicionista en cuanto a la composición del contenido conceptual, a saber, que los

constituyentes de las representaciones complejas no son condiciones separadamente necesarias ni conjuntamente suficientes para la pertenencia a la nueva categoría. Como ya se ha dicho, estas teorías comparten la idea de que sólo se requiere un cierto grado de similitud entre las instancias de la categoría, y que por tanto los rasgos constituyentes son sólo estadísticamente relevantes¹⁰². Es justamente este factor común el que nos lleva a su rechazo como teorías de la representación mental, dentro del marco propuesto en esta investigación, por cuanto no es posible que tal caracterización del contenido sea compatible con la noción de composicionalidad tal como se esbozó en el capítulo 1¹⁰³.

La teoría de prototipos afirma que los conceptos se individúan por su contenido, y que éste a su vez está determinado por las relaciones inferenciales estadísticamente relevantes que tiene con otros conceptos. Esto es, un concepto se individua por los rasgos o características, representados por otros conceptos, que los objetos que caen bajo el concepto tienden a tener. Esta teoría asume que para que un sujeto tenga el concepto debe saber que esas relaciones, aunque sólo con cierta relevancia estadística, se dan.

De acuerdo a lo anterior, el contenido de una representación compleja en la teoría de prototipos debiera heredar las características estadísticamente relevantes de los conceptos que la constituyen, si se respeta el principio de composicionalidad, sin embargo esto no ocurre en el caso de los prototipos. Si

¹⁰² Otra forma de decirlo es que estas teorías rechazan que las inferencias constitutivas de los conceptos complejos sean inferencias analíticas.

¹⁰³ Noción que requiere que las inferencias entre los conceptos constituyentes y el complejo sean de carácter analítico, en el sentido de que el contenido de la representación compleja se hereda de sus constituyentes.

comparamos el prototipo de una categoría¹⁰⁴ compleja con los prototipos de las categorías que la constituyen no resulta que los rasgos prototípicos sean los mismos. Más bien, el prototipo de la categoría compleja exhibe rasgos estadísticamente relevantes para la pertenencia a la misma, rasgos que podrían denominarse emergentes, que no corresponden a las características prototípicas de sus constituyentes. El ejemplo más utilizado en este caso es el del PEZ MASCOTA, cuyo prototipo no tiene las características estadísticamente relevantes ni de MASCOTA ni de PEZ.

Aún más, la composicionalidad permite explicar cómo somos capaces de tener o formar conceptos sobre cosas que nunca hemos visto, o una categoría demasiado amplia, sólo al combinar conceptos, para las cuales carecemos, por razones obvias, de un prototipo. En este caso uno puede pensar en conceptos tales como COSAS QUE PESAN MÁS DE 10 GRAMOS, LÁMPARA O RANA, MONARQUÍA ESTADOUNIDENSE, etc.

Por último, tal como lo señala Murphy (2002), una de las ventajas de la teoría de prototipos, sobre teorías definicionistas, es que explica la intransitividad en los juicios de categorización. En este caso un objeto puede ser categorizado como perteneciendo a una categoría que a su vez es una sub-clase de otra categoría más general, y sin embargo el mismo objeto no es categorizado como perteneciendo a la categoría más general. Dos ejemplos que cita Murphy son relativos al concepto MUEBLE. En el primer caso los participantes están de acuerdo en que los asientos de automóvil son un tipo de

¹⁰⁴ Sigo aquí la usanza de los proponentes de esta teoría en cuanto a no distinguir concepto/categoría.

silla, y que las sillas son un tipo de mueble, sin embargo no aceptan que los asientos de automóvil sean muebles. En el segundo caso los participantes categorizan el Big Ben como un tipo de *clock*¹⁰⁵, y que *clock* es un tipo de mueble (a diferencia de *watch*, un reloj que no es un tipo de mueble), sin embargo no aceptan que el Big Ben sea un mueble. De forma similar, en algunas inferencias la gente es incapaz de hacer ciertas generalizaciones aunque el principio de transitividad o inclusión de clases lo permita y asegure su verdad (Hampton, 1982). Por ejemplo, es mucho más común generalizar a todos los pájaros una característica que posean los pájaros prototípicos, pero no así hacia otro tipo de pájaros más bien atípicos. Así, los participantes infieren que si los petirrojos tienen un tipo especial de sistema coronario, entonces todos los pájaros lo tienen, pero sin embargo a partir de la misma premisa no infieren que los avestruces lo tienen (reconociendo los participantes que los avestruces son pájaros).

Ahora bien, esto último plantea un dilema:

- a) Si se acepta la intransitividad en los juicios de categorización como revelando la estructura del concepto, entonces no es sólo que los prototipos no compongan, sino que la composicionalidad simplemente no se daría para los conceptos, lo que llevaría a cuestionar este principio como requisito para una teoría de conceptos¹⁰⁶. Esta es la única salida

¹⁰⁵ No sé cómo traducir esta forma de decir “reloj”, que en inglés se diferencia de “watch”, también traducido al español como “reloj”, a fin de que se entienda el ejemplo.

¹⁰⁶ Y, según lo que he expuesto, a poner en duda la teoría de fondo supuesta en esta tesis. Como lo señalé en el capítulo 1, creo que es parte del marco hegemónico asumir la composicionalidad entendida de manera combinatorial. Si este fuera el caso, entonces la discusión queda más allá del alcance de la presente investigación.

que veo para la teoría de prototipos frente a la condición de composicionalidad. Por ejemplo, Murphy y Medin (1985) muestran que no todo concepto compuesto es la intersección de los componentes: HORSE RACE, OCEAN DRIVE, EXPERT REPAIR, ENGINE REPAIR¹⁰⁷.

- b) Asumir que los juicios de intransitividad no revelan la estructura de los conceptos. Aquí se salva la composicionalidad, pero la teoría de prototipos no funciona, pues los prototipos no componen: los conceptos componen, los prototipos no, por ende los conceptos no son prototipos.

Si bien las consideraciones hechas a la teoría de prototipos respecto del problema de la composicionalidad son similares a las que podrían hacerse a la teoría de ejemplares, en este caso el problema resulta aún más grave. Si la información asociada a un concepto son „memorias’ de instancias de la categoría, entonces resulta mucho más difícil que representaciones complejas se compongan a partir de las representaciones más simples que la constituyen.

Creo que más bien aquí surge un problema distinto para evaluar la teoría de ejemplares. Claramente en esta teoría un concepto como PEZ MASCOTA codificaría no las memorias de PEZ y de MASCOTA, sino más bien las instancias de peces mascota identificadas como tales en el aprendizaje del concepto. Así, PEZ MASCOTA no se compone de PEZ y MASCOTA, lo que

¹⁰⁷ Como ya lo señalé en el capítulo 1 cuando hablo de composicionalidad, estos casos se pueden evitar de distintos modos. El más rápido es no comprometiéndose con un paralelismo fuerte entre lenguaje y pensamiento. Pero también uno puede apelar a una noción de propiedades relacionales que puedan dar cuenta de una composicionalidad de conjunción de propiedades en este caso. HORSERACE puede ser tratado como HORMIGA GRANDE. El caso de Fodor contra los prototipos depende de dos cosas: primero mostrar que PETFISH no es como HORMIGA GRANDE, y que por lo tanto sigue siendo un reto para los teóricos de los prototipos, y segundo mostrar que si así fuera el caso HORMIGA GRANDE puede tratarse como un caso de conjunción de propiedades.

creo sería el caso para cualquier concepto que podamos imaginarnos, en el marco de esta teoría. Lo anterior, sin embargo, no significa que esta teoría no conciba los conceptos como estructurados, sino que más bien la información que portan los conceptos está estructurada, no a partir de rasgos o características, sino como un conjunto de instancias recordadas como ejemplos de la categoría. En este caso las relaciones inferenciales con otros conceptos no parecerían jugar ningún papel en la individuación de cada concepto. Así, cada concepto, aunque sea entendido como una entidad estructurada, no está compuesto de otros conceptos sino de las aludidas „memorias’. Si bien esta no es la interpretación común de esta teoría, creo que es lo que debieran decir sus proponentes si se ciñen a las descripciones que hacen de los conceptos cuando pretenden revelar su estructura a partir de los efectos del ejemplar mostrados en sus experimentos de categorización.

Sin embargo, si se acepta la interpretación esbozada en el párrafo anterior, cabría preguntarse qué diría uno de estos teóricos en el caso de que a un individuo se le pida categorizar objetos en una categoría nueva para él, pero que, al menos en la forma lingüística se componga de conceptos ya aprendidos por el sujeto. Por ejemplo, si se le pide que decida si ciertos objetos que le serán presentados son vacas verdes, ¿cuál sería la estrategia del sujeto? Según la interpretación dada, y que exime a esta teoría de dar cuenta de la composicionalidad pues niega su papel a nivel conceptual, el individuo simplemente no podría responder a la tarea al no tener un ejemplar „guardado’ en tal categoría (necesariamente si la categoría es nueva para el individuo, dada la interpretación). Ahora bien, como ésta no parece ser una idea plausible

sobre el comportamiento del individuo tendríamos que aceptar que de alguna forma compone el nuevo concepto a partir de sus constituyentes, VACA y VERDE, pero en este caso resulta difícil imaginar qué características tendría una combinación de los ejemplares que el individuo retiene en ambos conceptos, donde posiblemente todos los ejemplares de vacas que tenga no sean verdes, y donde los ejemplares de cosas verdes que tenga no incluyan ninguna vaca. Lamentablemente, tal como lo señala Murphy (2002), es difícil evaluar esta teoría en particular debido a que sus proponentes no han sido claros respecto de la naturaleza de los conceptos, y utilizar la expresión „memorias’ no responde satisfactoriamente los problemas que debe resolver o explicar una teoría de conceptos.

En el caso de la teoría-teoría parece insostenible el principio de composicionalidad tal como se ha expuesto. Al estar determinado el contenido de manera holista no es posible sostener algún tipo de “herencia” de contenido por parte de las representaciones complejas desde las más simples. Aun cuando se aceptara algún tipo de holismo local, es decir que la determinación del contenido de las representaciones está dado por su rol dentro de mini-teorías, sigue siendo una característica de esta teoría que el contenido del concepto componente también es afectado por su participación en el concepto compuesto¹⁰⁸.

Finalmente, para la teoría de Prinz, se predice que si no tenemos experiencia previa relativa a una categoría, y el contexto no nos sirve de guía

¹⁰⁸ Esta objeción se hace más clara si se considera tanto que la teoría-teoría es parte de las teorías psicológicas que rechazan a la teoría definicionista, como el compromiso que tiene con el cambio conceptual y, por tanto, con el rechazo de la tesis del continuo.

en la formación de la representación de la misma, entonces el proxitipo que se forma es similar al prototipo. Sin embargo, justamente una de las cosas que parece estar clara respecto de los prototipos es que no componen: el prototipo de una representación compleja no es función de los prototipos de sus componentes. Por lo tanto, los proxitipos así entendidos tampoco cumplen con el requisito de composicionalidad.

La principal estrategia de Prinz para resolver el problema de la composicionalidad es diferenciar entre una noción fuerte de composicionalidad y una débil (2002:291). Según la versión fuerte los conceptos **deben ser** composicionales, mientras que de acuerdo a la versión débil lo que se requiere es que **sean capaces de** combinación composicional. En la concepción fuerte el contenido de los conceptos debe ser insensible al contexto¹⁰⁹, a fin de que sea siempre el mismo contenido el que es aportado por un componente a su compuesto. Este requisito es imposible de satisfacer por parte de los proxitipos. Sin embargo, si se toma la concepción débil sólo se requiere que podamos generar conceptos complejos vía composición, pero no que siempre tenga que ser así.

Sin embargo, dada la noción de composicionalidad alcanzada en el capítulo 1, mi propuesta está comprometida con la noción fuerte de composicionalidad. Claramente esto no es una objeción en sí misma a las teorías que no asumen la arquitectura computacional clásica. Prinz puede

¹⁰⁹ El uso de expresiones que pertenecen propiamente a la filosofía del lenguaje puede resultar equívoco aquí. Para la presente discusión la “sensibilidad al contexto” debe entenderse como la posibilidad de que el contenido de la representación varíe de acuerdo a las computaciones o procesos en los cuales participa. Dada la noción de composicionalidad expresada en el capítulo 1, tal sensibilidad no está permitida dentro del marco de investigación en el cual se inscribe esta investigación.

darse por satisfecho con la composicionalidad débil pues está comprometido con una arquitectura conexionista. Como señalé en la introducción, sólo considero la propuesta de Prinz a modo de ejemplificar su necesidad de cambiar de arquitectura cognitiva, dada la noción de concepto que está ofreciendo. Creo que esto constituye otra forma de presentar la composicionalidad fuerte de la que habla Prinz como un constreñimiento de la arquitectura computacional clásica. De hecho, la propuesta de los proxitipos se aleja de las otras propuestas revisadas en este capítulo sólo en cuanto se considere como implementándose en otro tipo de arquitectura. Por ejemplo, en los casos específicos en que la teoría de los prototipos tenía este problema, las respuestas de Prinz son:

- a) Para el caso de los conceptos que no tienen prototipos, señala que es una cuestión empírica determinar si realmente la gente no tiene un prototipo asociado a tales conceptos (donde la noción de prototipo ya no es una lista de rasgos codificados en una representación mental, sino aquellas conexiones que han adquirido más peso dada la experiencia de los individuos);
- b) En cuanto a las propiedades emergentes de los conceptos complejos (respecto de los conceptos constituyentes de los mismos), se pueden tener tres estrategias:
 - i) que no son compuestos, al menos no siempre;

- ii) que tales propiedades sí son derivadas de los componentes, pero que sólo se hacen prominentes, y funcionan como detectores, en el concepto complejo, o
- iii) que son excepciones al requerimiento, pero no el caso normal.

Para hacer plausible sus respuestas, Prinz plantea que la composicionalidad debe entenderse en su versión débil, y que más bien la utilizamos para concebir representaciones complejas cuando no es posible utilizar otra forma, por ejemplo vía la experiencia directa de la categoría a representar. Aquí Prinz está pensando en los casos donde tenemos una descripción altamente compleja como LOS ANIMALES QUE PESAN MÁS DE 100 GRAMOS Y VIVEN EN ÁRBOLES DE HOJAS AMARILLAS QUE FLORECEN DURANTE LA PRIMAVERA EN EL SURESTE DE MEXICO CUANDO NO HA NEVADO EL INVIERNO ANTERIOR. En estos casos probablemente nuestras representaciones de la categoría no se adecuen de manera confiable a lo representado, pero en vistas a que no tenemos experiencia directa ni un contexto teórico al cual poder recurrir, y que no son el tipo de categorías que parecen de mucha utilidad en las tareas cognitivas, es lo mejor que podemos tener.

Creo que la respuesta de Prinz muestra justamente lo que señalé en el capítulo 1, en cuanto a que el principio de composicionalidad, entendido de manera fuerte, es un constreñimiento de la arquitectura clásica.¹¹⁰

¹¹⁰ La estrategia general de Prinz consiste en desechar la hipótesis del lenguaje el pensamiento. Su idea es que el error de Fodor, de quien reconoce su teoría semántica respecto del contenido que denomina

3.4. Conclusión

Según lo visto, los intentos de caracterizar el contenido cognitivo a partir de lo que se revelaría en las tareas de categorización implica la incapacidad de asumir la noción de composicionalidad fuerte tal como se presentó en el capítulo 1. Si bien esto no constituye un argumento definitivo en contra de estas teorías, espero que lo mostrado hasta ahora si deje claro que son incompatibles con el marco de fondo que se ha asumido. Si las hipótesis del sistema de símbolos físicos y del lenguaje del pensamiento son asumidas, entonces la noción de composicionalidad fuerte no es negociable. Si las teorías revisadas en este capítulo no son compatibles con tal noción de composicionalidad, entonces tampoco lo son con el marco que definen las dos hipótesis señaladas.

Ahora bien, así como respecto de los conceptos primitivos pareciera ser que el marco de fondo impone restricciones que no podrían cumplirse, en este caso surge con más fuerza esa impresión. Si no es posible dar cuenta de manera adecuada de dos fenómenos que a primera vista parecen ser fundamentales en cuanto a las motivaciones para postular una teoría de las

intencional, es asumir que las representaciones mentales son amodales, *language-like*, esto es que están instanciadas en un código que no rescata las propiedades sensoriales de sus relata. Sin embargo, cuando se enfrenta a casos de conceptos cuyos relata, si los hay, no parecen ser analizables en términos de propiedades sensoriales, su mejor estrategia es asumir que tenemos representaciones en formato modal de las palabras mismas, en cuanto expresiones materiales con propiedades sensoriales, y que son las representaciones de esas sensaciones, más el contexto (ahora si físico y social) donde fueron aprendidas, las que figuran en los pensamientos que involucran tales conceptos.

representaciones mentales, entonces uno debiera considerar la posibilidad de cuestionar los fundamentos del marco de fondo. De hecho, esta es la línea que sigue Prinz, en cuanto a rechazar el formato amodal de las representaciones, es decir la hipótesis del lenguaje del pensamiento, y que pareciera que muchos psicólogos cognitivos han tomado, aunque sin reconocerlo explícitamente¹¹¹.

Sin embargo, de forma similar a la estrategia del capítulo 2, en el capítulo que sigue intentaré mostrar que es posible dar cuenta de los fenómenos del contenido cognitivo y las tareas de categorización a partir de distinciones más finas respecto de lo que podría ser la forma de las representaciones mentales, por un lado, y por otro respecto del papel que juega el contenido en las inferencias que es posible realizar al poseer un concepto y las inferencias que se siguen del propio procesamiento de la información que requiere, como condición previa, la posesión de conceptos.

¹¹¹ Aun cuando en los desarrollos de las teorías de conceptos de inspiración psicológica parece suponerse el marco de investigación asumido en esta investigación, no es menos cierto que algunas de esas teorías resultan mucho más adecuadas si se asume una arquitectura cognitiva de tipo conexionista, como por ejemplo en el caso de la teoría de los ejemplares.

CAPÍTULO 4

Solución: Contenido y Forma

“De entrada, la regla es: si hay algún problema para el que pareciera que necesitas sentidos a fin de resolverlo, o bien hazlo con la sintaxis, o bien no lo hagas en absoluto.” Fodor, 2008:87.

Según lo visto en el capítulo anterior, los intentos provenientes de la psicología cognitiva de integrar el contenido cognitivo dentro del contenido conceptual implican considerar a este último como estructurado de una forma en la que se viola la condición de composicionalidad, según la cual en la medida en que nuestras capacidades cognitivas son ilimitadas, pero nuestros recursos cognitivos son finitos, la productividad y sistematicidad de la mente requiere de una estructura composicional de los conceptos. Esto es, los conceptos son los constituyentes de los pensamientos y, en indefinidamente muchos casos, de otros conceptos por lo que las representaciones mentales heredan sus contenidos a partir de los contenidos de sus constituyentes. Así, el intento antes descrito debe ser rechazado¹¹².

En este capítulo argumentaré que las nociones de contenido cognitivo y los intentos de introducir elementos dentro del contenido conceptual, por ejemplo a partir de lo que revelan los experimentos de categorización, no es necesaria, en un primer momento, y que resulta incompatible con la teoría de fondo que se ha asumido.

Respecto del contenido cognitivo y los problemas que éste conlleva para una semántica de los conceptos puramente referencialista, considero que esto puede ser explicado en términos de la forma/sintaxis de los conceptos. Para lograr lo anterior será necesaria una distinción clara entre contenido y forma

¹¹² Si bien Prinz (2002, Cap. 11) intenta mostrar que la condición de composicionalidad no es esencial, o que al menos se puede debilitar lo suficiente como para que su propuesta integre efectivamente el contenido cognitivo, no discutiré aquí dicho intento, pues la propuesta de Prinz implica el no aceptar la existencia de conceptos tal y como se suponen en la presente discusión.

basada en el marco de fondo de esta investigación, a saber, las tesis computacional y representacional de la mente. Esto implica ir más allá de las intuiciones a las que nos lleva la analogía con el lenguaje natural, por ejemplo separándose claramente de las discusiones en filosofía del lenguaje respecto de la semántica de las oraciones de atribución de actitudes proposicionales¹¹³.

En cuanto a la categorización, entendida ésta como la otra fuente de dificultades para la semántica referencialista defendida en esta tesis, la idea es que lo que revelan los experimentos de categorización y los casos de Frege supone la posesión de conceptos, y más bien nos dice algo acerca de su forma, en el sentido de cómo es que utilizamos los conceptos en cuanto a los procesos de pensamiento (computaciones) para los cuales están disponibles tales y cuales representaciones (acceso desde la memoria, reconocimiento de objetos, etc.). No es necesario introducir esas capacidades cognitivas, basadas en la forma de los conceptos, en la determinación del contenido. Aún más, como se mostró en el capítulo anterior tal introducción resulta incompatible con la identidad de contenido requerida por el supuesto computacional y el principio de composicionalidad ligado a éste (defendido en el capítulo 1).

Finalmente, luego de separar claramente lo que corresponde a la forma y el contenido del concepto, presento una estrategia de solución para hacer de la relación entre ambos componentes de los conceptos una relación no arbitraria. Brevemente, en cuanto la forma de los conceptos primitivos depende de ciertos mecanismos que fijan su referencia de manera no arbitraria,

¹¹³ Si bien podrían seguirse consecuencias para las discusiones en filosofía del lenguaje sobre la semántica de estas oraciones, considero que no es pertinente aquí entrar en estas cuestiones.

entonces es posible asegurar tal relación para los primitivos. Por lo tanto, en tanto los complejos derivan su contenido de los primitivos, esto asegura también una relación no arbitraria entre sus contenidos y las capacidades inferenciales asociadas a su posesión.

4.1. Contenido Cognitivo reconsiderado

En el capítulo anterior se introdujo la noción de contenido cognitivo para dar cuenta de lo que en la literatura se suele llamar los casos de Frege¹¹⁴. Lo que sucede en tales casos es que la pura referencia de un símbolo o representación parece no bastar para hacer diferencias en casos que involucran la importancia cognitiva para un sujeto que posee tal representación. Por lo tanto, no resulta posible dar cuenta de las diferencias en valor cognitivo de ciertos juicios de identidad que no resultan obvios (cognoscibles a priori) para un individuo. Lo anterior lleva a postular otra categoría semántica además de la referencia, a saber el sentido de una representación, destinada a cumplir la tarea de diferenciar conceptos correferenciales que sin embargo a nivel cognitivo deben distinguirse.

¹¹⁴ Como ya lo indiqué anteriormente, utilizo “casos de Frege” para referirme a la constelación de objeciones surgidas a partir de las consideraciones que llevaron a Frege a postular la categoría semántica de sentido, y en ningún caso corresponden a objeciones de Frege. De la misma forma, el rechazo de los sentidos en la determinación del contenido conceptual no corresponde necesariamente a todas las vertientes que surgen como interpretación de lo que quiso decir Frege por “sentido”, sino que principalmente a aquellas que sostienen que los sentidos son independientes de la referencia y por tanto tienen implicaciones internistas.

Una forma de prescindir de los sentidos al momento de caracterizar el contenido de las representaciones mentales es apelar a la forma o sintaxis de tales símbolos. En la medida en que los conceptos tienen tanto contenido como forma, en principio sería posible tener dos conceptos que se distingan no por su contenido sino por su forma, dejando así una salida para la semántica referencialista al problema del contenido cognitivo. De hecho, la cita que da inicio a este capítulo es antecedida por lo siguiente:

“... después de todo, un tema principal de la Teoría Computacional de la Mente¹¹⁵ es definir los procesos mentales sobre las propiedades sintácticas de las representaciones mentales, evitando así una psicología que está comprometida con los significados, con respecto a los sentidos, la Teoría Computacional de la Mente es un programa de reducción *eliminativista* ... A la larga, la psicología computacional es una especie de truco que Turing inventó para hacer que parezca que hay sentidos y que hacen cosas (aunque, estrictamente hablando, no existen (y por tanto) y no hacen).” Fodor, 2008: 87.

Ahora bien, lo anterior constituye una solución en principio. Sin embargo, es necesario revisar si es posible articularla de tal forma que responda realmente a los problemas planteados. Si consideramos que el problema surge a partir de Frege, entonces hay que recordar que él mismo en un primer momento tomó esta opción en su *Conceptografía* (1879), pero que justamente la introducción del sentido se debe a las fallas que encontró en tal solución.

¹¹⁵ Teoría Computacional de la Mente, que Fodor utiliza como idéntica a Teoría Representacionista de la Mente.

En lo que sigue de este apartado mostraré cómo es posible, dentro del marco de investigación aquí asumido, hacer que funcione la solución sintáctica y, por tanto, deshacerse de los sentidos. En primer lugar trazaré la distinción entre contenido y forma a nivel computacional para luego ver cómo es que esta funcionaría para el caso de los conceptos. En el siguiente apartado consideraré los problemas que presenta la evidencia surgida de las tareas de categorización en psicología.

4.1.1. Computación: Contenido y Sintaxis.

De forma similar a lo que se presentó en el capítulo 1 respecto del principio de composicionalidad, al trazar la distinción entre contenido y forma de los símbolos físicos que constituirían los elementos básicos de nuestra mente¹¹⁶ resulta necesario alejarse de las discusiones relativas al lenguaje natural y concentrarse en la perspectiva computacionalista.

En el marco computacional los procesos que se realizan sobre los símbolos se hacen en virtud de su forma o sintaxis. En este caso la forma corresponde a la configuración física del símbolo y es responsable de sus poderes causales. Si entendemos la eficacia causal de un símbolo como el conjunto de procesos en los cuales éste puede participar, debido a la forma

¹¹⁶ Junto a las reglas de procesamiento, dado el marco de investigación asumido.

que tiene, y si ese conjunto de procesos al nivel cognitivo a su vez se entiende como las inferencias posibles del símbolo en cuestión, entonces parece plausible asumir que los papeles inferenciales que definen la eficacia causal del símbolo son sensibles a su forma y no a su contenido.

Considérese un sistema computacional simple, es decir una máquina física capaz de recibir un insumo físico que porta información, procesar dicha información en virtud de su forma física, y entregar un resultado que respete el contenido del insumo inicial. Por ejemplo, la máquina expendedora de café recibe monedas de distintos valores, en virtud de su sensibilidad a la forma de las monedas, literalmente, se inician una serie de procesos o computaciones internas que entregan como resultado un café y el cambio en caso de que el dinero ingresado a la máquina sea superior al costo del café. En este caso simple resulta claro cómo las computaciones de la máquina son “ciegas” al contenido de los símbolos físicos, en este caso el valor de las monedas tanto del insumo como del cambio entregado, pero sin embargo respetan tal contenido.

Un caso análogo puede hacerse para el caso de una calculadora. Los contenidos sobre los que opera la máquina, y a los cuales es ciega, son cantidades, y las computaciones que realiza sobre esas cantidades deben ser sensibles a las formas de los símbolos. Así, si bien cuando se piensa acerca del lenguaje matemático uno podría no tener claridad respecto de si “al ser sumado con 6, da 9”, “al multiplicarse por 5, da 15”, “al restarle 2, queda 1” son características que definen el contenido del símbolo “3” o bien si sólo

corresponden a su sintaxis en dicho lenguaje, parece claro que al nivel de la calculadora las operaciones en las cuales participa el símbolo “3” no pueden ser parte de su contenido sino que son parte de sus papeles inferenciales, por tanto de su eficacia causal dentro de la economía cognitiva de la máquina, y por lo mismo, finalmente, características a las cuales es sensible el procesamiento de la información. De lo anterior se sigue que, al menos al nivel computacional, es decir en la caracterización funcional del procesamiento, la forma corresponde a los papeles inferenciales en los cuales puede participar un símbolo¹¹⁷.

4.1.2. Conceptos complejos

Una de las formulaciones ya clásicas del problema de Frege consiste en el enunciado de identidad “El lucero de la mañana es el lucero de la tarde”. Desde la perspectiva puramente referencial donde el significado de los términos se agota en la referencia de los mismos, la identidad expresada no aporta ningún conocimiento pues sólo se está enunciando que un objeto, Venus en este caso, es idéntico a sí mismo. Sin embargo resulta claro que el enunciado en cuestión afirma algo distinto que el principio de identidad. Así,

¹¹⁷ En este sentido la individuación funcional de los conceptos parece ser irremediabilmente holista. Sin embargo, en la medida en que puedan postularse, y se obtenga evidencia para, módulos cognitivos especializados, podría sostener que algunos conceptos dentro de estos módulos, además de ser innatos, podrían individuarse por sus relaciones, tanto de contenido como de forma, con otros conceptos dentro del sistema cognitivo modular. En este caso habría un holismo local y no involucraría inferencias posibles fuera del procesamiento modular. Otra solución, que expongo y favorezco en este capítulo, consiste en asumir que sólo las inferencias analíticas son constitutivas del contenido del concepto, y que se derivan de la forma de las representaciones complejas, en tanto esta forma parte de la composición de tales representaciones.

“lucero de la mañana” y “lucero de la tarde”, a pesar de tener la misma referencia, deben distinguirse.

En primer lugar quiero dejar de lado la cuestión acerca de si las oraciones del lenguaje natural expresan pensamientos, entendidos estos como representaciones mentales complejas que a su vez expresan proposiciones que son las portadoras en última instancia del contenido en el sentido de condiciones de verdad. Esto es, aceptemos por bien del argumento que en el lenguaje natural podemos expresar una proposición completa por medio de la emisión de un enunciado como el anterior, y que a tal enunciado corresponde una representación mental compleja de forma similar a la del enunciado. Entonces tenemos una representación mental similar al enunciado, compuesta de dos conceptos complejos y unidos por la relación de identidad. La pregunta se vuelve entonces en cómo podemos distinguir los conceptos complejos LUCERO DE LA MAÑANA y LUCERO DE LA TARDE. Parece obvio que ambas representaciones son distintas pues una incluye como constituyente el concepto TARDE y la otra el concepto MAÑANA. Para poseer el primer concepto complejo es necesario poseer el concepto de MAÑANA, sin embargo es perfectamente posible poseer el segundo concepto complejo aun cuando se carezca del concepto MAÑANA. Claramente TARDE y MAÑANA son conceptos distintos, tanto en su contenido como en su forma, entendida computacionalmente, por lo que no es de sorprender que los conceptos complejos que respectivamente constituyen sean a su vez distintos.

En la medida en que aceptemos que los conceptos complejos se constituyen exhaustivamente por el contenido y la forma de sus constituyentes entonces no parece haber problema para distinguir conceptos complejos correferenciales, en tanto podamos distinguir en su estructura elementos constituyentes distintos.

4.1.3. Conceptos primitivos

La solución a los casos de Frege esbozada en el apartado anterior no está disponible para el caso de los conceptos primitivos pues éstos justamente carecen de estructura. No es posible apelar a diferencias en los constituyentes para explicar la diferencia entre conceptos primitivos correferenciales.

Lo primero que se debe notar es la dificultad de encontrar conceptos primitivos correferenciales que sean distintos. La mayor parte de la literatura en filosofía del lenguaje concerniente a este asunto suele centrarse en términos singulares y no en el caso de nombres comunes como de los que se trata en este trabajo. Conceptos como TRIANGULAR y TRILATERAL, CORDADO y RENADO¹¹⁸, parecen asemejarse a los casos de conceptos complejos revisados en el apartado anterior. De la misma forma, si aceptáramos que para las expresiones del inglés “groundhog”, “woodchuck”, y “land-beaver” (distintos

¹¹⁸ “cordado” como abreviatura de “criatura con corazón” y “renado” para “criatura con riñones”.

nombres que se le dan a la marmota) existen conceptos distintos, creo que también puede rastrearse la diferencia a los componentes de la misma.

Pese a la dificultad señalada, consideraré el caso de un par de términos singulares correferenciales, asumiendo que hay dos representaciones mentales primitivas correspondientes a cada uno. Creo que lo que se diga respecto de este caso valdrá, *mutatis mutandi*, para el caso de dos conceptos primitivos correferenciales correspondientes a nombres comunes. Lo importante a considerar es que mi propuesta, en tanto externista, está comprometida con el que no pueden existir dos conceptos primitivos distintos cuya referencia sea la misma y cuya diferencia no sea explicable en términos de una diferencia de forma. Asimismo, mi propuesta no puede aceptar la existencia de conceptos primitivos vacíos, pues la falta de referencia de un concepto sólo podría deberse a una falla debido a la sistematicidad del pensamiento, de acuerdo a la noción de composicionalidad con la que estoy comprometido, por lo que un concepto referencialmente vacío debe ser complejo.

Considérese el caso de los nombres propios “Pablo Neruda” y “Neftalí Reyes”. Si se asume una semántica referencialista, y que los nombres propios no tienen estructura en el sentido antes descrito, entonces ambos términos tienen el mismo significado pues refieren a un mismo individuo. Sin embargo es posible para un individuo racional dudar o incluso negar el enunciado “Pablo Neruda es Neftalí Reyes”. Digamos que sabe que Pablo Neruda es el autor de “20 poemas de amor y una canción desesperada”, y que también ha escuchado

hablar de un hijo de ferrocarrilero de la Araucanía chilena llamado Neftalí Reyes, pero que nunca ha pensado que se trata del mismo individuo¹¹⁹.

Ahora bien, asumamos que tal individuo tiene una representación mental cuando utiliza “Pablo Neruda”, del mismo modo que tiene una cuando utiliza “Neftalí Reyes”¹²⁰. El problema del contenido cognitivo en este caso resulta inevitable: el individuo debe tener dos representaciones mentales distintas a fin de poder tener la duda o negar la identidad¹²¹. Aún más, la diferencia entre las representaciones mentales no se puede deber a diferencias en el contenido referencial ni tampoco a la estructura, pues por ejemplo alguien que acepta la identidad producto de su conocimiento de la historia de la literatura chilena no sostendrá la distinción debido a la mera diferencia en las letras de los nombres, justamente porque utiliza ambos nombres para expresar el contenido de la única representación mental que tiene para el portador de ellos¹²².

En tanto he tomado un ejemplo proveniente de las discusiones en filosofía del lenguaje, cabe considerar aquí ciertas soluciones que si bien son ya propias de la filosofía de la mente, se encuentran en estrecha conexión con los problemas descritos y estudiados en filosofía del lenguaje. Una forma de intentar resolver el problema de los nombres correferenciales consiste en

¹¹⁹ Las consideraciones que siguen respecto de este caso se aplican de la misma forma a los casos “Paderewski” de Kipke, en cuanto la diferencia entre los nombres del lenguaje natural (“Pablo Neruda” v/s “Neftalí Reyes”) no hace una distinción al nivel de las representaciones mentales asociadas al uso de ellos.

¹²⁰ Nuevamente aquí me valgo del uso del lenguaje natural a fin de hacer más claro el ejemplo, pero nada de lo que se sigue depende de adoptar alguna teoría semántica en particular respecto del lenguaje natural. Las consecuencias para la filosofía del lenguaje que pudieran extraerse de la noción de contenido conceptual defendida en esta tesis no forman parte de la misma.

¹²¹ Negar la identidad no como acto de habla sino al nivel del pensamiento.

¹²² Estrictamente hablando, los conceptos *qua* particulares mentales no tienen letras, por lo que distinciones a nivel de pronunciación o escritura no corresponden en este nivel.

apelar a la noción de “archivo” y “nombre de archivo”, que se puede rastrear en Treisman y Schmidt (1982), G. Evans (1982) y J. Perry (1980). La idea es que la mente abre un archivo y lo nombra cada vez que se encuentra con un objeto o individuo. En la medida que obtiene información se busca el archivo, por medio del nombre, y se almacena allí esa nueva información (también puede borrarse información del archivo si fuese el caso). Lo importante es que la representación mental se identifica con el nombre del archivo y no con el archivo, éste último sólo tiene información asociada pero que no constituye el contenido del concepto. En el caso del individuo que duda o niega la identidad se dirá que tiene dos archivos distintos etiquetados bajo distintos nombres, donde la distinción debe ser formal o sintáctica para explicar cómo es que la mente reconoce o es sensible a uno u otro según sea el caso. En el caso del individuo que acepta la identidad dado su conocimiento, se puede decir que o bien almacenó la información “es también llamado Neftalí Reyes”, si abrió primero el archivo bajo el nombre PABLO NERUDA, o bien que tenía en un principio dos archivos y que luego de tener el conocimiento de la identidad los junto en uno solo.

Esta solución tiene el problema de volver a poner los papeles inferenciales, ligados a los archivos como bases del conocimiento ligado al concepto, demasiado cerca del contenido del mismo. De hecho, en principio parece estar sujeta a la crítica de Prinz ya presentada, en cuanto las capacidades epistémicas del individuo parecen depender de cuestiones semánticas y no puramente formales.

Kit Fine (2007) ofrece una solución que denomina “relacionista” para este tipo de problemas, según la cual es posible que sostener el externismo sin hacer tantas concesiones al papel inferencial. En este caso no es la noción de información en archivos la que juega el papel de poner en contacto a dos representaciones, sino que es la noción relacional de “coordinación” entre representaciones la que lo hace. La idea de Fine consiste en asumir que la estrategia por defecto de los hablantes es considerar cada emisión de un nombre propio como refiriéndose a un individuo distinto, a menos que hayan buenas razones para inferir lo contrario. Así, en el caso del individuo que no tiene motivos para asumir que se está hablando del mismo individuo, la propuesta relacional asume que hay dos conceptos o representaciones que no están coordinadas. En el caso del individuo que posee la información relativa a la verdad del enunciado de identidad, se dice que las representaciones se coordinan. Lo importante para mi propuesta en este caso es el supuesto de que la estrategia por defecto consiste en asumir representaciones distintas para emisiones distintas, ya sea de nombres propios distintos pero correferenciales, o incluso de emisiones del mismo nombre (para referir al mismo individuo) en las cuales el intérprete no tiene razones para asumir que se le está hablando del mismo individuo, por ejemplo en el caso “Paderewski”. Además, la coordinación así entendida parece sugerir que su papel no es semántico sino sintáctico. Pero incluso si la noción de coordinación es considerada como semántica, ésta es extrínseca a la representación, por lo tanto no es constitutiva de la misma sino sólo del estado mental complejo en el cual participa. De esta forma, la coordinación, dentro de mi propuesta, podría tomarse como parte de las reglas que constituyen los estados mentales

entendidos como representaciones complejas, sin alterar el contenido de las representaciones constituyentes.

Si bien creo que la postura de Fine es al menos compatible con mi propuesta, lo que me importa es que, por lo ya dicho en el apartado sobre la distinción entre contenido y forma a nivel computacional, si un individuo emplea dos representaciones mentales, lo que se evidencia en este caso en su comportamiento lingüístico, entonces necesariamente si la distinción no existe al nivel del contenido referencial ésta debe encontrarse al nivel sintáctico. En este caso el individuo tiene que tener dos representaciones mentales distintas, por su forma, para explicar por qué realiza las inferencias que de hecho realiza. Tal como se señaló en el capítulo 2, los conceptos primitivos lo son en cuanto sus condiciones de individuación no dependen de la existencia, y sus relaciones, con otros conceptos, pero esto no implica que uno no pueda distinguir componentes en ellos, a saber en términos de su contenido y su forma¹²³.

4.2. Categorización: conceptos y su uso

Lo presentado en el apartado anterior nos da una respuesta a los problemas que se presentan para una teoría referencialista del contenido. Lo

¹²³ Así, si se encontrará un caso de conceptos primitivos correferenciales, mi postura está comprometida con que la diferencia debe estar en la forma de los mismos, y por ende en sus papeles inferenciales. La otra posibilidad es que no sean correferenciales, aun cuando puedan ser co-extensivos, en este caso la distinción ocurriría a nivel semántico, pero asumiendo que la referencia es la propiedad que expresa el concepto, y no la extensión de los mismos. Esta solución estaría disponible para el caso de CORDADO y RENADO, donde la mera co-extensión, en cuanto hecho contingente, no es suficiente para decir que ambos conceptos tienen el mismo contenido.

anterior no significa dejar de lado completamente el contenido cognitivo. Más aún, en cuanto lo que se denomina contenido cognitivo es dejado fuera de la caracterización semántica del contenido conceptual pero se inserta en la individuación de los conceptos vía su forma, entonces resulta evidente que lo que reflejan los experimentos de categorización si nos dicen algo respecto de los conceptos.

Tal como Prinz lo hace notar, tenemos ciertos mecanismos que nos permiten formar creencias acerca de qué objetos caen bajo los conceptos, y el problema que intenté destacar es justamente que las aproximaciones psicológicas a la teoría de conceptos suelen identificar dichos mecanismos con los conceptos mismos. Lo que se sigue de lo mostrado anteriormente es que tales mecanismos tienen que identificarse con la forma del concepto, pues es justamente ésta la que es responsable causalmente, aunque individuada en términos funcionales, con las inferencias que se realizan en el procesamiento cognitivo. Como ya se vio, tales inferencias incluyen todos los procesos cognitivos en los cuales participa el concepto, especialmente su participación en creencias y deseos, tanto en su fijación como en la producción de conducta inteligente. Por lo mismo, en términos de explicar la conducta de un individuo por medio de la atribución de una determinada representación mental no es suficiente individuar su contenido, sino también los procesos inferenciales en los cuales participa¹²⁴.

¹²⁴ Por eso tal vez, al notar que hay dos lecturas de las oraciones de atribución de actitudes proposicionales, a saber una transparente y una oblicua, en un caso lo que cuenta es la pura referencia y en el otro el cómo el individuo tiene re-presentado el referente. La distinción entre “concepto” y “concepción” tal vez también pueda explicarse de este modo.

Por otra parte, Fodor insiste en que ninguna de las relaciones inferenciales en las cuales participa un concepto es constitutiva de su identidad. Para individuar un concepto todo, y sólo, lo que cuenta es su extensión o referencia. Sin embargo acepta que para la predicción de la conducta es necesario apelar a todas las creencias, deseos, esperanzas, etc., en las cuales se involucra el concepto, esto es, que apelar sólo al contenido individuado referencialmente no alcanza para la psicología.

Creo que una posibilidad no siempre considerada es que así como Fodor y Prinz caracterizan la categorización de formas completamente distintas, el primero centrado en la relación semántica¹²⁵ y el segundo como la contraparte epistémica de la referencia, es posible que la discusión precedente se deba a un desacuerdo respecto de qué es lo que se está individuando. Sin embargo la cuestión parece ir más allá de un problema terminológico, Fodor insiste en que la contraparte epistémica no es esencial a la individuación conceptual. Así podemos leer la advertencia de Fodor (1998:25), respecto de la condición no-negociable de la categorización en cuanto a que no es tendencioso asumir que los conceptos tienen sus condiciones de satisfacción, a nivel semántico, esencialmente, es decir que nada podría ser el concepto ÁRBOL a menos que fuera satisfecho por todos y sólo los árboles. Ahora bien, las condiciones de satisfacción semántica no necesariamente son las mismas

¹²⁵ La condición de categorización formulada por Fodor (1998, Cap. 2) es completamente distinta de la formulación de Prinz. Según Fodor, “los conceptos son categorías y son rutinariamente empleados como tales”, su explicación es que los conceptos funcionan como categorías en cuanto se aplican a cosas del mundo, o bien que las cosas del mundo ‘caen’ bajo ellos. Así, por ejemplo, el concepto ÁRBOL se aplica al individuo que crece en el patio de mi casa y a un número indefinido de individuos que satisfacen el concepto. Por otro lado, cuando digo del individuo que crece en el patio de mi casa que es un ‘árbol’, estoy diciendo que ‘cae’ bajo ese concepto, que forma parte su extensión. En este sentido, la categorización está estrechamente vinculada a la evaluabilidad semántica, en cuanto uno puede categorizar un objeto como perteneciendo a la extensión de un concepto de forma errónea o correcta.

que las condiciones de confirmación, las primeras son metafísicas y esenciales, y las segundas epistémicas, generalmente mediadas por teorías y, por lo tanto, holistas. Así, es perfectamente compatible ser atomista respecto de las condiciones de satisfacción y ser holista respecto de las condiciones de confirmación. Aquí Fodor claramente reduce la condición de categorización a lo que podría llamarse 'Evaluabilidad Semántica' o 'Determinación de la Referencia', en el sentido de que una aplicación de un concepto sobre un objeto puede ser evaluada como correcta o incorrecta, o un enunciado donde se atribuya a un sujeto un concepto, como verdadero o falso. Al mismo tiempo parece desechar como característica para la individuación conceptual las relaciones epistémicas en las que participa o puede participar la representación mental¹²⁶.

Creo que el identificar el concepto únicamente con su contenido es lo que explica la divergencia que es tomada no como meramente terminológica. De forma similar las disputas entre los teóricos de los prototipos y los definicionistas suelen girar en torno a qué es un concepto, aunque en última instancia son acerca de la caracterización del contenido conceptual. Así por ejemplo Rey (1983) desecha las críticas de los teóricos de los prototipos por estar centrados en aspectos epistémicos y no en las condiciones metafísicamente necesarias para ser un concepto, mientras que por su parte los teóricos de los prototipos señalan que la teoría definicionista no tiene realidad psicológica (Murphy, 2002).

¹²⁶ Y por ende afectar la condición de publicidad que requiere Fodor para hacer sostenible su atomismo respecto de los conceptos léxicos. Tal problema está ligado a la posibilidad de una psicología científica que postule leyes intencionales y no es esencial a lo presentado en esta tesis.

Sin embargo, el propio Murphy (2002), enfrentado a la evidencia de tareas de categorización ofrecidas dentro de las teorías de inspiración psicológica (prototipos, estereotipos y teoría-teoría), sugiere que tal vez sea necesario separar lo que tiene que ver con el acceso al concepto, que podría ser un asunto de memoria y otras capacidades cognitivas, de lo que tiene que ver con el contenido conceptual, y que tal vez lo que revelan los experimentos de categorización no es acerca del contenido sino del acceso. En este sentido puede pensarse que muchas de las características exhibidas en los experimentos de categorización respondan más bien a los procesos que operan sobre las representaciones, dadas la naturaleza de la tarea y los recursos disponibles. Esto explicaría por qué según cómo sea diseñada la tarea se obtienen resultados compatibles con la teoría que se esté testeando. Más importante aún, es posible obtener evidencia que a primera vista resulta incompatible con los postulados de las otras teorías, como en el caso de los juicios de tipicidad para las teorías definicionistas.

Si esto es así, entonces tenemos buenas razones para pensar que es posible considerar el contenido conceptual desde la perspectiva referencialista y tomar lo que se ha llamado contenido cognitivo como aquello que responde a lo que he denominado la forma de los conceptos. Hacer esta distinción en principio parece resolver el dilema respecto de la categorización, entendida psicológicamente, y la composicionalidad del contenido de los conceptos complejos. Sin embargo persiste la divergencia sobre qué es lo que interesa cuando individuamos conceptos, o bien cómo hacer coincidir en una teoría ambos aspectos de la cuestión.

4.3. Contenido y forma re-unidos

Un problema de mantener totalmente separados contenido y forma consiste en que su relación podría ser perfectamente arbitraria, de forma similar al problema de la adquisición de conceptos primitivos y la relación evidencial que parece guiar tal adquisición. En principio uno podría tener un concepto con el contenido *gato*¹²⁷ y que su forma nos llevara a inferir que ningún gato del mundo real sea categorizado bajo el concepto GATO. A fin de cuentas, si contenido y forma van por separado, y la forma corresponde a los papeles inferenciales, nada parece impedir esa posibilidad.

Creo que al menos parte de la respuesta a la divergencia mencionada y, específicamente en cuanto a cómo hacer coincidir ambos aspectos de los conceptos, consiste en mirar hacia los conceptos primitivos, tal como se señaló en el capítulo 2 al hablar acerca de una manera de entender el concepto ROJO como primitivo no-innato. Otra posibilidad, teniendo presente la distinción entre contenido y forma ya alcanzada, consiste en pensar que algunos conceptos primitivos sean innatos, que su contenido esté determinado referencialmente, en términos de que evolutivamente se desarrollaron para tener un comercio con el ambiente que resultara ventajoso para los que lo poseyeran, y que su forma esté determinada por las características que tal representación está destinada a cumplir dentro de la economía cognitiva de individuos como

¹²⁷ Utilizo *cursivas* para mencionar el contenido.

nosotros. En tanto esta forma es producto de la evolución, es posible pensar que su determinación se deba a procesos ellos mismos no-conceptuales, tal vez describibles a nivel neurofisiológico. En particular, las teorías teleológicas del contenido intentan mostrar una relación entre el funcionamiento del sistema cognitivo y su relación tanto con el ambiente como con las respuestas conductuales de los individuos que los poseen. Por ejemplo, la biosemántica de Millikan (2000) postula que hay mecanismos encargados de la detección de rasgos del ambiente, que identifica como mecanismos productores de la representación o indicadores, y mecanismos encargados de consumir esa representación de acuerdo a las necesidades del organismo. Ambos mecanismos serían el producto de la selección natural, por su contribución a la adaptabilidad del organismo, y así asegurarían una relación no arbitraria entre el rasgo ambiental indicado y el contenido tal como es usado por el organismo. Creo que esta distinción podría servir para distinguir conceptos primitivos adquiridos a través de la experiencia (y determinados por mecanismos más básicos) y conceptos primitivos innatos que determinan la forma en que conceptualizamos el mundo. En este mismo sentido se podría dar cuenta de la importancia de los mecanismos que van de lo superior a lo inferior (*top-down*) tal como lo plantea Segal (1989) en su disputa con Burge (1982) acerca del externismo, sin tener que tocar la semántica de todos o la gran mayoría de los conceptos.

Aun si no pudiera tenerse una historia semántica del tipo ofrecido por Millikan para todos los conceptos primitivos, en la medida que mi propuesta acepta que puedan haber algunos conceptos innatos que se interdefinan y que

determinen de alguna manera la forma que tienen mentes como las nuestras de procesar la información, y así individuarse por su papel conceptual únicamente, entonces es posible encontrar espacio para este tipo de argumentos. Además, no todas las propuestas teleológicas son tan ambiciosas como la de Millikan¹²⁸. Por ejemplo Sterelny (1990) plantea una propuesta similar a la de Millikan pero sólo al nivel de los primitivos, y acepta una arquitectura combinatorial para los conceptos complejos. También Neander (2006) intenta una semántica “modesta”, sólo para los primitivos, que combina rasgos teleológicos e informacionales. Ambas posiciones resultan compatibles con lo propuesto en esta tesis.

Finalmente, para retomar la distinción hecha en el capítulo 2 entre lo conceptual y lo no-conceptual, o más bien pre-conceptual, un buen ejemplo de cómo podría esto funcionar se encuentra en el concepto OBJETO. Los índices visuales de Pylyshyn (2000) podrían jugar un papel no-conceptual en nuestra detección de objetos. La propuesta de Pylyshyn consiste en mostrar que nuestro sistema visual está capacitado para detectar objetos visuales o proto-objetos de una forma que no requiere ni utiliza capacidades conceptuales. Esta capacidad se limita a cierto número de proto-objetos (4 en adultos humanos) y a ciertas características (continuidad espacio-temporal, por ejemplo), y de hecho omite otras características de orden conceptual (el color, por ejemplo). Así, en el caso de los insumos que aporta el sistema visual al aparato cognitivo propiamente conceptual restringe las posibilidades de éste, y posiblemente determina parte del contenido del concepto objeto. Esta determinación, en

¹²⁸ Debido a mi compromiso con la arquitectura clásica sólo puedo tomar las ideas de Millikan como guías para construir una semántica de los primitivos, pues ella rechaza la hipótesis del lenguaje del pensamiento (1993, Caps. 3-5) y parece optar por una arquitectura conexionista.

tanto debe estar sustentada en algún tipo de mecanismo causalmente eficaz dentro de la economía cognitiva, debe reflejarse en la forma del insumo, y por ende también determinar la forma del concepto. De acuerdo a esto, el concepto OBJETO podría estar definido formalmente por los rasgos que operan a ese nivel pre-conceptual y constituirse así, por ejemplo, en el concepto de objeto que ha sido estudiado por Spelke (1993), el que sería innato y primitivo, además de operar dentro de un módulo cognitivo por lo que no se definiría holísticamente. Aún más, es probable que la aparición de ese concepto tenga que ver con las necesidades de los organismos que lo poseemos, y en ese sentido su configuración responda a presiones de orden superior-inferior. Otra posibilidad es que el aparato pre-conceptual que guía nuestra manera de conceptualizar los objetos simplemente termine en un proceso donde se adquiere el concepto OBJETO, y así éste último no sea innato sino que sólo esté determinado por capacidades innatas no-conceptuales¹²⁹.

Si algunas de las historias esbozadas en el párrafo anterior resulta ser la correcta respecto del concepto OBJETO (dentro de las esbozadas en el párrafo anterior), entonces es posible ofrecer una historia plausible acerca de la relación no-arbitraria entre el contenido y la forma de los conceptos primitivos, y por composicionalidad podemos asumir que tal falta de arbitrariedad se conserva en los conceptos complejos constituidos a partir de los primitivos. Aún más, hay ciertas inferencias privilegiadas que permiten individuar los conceptos complejos, y así evitar cualquier tipo de holismo semántico. Dada la noción de

¹²⁹ En el sentido defendido en el capítulo 2. Primitivo pues su posesión no depende de otros conceptos, pero no innato porque su adquisición estaría mediada por la operación de los recursos no conceptuales postulados por Pylyshyn. 2000.

composicionalidad del sistema de símbolos físicos defendida aquí, resulta que si hay inferencias analíticas o constitutivas de contenido¹³⁰, y son sólo estas las que individúan la forma del concepto complejo.

4.4. Conclusión

En el presente capítulo he explorado una forma de dar cuenta de los problemas que presentaban el contenido cognitivo y la categorización para una semántica referencialista de los conceptos. La idea fundamental para encontrar la solución pasa por tomarse en serio la distinción entre contenido y forma de una representación, y en particular por considerar tal distinción dentro de la perspectiva computacionalista que sirve como marco de fondo de esta investigación.

En términos computacionales cualquier diferencia entre representaciones debe expresarse a nivel formal, bajo el supuesto de que los procesos sólo son sensibles a las propiedades formales de los símbolos, y ciegos a su semántica. De este modo, en tanto se identifiquen los papeles causales de las representaciones mentales con sus papeles inferenciales, debido a la individuación funcional de la mente, resulta necesario considerar la

¹³⁰ Ahora bien, al entender “inferencia analítica” de este modo, privilegiando la noción de constitución de contenido en el sentido propuesto, se pierde en gran medida la publicidad de los conceptos complejos. En la medida que es posible adquirir conceptos primitivos por medio de la experiencia, y que estos primitivos deben corresponder a una minoría dentro del total de conceptos de un individuo cualquiera, entonces la variabilidad en los conceptos complejos formados puede ser muy grande de un individuo a otro. A la hora de individuar conceptos con fines de una explicación psicológica que subsuma a una gran cantidad de individuos, lo más que podría proporcionarse es un número limitado de condiciones necesarias que resultarían poco informativas respecto de la conducta que pudiera predecirse.

sintaxis de los símbolos como el conjunto de inferencias en las cuales puede participar, a menos que uno tenga una forma de distinguir entre inferencias privilegiadas y no privilegiadas, por ejemplo inferencias analíticas y no analíticas.

Así, si el contenido cognitivo, y por ende los procesos de categorización ligados a su posesión se expresan en capacidades de procesamiento, este rasgo de los conceptos debe identificarse con su sintaxis, entendida al modo computacional. De esta forma la semántica referencialista no se ve afectada por los usos cognitivos o epistémicos que involucran a los conceptos. Para el caso de los conceptos complejos la diferencia entre representaciones correferenciales puede explicarse por una diferencia en los constituyentes de los mismos. Para el caso de los primitivos, por otra parte, la diferencia debe estar a nivel formal. Por lo ya dicho respecto de la sintaxis computacionalmente entendida, si hay una diferencia en las inferencias que permiten dos conceptos, aun cuando sean correferenciales, entonces esta debe estar reflejada en la forma de los mismos.

Sin embargo, al llevar los papeles inferenciales a la forma conceptual, resulta necesario reconocer que tales inferencias son constitutivas del concepto, si bien no de su contenido, por lo que no es posible desechar totalmente estos rasgos de la individuación conceptual. Un concepto es lo que es en virtud tanto del contenido que tiene como de la forma que posee. El contenido da cuenta de su carácter intencional, y la forma explica sus rasgos psicológicos o la importancia cognitiva/epistémica que posee para el agente.

Lo anterior no constituye una semántica de dos factores en tanto el contenido es sólo la referencia para los conceptos primitivos, y la referencia de los primitivos que lo componen más su forma¹³¹ para los complejos. Tampoco es una mera cuestión terminológica respecto de qué es lo que merece llamarse contenido. Si queremos rescatar el carácter intencional de la mente tal como se presentó en el capítulo 1, entonces la semántica externista es necesaria. Si queremos salvar el carácter computacional de la mente, entonces los papeles inferenciales tienen que ser entendidos sintácticamente y no como parte del contenido.

Por último, se presentó una forma en que podría verse la relación entre contenido y forma que no resultara totalmente arbitraria. Asumiendo que algunos de nuestros mecanismos cognitivos han evolucionado, es posible pensar que tanto el contenido como la forma de algunos conceptos estén determinados de manera que nos proporcionen información respecto del ambiente que a su vez podamos utilizar de manera provechosa en nuestros procesos cognitivos. Si esto es así, y dado el principio de composicionalidad asumido, entonces resulta que si podemos distinguir tipos de inferencias en la forma de los conceptos. Sólo las inferencias que forman el concepto individualizan su forma, y explican el potencial inferencial del mismo, sin que tal potencial pase a formar parte de la individuación del concepto. Tales inferencias resultan ser analíticas, pues están ligadas al contenido del concepto complejo, pero lo hacen en virtud de la forma, tanto del complejo como de los primitivos

¹³¹ La forma de la representación compleja.

constituyentes. Por lo tanto la relación entre contenido y forma no sólo no es arbitraria para los conceptos primitivos, sino que ésta se traspasa vía composicionalidad a los conceptos complejos.

CONCLUSIÓN

Después de todo, si el cerebro fuera tan simple como para que una sola aproximación pudiera revelar sus secretos, entonces nosotros seríamos tan simples que no podríamos hacer el trabajo!¹³²

¹³² Tomado de A. Clark (2001), quien señala no recordar dónde ni de quién lo escucho.

A lo largo de esta tesis he explorado distintas posibilidades para formular una teoría de los conceptos dentro del marco hegemónico en ciencia cognitiva. Con base en las restricciones que impone dicho marco, en el capítulo 1 he considerado que, dados los supuestos computacional y representacionista, una semántica externista o referencialista es la mejor alternativa para individuar el contenido de los conceptos primitivos. Además, si se acepta que la noción de composicionalidad expuesta en ese capítulo debe ser respetada a fin de no violar las restricciones del marco de investigación basado en las hipótesis del sistema de símbolos físicos y del lenguaje del pensamiento, entonces resulta que tal semántica se aplica a todos los conceptos, pues los complejos no hacen más que heredar el contenido de sus constituyentes.

De acuerdo a esta postura respecto del contenido conceptual tenemos que los conceptos primitivos pueden ser o bien innatos o bien pueden formarse a partir de medios no conceptuales. Por ejemplo, pueden adquirir su contenido a través de mecanismos que no operan sobre representaciones conceptuales. Tales mecanismos pueden estar determinados genéticamente para responder selectivamente a ciertos elementos externos a la mente, y dar como resultado la adquisición de un concepto primitivo. Aún más, los papeles causales en los cuales participarían estos conceptos estarían también determinados por evolución, lo que explicaría su relación no-arbitraria con el contenido, la que a su vez se traspasa vía composicionalidad a los conceptos complejos que constituyen. Además, en el caso de un concepto innato, si lo hubiera, podría ser el caso que éste se interdefiniera con otros conceptos

innatos, y por lo tanto ninguno de ellos fuera primitivo, en el caso de que todos ellos estuvieran involucrados en el funcionamiento de un módulo cognitivo. Esto permite distinguir entre concepto primitivo y concepto innato.

Respecto de los conceptos complejos, su contenido estaría determinado por la referencia de sus partes más las inferencias que lo constituyen, esto es, las relaciones causales con los conceptos constituyentes. Tales relaciones causales, entendidas en este marco como computaciones constitutivas del concepto, terminan siendo un tipo de inferencia privilegiada, específicamente inferencias analíticas pues son verdaderas en virtud de su contenido¹³³. Por lo tanto, toda otra inferencia en la cual participe o pueda participar el concepto complejo así constituido no será constitutiva ni de su contenido ni de su forma.

En cuanto a la forma de los conceptos, la investigación aquí presentada me ha llevado a considerarla como relacionada con el papel causal de las representaciones en las computaciones que participan. Para el caso de los primitivos tal forma está determinada por los mecanismos y representaciones no-conceptuales, si es el caso, que permiten la adquisición del concepto, y es en virtud de ésta que la representación mental tiene los poderes causales que exhibe. En tanto estos mecanismos pueden ser innatos y haber sido seleccionados por evolución, y para el caso en que participaran representaciones mentales no-conceptuales éstas a su vez dependen del mismo tipo de mecanismos, entonces resulta que la relación entre la forma y el

¹³³ Así, estoy comprometido con la inferencia VINO BLANCO → BLANCO, pero no necesariamente con VINO BLANCO → LÍQUIDO. En el segundo caso la inferencia sería analítica si y sólo si LÍQUIDO es un constituyente de VINO. Si bien creo que resultaría razonable asumir tal composición del concepto VINO, en lo que se propone en esta tesis no hay un compromiso con el análisis de ningún concepto léxico en particular.

contenido de los primitivos es no arbitraria. Así, no es un misterio cómo el contenido del concepto OBJETO y las inferencias en las cuales participa tal concepto en virtud de su forma, sean tales que las computaciones en las que participa respecten ese contenido. Para el caso de las representaciones complejas, por su parte, tanto contenido como forma serán heredados de sus constituyentes, por lo que la relación de no-arbitrariedad se mantiene en su caso.

Con relación a las teorías de conceptos tradicionales, la presente propuesta puede ser vista como una versión de la teoría definicionista, en tanto asume que los conceptos complejos se individúan por sus relaciones analíticas con sus constituyentes. Sin embargo, respecto de los conceptos léxicos en particular, esta postura no está comprometida con la complejidad de todos o la mayoría de los conceptos léxicos, aunque parecería razonable asumir que para a gran mayoría de estos conceptos no existen mecanismos del tipo descrito para los primitivos.

Respecto de la evidencia psicológica revelada en las tareas de categorización, lo que se propone es ver en ella capacidades asociadas a la posesión y uso de conceptos en tareas cognitivas. En este sentido las tareas de categorización exhiben más bien características acerca de los procesos mentales, pero por lo mismo éstos requieren de la posesión de conceptos por parte de los sujetos para poder realizar las tareas. Como ya se dijo anteriormente, las computaciones (procesos cognitivos, inferencias) operan sobre las representaciones, y no las configuran.

De acuerdo a esto, la teoría propuesta no es una que asuma una semántica de aspecto dual, como ya se dijo, no tampoco puede ser considerada una versión híbrida como la de Osherson y Smith (1981), según la cual los conceptos tienen dos partes, un núcleo y un procedimiento de identificación. Esta es una manera de dar cabida a los procesos de categorización suponiendo que el núcleo consiste de una definición y que el procedimiento de identificación consiste en un prototipo¹³⁴. La presente propuesta no da cabida al procedimiento de identificación señalado por Osherson y Smith dentro de lo que constituye al concepto, excepto en el caso de conceptos primitivos donde tales procedimientos son responsabilidad de medios no-conceptuales.

Finalmente, un último tema a plantear antes de terminar la tesis se refiere a la relación entre los conceptos y lo que revelan las tareas de categorización. Para Machery (2009) esto es un indicador claro de que hay en juego dos proyectos que no parecen tener relación entre sí. Por un lado está la capacidad referencial de los conceptos, relativa a la individuación de su contenido en términos de evaluabilidad semántica, y por otra el funcionamiento de los conceptos en la economía cognitiva. Más que ver una disociación total entre ambos aspectos relacionados con la posesión de conceptos, mi postura

¹³⁴ Tal como lo exponen los autores: “Podemos distinguir entre el núcleo de un concepto y su procedimiento de identificación; el núcleo tiene que ver con aquellos aspectos de un concepto que explican su relación con otros conceptos, y con los pensamientos, mientras que el procedimiento de identificación especifica el tipo de información utilizada para tomar decisiones rápidas sobre la membresía ... Podemos ilustrar esto con el concepto MUJER. Su núcleo puede contener información sobre la presencia de un sistema reproductivo, mientras que su procedimiento de identificación puede contener información acerca de la forma del cuerpo, longitud del pelo y el tono de voz. Teniendo en cuenta esta distinción, es posible que algunas de las teorías tradicionales de conceptos caractericen correctamente el núcleo, mientras que la teoría de los prototipos caracteriza un importante procedimiento de identificación”. (1981: 57)

es que deben haber restricciones o constreñimientos que se siguen para cada proyecto de los resultados que consiga el otro¹³⁵. De hecho, Rosch (1978: 199-200) es más cuidadosa de lo que han sido muchos de los seguidores de sus trabajos, lo que ha dado lugar a confusiones que se deben aclarar. Ella señala que lo que presentado respecto de la importancia de los prototipos en las categorías del mundo real y los efectos de *tipicidad*, indica que los prototipos deben tener algún lugar en las teorías psicológicas de la representación, el procesamiento y el aprendizaje. Sin embargo, los prototipos mismos no constituyen ningún modelo particular de la representación, el procesamiento o el aprendizaje. Específicamente, hay cuatro temas que merecen aclararse:

1. Hablar de *un prototipo* es una ficción gramatical adoptada por razones de conveniencia. De lo que se habla en realidad en sus experimentos es de juicios acerca de los grados de *tipicidad*. En el caso de las categorías¹³⁶ del lenguaje natural hablar de una única entidad que es el prototipo es o bien un gran malentendido de los datos empíricos o bien de una teoría encubierta acerca de la representación mental.
2. Los prototipos no constituyen ningún modelo particular del procesamiento de las categorías. Un prototipo puede ser descrito como una lista de rasgos, una descripción estructural o un ejemplar. Por el sólo hecho de que los individuos sean capaces

¹³⁵ Una discusión más amplia de mi rechazo a la postura de Machery (2009) se encuentra en el apéndice de esta tesis.

¹³⁶ Recuérdese que en la terminología psicológica “categoría” y “concepto” suelen ser utilizados indistintamente.

de reconocer prototipos no se sigue que distintos procesos realizados sobre categorías deban responder a un único modelo de procesamiento. Lo que los hechos acerca de la *tipicidad* si contribuyen a una teoría del procesamiento son constreñimientos: los modelos de procesamiento no deberían ser inconsistentes con los hechos conocidos acerca de los prototipos.

3. Los prototipos no constituyen una teoría de la representación de categorías. Del mismo modo que los prototipos admiten distintas descripciones, también admiten distintas formas de representación, proposicional o visual, por ejemplo. Por lo mismo, nuevamente los prototipos sólo pueden ser considerados una fuente de constreñimientos sobre las teorías de la representación. En este sentido, la representación de una categoría por medio de la sola conjunción suficiente de atributos necesarios parecería en principio incapaz de dar lugar a los efectos del prototipo.
4. Si bien los prototipos deben ser aprendidos, ellos mismos no constituyen ninguna teoría particular del aprendizaje.

Las aclaraciones de Rosch respecto de lo que sí dicen sus experimentos sobre categorización me permiten articular una posible respuesta, mas no argumento, a la objeción que podría surgir a mi propuesta de lo presentado recientemente por Carey (2009) sobre la tesis de la continuidad. En cuanto mi propuesta dice que el contenido de la mayoría los conceptos debe pensarse de

manera externista, debido a la caracterización que he hecho de los primitivos más el principio de composicionalidad, si fuese posible que el sistema de símbolos adquiriera poderes representacionales más allá de su dotación primitiva, entonces mi propuesta falla en lo que a composicionalidad se refiere, y por tanto en lo relativo a la caracterización externista del contenido.

La idea de Carey es que hay un tipo de aprendizaje de conceptos que no responde al mero testeo de hipótesis basadas en la posesión de conceptos previos. El modelo de aprendizaje que tiene en mente lo denomina “bootstrapping quineano”, similar al proceso de formación de nuevos términos en las teorías científicas. El asunto relevante para mi propuesta es que si este fuera el caso, entonces al menos algunos conceptos complejos no heredarían su contenido de los conceptos involucrados en su adquisición. Este es justamente el argumento de Carey para sostener que tal proceso refuta la tesis de la continuidad.

Ahora bien, desde mi propuesta creo que tengo al menos dos estrategias para enfrentar el problema:

- a) A diferencia de Carey considero como conceptos primitivos aquellos cuyo contenido no está constituido a partir de otros conceptos, pero acepto que pueda estar constituido a partir de otro tipo de representaciones mentales. Para Carey, sin embargo, todo concepto primitivo es innato, en el sentido de que no se constituye a partir de ningún tipo de representación mental, sea esta conceptual o no. Así, en

mi caso podría estipular que si el contenido de un concepto no es rastreable a otros conceptos, entonces tal concepto debe ser considerado primitivo. En este caso habría que aceptar la posibilidad de primitivos que no fuesen innatos, en el sentido de que hayan mecanismos seleccionados por evolución encargados de su adquisición, y que puedan ser adquiridos (y no aprendidos) al momento de formular o aprender una teoría científica particular¹³⁷.

- b) Basándome en la distinción entre individuación de conceptos, por una parte, y los procesos cognitivos que operan sobre éstos (y que por lo tanto requieren de los conceptos como condición previa para su operación), es posible pensar que el proceso de *bootstrapping* que describe Carey sea acerca de cómo utilizamos los conceptos, por ejemplo al momento de formar o aprender teorías científicas. Tal como Rosch señala, la evidencia psicológica sobre cómo utilizamos los conceptos para realizar inferencias no necesariamente nos revela un modelo particular de representación, procesamiento o aprendizaje.

Además de las dos estrategias antes descritas, creo que el peso de la prueba recae sobre Carey, tanto en lo relativo a mostrar que su evidencia es a favor de un tipo de representación mental y modelo de aprendizaje, como respecto de modelar una implementación computacional de tales símbolos y procesos. Sobre el primer tema creo haber dado razones para creer que lo que

¹³⁷ Esto además requiere aceptar que hay inconmensurabilidad entre teorías científicas, que el conocimiento científico no es acumulativo, etc. Sólo quiero señalar aquí que tales nociones son aun fuente de debate en filosofía de la ciencia y que, más bien, pueden ser una fuente de más problemas antes que de soluciones.

dice Carey no afecta la propuesta de esta tesis. Sobre la implementación computacional de procesos inferenciales no-demostrativos, abductivos, o de *bootstrapping quineano*, tanto para distinguir cada uno de estos procesos, como para detallar en qué consisten, requeriría otra tesis.

Al final, si nuestro cerebro fuera tan simple como para que una sola aproximación pudiera revelar sus secretos, entonces nosotros seríamos tan simples que no podríamos hacer el trabajo.

APÉNDICE: MACHERY 2009

Recientemente Machery (2009) ha señalado que el concepto de “concepto” que manejan los psicólogos cognitivos podría no corresponder a una clase natural, en el sentido de ser explicativamente relevante para la investigación que ellos desarrollan. En una línea de argumentación similar a la Osherson y Smith (1981), pero tomando una postura radical acerca de la distinción entre lo que revelan los experimentos en tareas de categorización por una parte, y por otra la individuación de conceptos por su contenido, su propuesta consiste en disociar completamente ambas líneas de investigación.

La noción de “concepto” que Machery atribuye a los psicólogos cognitivos es la siguiente:

(C_p) Un concepto de x es un cuerpo de conocimiento acerca de x que está almacenado en la memoria de largo plazo y que es usado por *default* en los procesos que subyacen la mayoría, sino todas, las competencias cognitivas más superiores cuando éstos procesos tienen como resultados juicios acerca de x.

Por el contrario, la aproximación a los conceptos que Machery declara filosófica es:

(C_f) Tener un concepto de x es ser capaz de tener actitudes proposicionales acerca de x como x.

Para Machery la diferencia en la concepción de “concepto” marca divergencias sustanciales, debido a las distintas metas explicativas que de hecho tienen.

Sin embargo, el reconocer esta diferencia lleva a Machery, contrariamente a lo defendido a en esta exposición, a disociar completamente ambas líneas de investigación. Así, considera que los debates entre filósofos y psicólogos cognitivos son una especie de sin sentido. En tanto para los filósofos una teoría de conceptos es una teoría semántica para nuestras actitudes proposicionales, en cuanto explica cómo nuestros pensamientos pueden tener el contenido que tienen, las disputas con los psicólogos cognitivos no son reales debates en cuanto los últimos están interesados en cómo es que utilizamos esos conceptos. Según Machery, las críticas de autores como Margolis (1995) y Fodor (1994, 1998), en cuanto a que como teorías semánticas las teorías de conceptos basadas en los experimentos de categorización no son sostenibles, simplemente se equivocan pues tales teorías no intentan responder la pregunta filosófica acerca de los conceptos.

Si esto es así, entonces muchos argumentos esgrimidos desde la filosofía quedarían invalidados, como los de Rey (1983), Margolis (1994, 1995), Fodor (1994, 1998) y Laurence y Margolis (1999). Estas objeciones tendrían sentido sólo si los psicólogos intentaran responder la pregunta acerca de cómo podemos tener actitudes proposicionales acerca de los objetos de nuestras actitudes. Por ejemplo, la crítica a la teoría de prototipos de Margolis (1994) señala que la teoría de prototipos es inconsistente con las ideas de Kripke y

Putnam acerca de la naturaleza histórico-causal de la referencia. Como se dijo en el capítulo 1, para Kripke y Putnam un término puede referir aun cuando el hablante que usa ese término pueda tener unas pocas creencias verdaderas acerca del referente del término. La objeción de Margolis es que esta propiedad es inconsistente con la teoría de prototipos pues en ella se afirmarían que tener actitudes proposicionales acerca de x como tal es tener conocimiento acerca de cuáles propiedades son típicamente verdaderas de x .

Machery argumenta que la objeción de Margolis no es válida si uno reconoce que la teoría de prototipos, al igual que otras teorías psicológicas de los conceptos, no intenta explicar cómo es que podemos tener actitudes acerca de los objetos de nuestras actitudes. Del mismo modo, las críticas hechas a teorías filosóficas de los conceptos relativas a su incapacidad para explicar cómo razonamos, categorizamos, hacemos analogías o realizamos inducciones, como la de Prinz (2002) a Fodor, tampoco tienen sentido pues no es su meta proporcionar tales explicaciones.

Creo que lo mostrado en esta tesis puede ser visto como un reconocimiento de la diferencia importante entre lo que investigan los psicólogos cognitivos y la tarea de los filósofos en ciencia cognitiva, pero que sin embargo es capaz de reunir ambas tal y como el marco de investigación hegemónico había supuesto desde un principio: la filosofía se encarga de la semántica de los símbolos y la psicología cognitiva se encarga de describir los procesos que operan sobre las representaciones (y las ciencias de la computación se encargan de mostrar cómo esto es realizable en un dispositivo

físico). Ésta es una forma de lo que Machery llama la “concepción fundacional”, según la cual los psicólogos asumen que cuando razonamos, categorizamos, hacemos analogías, etc., acerca de algo, usamos algún tipo de conocimiento acerca de ese algo. Sin embargo los psicólogos no explican cómo es que podemos tener un conocimiento o cualquier tipo de actitud proposicional acerca de cualquier cosa. Esto sugeriría que la filosofía podría jugar un papel fundacional en la teoría de conceptos: el trabajo de los filósofos es explicar lo que los psicólogos asumen, esto es, que podemos tener actitudes proposicionales acerca de los objetos de nuestras actitudes.

Sin embargo, Machery ofrece un argumento en contra de esta concepción. Según él, la forma en que los filósofos sustentan sus afirmaciones acerca del contenido conceptual se basa en experimentos mentales que describen situaciones extrañas¹³⁸, y donde se testan nuestras intuiciones respecto de la atribución de estados mentales. La necesidad de los experimentos mentales se debe a que en situaciones de la vida real u ordinaria las teorías del contenido en competencia adscriben los mismos conceptos a los individuos. En este sentido el experimento mental juega el papel de experimento crucial para decidir entre teorías rivales.

El ejemplo que toma Machery es el experimento mental que propone Tyler Burge (1979), donde un individuo, Oscar, está convencido de que tiene artritis en el muslo. Lo que pide Burge del lector es que imagine una situación donde todo es casi idéntico a la situación de Oscar. En esta segunda situación,

¹³⁸ Desarrollo aquí el ejemplo de Burge (1979) y la crítica de Machery, pues esta objeción se dirige directamente a uno de los argumentos que utilicé en el capítulo 1 para justificar la adopción de una teoría externista respecto del contenido conceptual.

sin embargo, hay una diferencia: en la comunidad lingüística el término “artritis” es usado para referir tanto a dolores en las articulaciones como a dolores en el muslo. Lo que se persigue es que las intuiciones del lector lo lleven a pensar que en el primer caso la creencia de Oscar es falsa, y que en la segunda situación Oscar no tiene creencias cuyo objeto sea la artritis. Así, sólo en la primera situación Oscar tiene el concepto ARTRITIS, aun cuando la única diferencia entre las dos situaciones radica en el lenguaje que habla la comunidad lingüística de Oscar. Burge concluye que la capacidad para tener actitudes proposicionales acerca de los objetos de nuestras actitudes, en este caso superviene sobre hechos sociales, y no sobre las creencias que pueda tener Oscar, pues por el diseño del experimento el individuo Oscar tiene las mismas creencias acerca de su concepto ARTRITIS en ambas situaciones, y aun así su concepto es distinto pues el objeto o propiedad al que refiere es distinto.

La crítica de Machery a esta posición es que la metodología utilizada asume que la adscripción de actitudes proposicionales es uniforme. Argumenta que si hubiera alguna variación en las adscripciones de actitudes proposicionales que realizan los individuos y si no hubiera una razón para preferir un tipo de adscripción sobre la otra, sería poco claro qué hacer al construir una teoría de las representaciones mentales. Suponiendo que distintos lectores pudieran estar en desacuerdo respecto de si atribuir o no a Oscar creencias sobre la artritis en la segunda situación, entonces no estaría claro que la posesión de conceptos realmente superviene sobre hechos sociales. Pues bien, según Machery no se ha determinado aún si existe tal

variación, y en tanto no sea empíricamente establecido que no la hay o que no resulta relevante, entonces es por lo menos cuestionable que la tarea de construir una teoría de los conceptos filosófica pueda ser completada por los métodos filosóficos tradicionales.

Según mi parecer la objeción de Machery descansa sobre supuestos al menos cuestionables. En primer lugar, la idea de “establecer empíricamente” variaciones en las intuiciones de los hablantes recurriendo a lo que hoy se conoce como filosofía experimental no es algo cuyo consenso en la comunidad filosófica esté asegurado, por decir lo menos. Además, apela a una forma de entender el trabajo filosófico que no agota todas las posibilidades metodológicas. Aún más, como el mismo lo señala, si rechazáramos la “metodología favorita de los filósofos” esto no implicaría el rechazo de la concepción fundacionalista antes aludida. Si existieran otros métodos para estudiar la relación entre los pensamientos y sus objetos la implementación de la concepción fundacionalista sería posible, pero no existiría en la actualidad un método alternativo reconocido, según Machery. Sin embargo, creo que los desarrollos de semánticas informacionales vinculadas a la ciencia cognitiva, tal y como se han presentado en esta investigación, si constituyen un método alternativo, y justamente este trabajo intenta mostrar un forma en tal trabajo puede ser completado, o al menos llevado a cabo sin tener que recurrir a la metodología tradicional señalada por Machery¹³⁹.

¹³⁹ Como lo he señalado a lo largo de toda la tesis, no es mi intención justificar afirmaciones acerca de la relación entre lenguaje y pensamiento, sino de la relación entre pensamiento y mundo, por lo que este trabajo, a lo sumo, podría servir como fuente de posibilidades para una justificación de los argumentos esgrimidos por Kripke, Putnam y Burge a favor de una determinación externista del contenido de las oraciones de (atribución de) actitudes proposicionales.

REFERENCIAS

- Armstrong, S. L., Gleitman, L. R., and Gleitman, H. (1983) "What some concepts might not be". *Cognition* 13:263–308.
- Ayede, M. (1998) "Fodor on Concepts and Frege Puzzles". *Pacific Philosophical Quarterly* 79, pp.289-294.
- Barsalou, L.W. (1993) "Flexibility, structure and linguistic vagary in concepts: manifestations of a compositional system of perceptual symbols." En Collins, A.C., Gathercole, S.E. & Conway, M.A. eds. (1993) *Theories of Memory*. 29-101. London: Lawrence Erlbaum Associates.
- Barsalou, L.W. (1999) "Perceptual Symbol Systems." *Behavioral and Brain Sciences* 22: 577-660.
- Barsalou, L. W., Simmons, W. K., Barbey, A., and Wilson, C. D. (2003) "Grounding conceptual knowledge in modality-specific systems". *Trends in Cognitive Sciences* 7:84–91.
- Block, N. (1978) "Troubles with Functionalism". *Minnesota Studies in the Philosophy of Science* 9:261-325.
- Bloom, P. (2000) "Object names and Other Common Nouns". En Bloom, P. (2002) *How Children Learn The Meanings of Words*. Cambridge, Mass.: MIT Press, 89-120.
- Boccardi, E. (2009) "Who's driving the syntactic engine". *Journal for General Philosophy of Science* 40 (1): 23-50.
- Bourne, L. (1970). Knowing and using concepts. *Psychological Review*, 77, 6, 546-556.
- Brooks, R. A. [1991] (1999). "Intelligence without representation". *Artificial Intelligence* 47:139–159.
- Burge, T. (1979) "Individualism and the Mental". *Midwest Studies in Philosophy*, IV, 73-121.
- Burge, T. (1986) "Individualism and the Psychology". *Philosophical Review* 95, 3-45.
- Carey, S. (1985). *Conceptual Change in Childhood*. MIT Press.
- Carey, S. (2009) *The Origins of Concepts*. Oxford University Press.
- Carey, S. E., and Spelke, E. S. (1994) "Domain specific knowledge and conceptual Change". En *Mapping the mind: Domain specificity in cognition and culture*, ed. L. A. Hirschfeld and S. A. Gelman, 169–200. Cambridge: Cambridge University Press.
- Carston, R. (2008) "Linguistic communication and the semantics/pragmatics distinction" *Synthese* 165: 321-345.
- Chomsky, N. (1959) "Verbal behaviour by B-F- Skinner" (Review). *Language*, Vol. 35, N°1, pp.26-58.
- Chomsky, N. (1995) "Language and nature". *Mind* 104:1–61.
- Churchland, P. M. (1981) "Eliminative materialism and the propositional attitudes". *Journal of Philosophy* 78:67–90.
- Clark, A. (1993) *Associative engines: Connectionism, concepts, and representational change*. Cambridge, MA: MIT Press.

- Clark, A. (2001) *Mindware: an introduction to the philosophy of cognitive science*. New York: Oxford University Press.
- Clark, A. & Chalmers, D. (1998) "The Extended Mind". *Analysis* Vol.58, N°1:7-19.
- Cohen, Jonathan (2004) "Objects, places and perception." *Philosophical Psychology*, Vol. 17, N°4, 2004. pp. 471-495.
- Corter, J.E. & Gluck, M.A. (1992). Explaining basic categories: Feature predictability and information. *Psychological Bulletin*, 111, 291-303.
- Cowie, F. (1999) *What's Within: Nativism Reconsidered*. Oxford: Oxford University Press.
- Cummins, R. (1989) *Meaning and Mental Representation*. Cambridge, MA: MIT Press.
- Dretske, F. (1981) *Knowledge and the Flow of Information*. Cambridge, MA: MIT Press.
- Evans, G. (1982) *Varieties of Reference*. Oxford: Clarendon Press.
- Ezcurdia, M. (1995) "Modos de presentación y modos de determinación". *Crítica* 80: 57-96.
- Ezcurdia, M. (2001) "Thinking about myself". En Brook, A. & DeVidi, R.C. (Comps.) (2001) *Self-reference and Self-awareness*, pp. 179-203. John Benjamins Publisher, Philadelphia.
- Fine, K. (2007) *Semantic Relationalism*. Blackwell.
- Fodor, J.A. (1975) *The Language of Thought*. Cambridge: Harvard University Press.
- Fodor, J.A. (1987) *Psychosemantics: the Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT.
- Fodor, J.A. (1990c) A Theory of Content II: The Theory. En *A Theory of Content and Other Essays*. Cambridge, MA: MIT Press.
- Fodor, J.A. (1994) "Concepts. A Potboiler". *Cognition* 50: 95-113.
- Fodor, J.A. (1998) *Concepts: Where Cognitive Science Were Wrong*. Oxford University Press.
- Fodor, J.A. (2001a) "Doing Without What's Within: Fiona Cowie's Critique of Nativism. *Mind* 110, N° 437, pp. 99-148.
- Fodor, J.A. (2001b). "Language, Thought, and Compositionality," *Mind and Language*, 16: 1-15.
- Fodor, J.A. (2003). *Hume Variations*. Oxford: Clarendon Press.
- Fodor, J.A. (2004) "Having Concepts: A Brief Refutation of the Twentieth Century". *Mind and Language* 19, pp.29-47.
- Fodor, J.A. (2008) *LOT 2: THE LANGUAGE OF THOUGHT REVISITED*. Oxford: Clarendon Press.
- Fodor, J.A. & Lepore, E. (1991) "Why Meaning (Probably) Isn't Conceptual Role". En Fodor, J.A. And Lepore, E. (2002) *The Compositionality Papers*. Oxford: Clarendon Press.
- Fodor, J.A. & Lepore, E. (1996) "The Pet Fish and the Red Herring: Why Concepts Aren't Prototypes". En Fodor, J. And Lepore, E. (2002) *The Compositionality Papers*. Oxford: Clarendon Press.
- Fodor, J.A. & Lepore, E. (1999) "All at Sea in Semantic Space: Churchland on Meaning Similarity". En Fodor, J. And Lepore, E. (2002) *The Compositionality Papers*. Oxford: Clarendon Press.

- Fodor, J.A. & Lepore, E. (2002) *The Compositionality Papers*. Oxford: Clarendon Press.
- Fodor, J.A. & Pylyshyn, Z. (1998) "Connectionism and Cognitive Architecture: a critical analysis". *Cognition* 28:3-71.
- Frege, G. (1879/1972) *CONCEPTOGRAFÍA Los fundamentos de la aritmética*. Padilla, H. (Trad.). Ciudad de México: Universidad Nacional Autónoma de México, Instituto de Investigaciones Filosóficas.
- Frege, G. (1892) "Sobre Sentido y Denotación". En Frege, G. 1972, *Lógica y Semántica*. Introducción, traducción y selección bibliográfica de Alfonso Gómez-Lobo. Valparaíso: Ediciones Universitarias de Valparaíso.
- Gelman, S. A. (1988) "The development of induction within natural kinds and artifacts categories". *Cognitive Psychology* 20:65–95.
- Gelman, S. A., and Markman, E. (1986) "Categories and induction in young children". *Cognition* 23:183–209.
- Gopnik, A., and Meltzoff, A. N. (1997) *Words, thoughts, and theories*. Cambridge, MA: MIT Press.
- Hampton, J.A. (1982) "A Demonstration of intransitivity in natural categories". *Cognition* 12, 151-164.
- Haugeland, J. (1981) "Semantic engines: an introduction to mind design". En J. Haugeland (ed.) *Mind Desing*. Cambridge MA: MIT Press.
- Heit, E. (2001). Background knowledge and models of categorization. En *Similarity and Categorization* (eds. Hahn & Ramscar), 155-178. Oxford University Press.
- Hull, C.L. (1920) "Quantitative aspects of evolution of concepts". *Psychological Monographs*, XXVIII.
- Hume, D. [1748] (1975) *Enquiries concerning human understanding and concerning the principles of morals*. Ed. P. H. Nidditch. Oxford: Oxford University Press.
- Jackendoff, R. (1989). What Is a Concept, That a person May Grasp It? *Mind and Language*, 4, 68-102.
- Kant, I. (1781/1994) *Crítica de la Razón Pura*. Prólogo, traducción, notas e índices Pedro Ribas (3ª Ed.). Madrid: Alfabeta.
- Kim, J. (1992) "Multiple realization and the metaphysics of reduction". *Philosophy and Phenomenological Research* 52:1–26.
- Kripke, S. (1972) *Naming and Necessity*. Cambridge, MA: MIT PRESS.
- Locke, J. [1690] (1979) *An essay concerning human understanding*. Ed. P. H. Nidditch. Oxford: Oxford University Press.
- Laurence, S. & Margolis, E. (1999) "Concepts and cognitive science". En Margolis & Laurence eds. (1999) *CONCEPTS: core readings*. Cambridge, MA: MIT PRESS.
- Machery, E. (2009) *Doing without Concepts*. Oxford: Oxford University Press.
- Margolis, E. (1994) "A reassessment of the shift from the classical theory of concepts to prototype theory". *Cognition* 51:73-89.
- Margolis, E. (1995) "The significance of the theory analogy in the psychological study of concepts." *Mind and Language* 10: 45-71.
- Margolis, E. y Laurence, S. eds. (1999). *CONCEPTS: Core Readings*. Cambridge, MA: MIT.

- Margolis, E. y Laurence, S. (2004) "Concepts". En T.A. Warfield, T.A. & Stich, S.P. eds. (2004) *The Blackwell guide to the philosophy of mind*. 190-213. Oxford: Blackwell.
- Markman, A. B. & Gentner, D. (1993). Structural alignment during similarity comparisons. *Cognitive Psychology*, 25, 431- 467.
- Medin, D.L. (1989) "Concepts and conceptual structure". *American Psychologist* 44:1469-1481.
- Mervis, C. B., Catlin, J., and Rosch, E. (1976). "Relationships among goodness-of-example, category norms, and word frequency". *Bulletin of the Psychonomic Society*, 7, 283–284.
- Millikan, R. (1984) *Language, Thought and Other Biological Categories*. Cambridge, MA: MIT Press.
- Millikan, R. (1993) *White Queen Psychology and Other Essays for Alice*. Cambridge, MA: MIT Press.
- Millikan, R. (2000) "A Common Structure for Concepts of Individuals, Stuffs, and Basic Kinds: More Mama, More Milk and More Mouse. *Behavioral and Brain Sciences* 22, pp.56-65.
- Murphy, G. L. (2002). *The big book of concepts*. MIT Press, Cambridge, MA.
- Murphy, G. & Medin, D. (1985). *The role of theories in conceptual coherence*. *Psychological Review*, 92. 289-316.
- Neander, K. (2006) "Content for Cognitive Science". En McDonald, G. & Papineau, D. (2006) *Teleosemantics*. Oxford: Oxford University Press.
- Newell, A. y Simon, H. (1976) "Computer science as an empirical inquiry: Symbols and search". *Communications of the Association for Computing Machinery*, 19, 113-126. Reimpreso en J. Haugeland (ed.) *Mind Desing*. Cambridge MA: MIT Press.
- Osherson, D.N. & Smith E.E. (1981) "On the adequacy of prototype theory as a theory of concepts". *Cognition* 9:35–58.
- Peacocke, C. (1992) *A study of Concepts*. Cambridge, MA: MIT Press.
- Perry, J. (1980) "A problem about continued belief". En J. Perry (1993) *The Problem of Essential Indexicals and other Essays*. Oxford: Oxford University Press.
- Pietroski, P.M. (2005) *Events and Semantic Architecture*. Oxford: Oxford University Press.
- Pinker, S. (1997) *How the Mind Works*. New York: W.W. Norton & Company.
- Pinker, S. & Prince, A. (1999) "The nature of human concepts: evidence from an unusual source". En Jackendoff, R., Bloom, P. & Wynn, K. eds. (1999) *Language, Logic and Concepts*. 221-261. Cambridge, MA:MIT PRESS.
- Posner, M.I. & Keele, S.W. (1968) "On the genesis of abstract ideas". *Journal of Experimental Psychology*, 77, 353-363.
- Posner, M.I. & Keele, S.W. (1970) "Retention of abstract ideas". *Journal of Experimental Psychology*, 83, 304-308.
- Prinz, J.J. (2002). *Furnishing the Mind: Concepts and Their Perceptual Basis*. Cambridge, MA:MIT Press.
- Putnam, H. (1975) The meaning of "meaning". En H. Putnam, *Philosophical Papers*, vol. 2: *Mind, Language, and Reality*. Cambridge: Cambridge University Press.

- Pylyshyn, Z. (2000) "Situating vision in the world". *Trends in Cognitive Science*, 4(5), 2000, pp. 197-207.
- Quine, W.V.O. (1951) "Two dogmas of empiricism". *Philosophical Review* 60:20-43.
- Quine, W.V.O. (1960) *Word and Object*. Cambridge, MA: MIT PRESS.
- Reed, S. K. (1972). Pattern recognition and categorization. *Cognitive Psychology*, 3, 382-407.
- Rey, G. (1983) "Concepts and Stereotypes". *Cognition* 15.
- Recanati, F. (2003) "Indexical Concepts and Compositionality". En García-Carpintero, M. & Macia, J. (Eds.), (2003), *Two-Dimensional Semantics*. Oxford: Clarendon Press, pp. 249-257.
- Rodríguez, M. (2006) *SOBRE CONCEPTOS: ATOMISMO INFORMACIONAL*. Tesis para optar al grado de Magíster en Filosofía, mención Epistemología. Profesor Patrocinante: Guido Vallejos O. Universidad de Chile, Facultad de Filosofía y Humanidades, Escuela de Postgrado. Santiago, Chile.
- Rodríguez, M. (2007) "Sobre Conceptos Primitivos. Atomismo Informacional". *Summa Psicológica UST*, vol. 4, n° 1, pp. 31-45.
- Rosch, E. (1973). "On the internal structure of perceptual and semantic categories". In T. E. Moore (Ed.), *Cognitive Development and the Acquisition of Language* (pp. 111–144). New York: Academic Press.
- Rosch, E. (1978). Principles of categorization. En E. Rosch y B. Lloyd (Eds.) *Cognition and Categorization*. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Rosch, E. & Mervis, C. (1975). Family resemblances: studies in the internal structure of categories. *Cognitive Psychology*, 7, 573-605.
- Rosch, E., Mervis, C. B., Gray, W., Johnson, D., & Boyes-Braem, P. (1976). Basic objects in natural categories. *Cognitive Psychology*, 8, 382-439.
- Segal, G. (1989) "Seeing what is not there". *The Philosophical Review*, Vol. 98 N° 2, 189-214.
- Segal, G. (2000) *A Slim Book About Narrow Content*. Cambridge, MA: MIT Press.
- Smith, E. & Medin, D. (1981). *Categories and concepts*. Cambridge, MA: Harvard University Press.
- Spelke, E., "Object Perception", in Goldman, A.I. (1993), *Readings in Philosophy and Cognitive Science*, Cambridge, Mass.: MIT Press, 447-460.
- Solomon, K.O., Medin, D.L. & Lynch, E.L. (1999) "Concepts do more than categorize". *Trends in Cognitive Science* 3:99-105.
- Sterlny, K. (1990) *The Representational Theory of Mind: An Introduction*. Cambridge, MA: Blackwell.
- Szabó, Z.G. (2002) *Problems of Compositionality*. New York, Garland.
- Thelen, E., and Smith, L. (1994) *A dynamic systems approach to the development of cognition and action*. Cambridge, MA: MIT Press.
- Treisman, A. & Schmidt, H. (1982) "Illusory Conjunction in the Perception of Objects". *Cognitive Psychology*, Vol 14 N° 1: 107-41.
- Tversky, A. (1977) Features of similarity. *Psychological Review*, 84, 327-352.
- Vallejos, G. (2008) *Conceptos y Ciencia Cognitiva*. Santiago de Chile, Bravo y Allende.

- Van Gelder, T.J. (1990) "Compositionality: A connectionist variation on a classical theme." *Cognitive Science* 14: 355-384.
- Van Gelder, T.J. (1995) "What might cognition be, if not computation?" *Journal of Philosophy* 91:345-381.
- Wisniewski, E. J. & Medin, D. L. (1994). On the interaction of theory and data in concept learning. *Cognitive Science*, 18, 221- 281.
- Wittgenstein, L. (1953) *Philosophical Investigations*. Ed. G.E.M. Anscombe. Oxford: Blackwell.