Máquinas y lenguaje
¿Puede una máquina comprender lo que otro dice? Una respuesta desde Wittgenstein

Dennis Yadira Jaimes Jiménez

Trabajo de grado para optar el título de Magíster en filosofía

Director
Dairon Alfonso Rodríguez Ramírez
PhD. Humanidades

Universidad Industrial de Santander
Facultad de Ciencias Humanas
Escuela de filosofía
Maestría en filosofía
Bucaramanga
2021

Contenido

Introducción 6

1. Humanos que *siguen reglas* y máquinas que "siguen reglas" 11

1.1 El algoritmo y la máquina 19

1.1.1 Concepto general de 'algoritmo' 19

1.1.2 Máquinas que ejecutan algoritmos 21

1.2 Humanos que siguen reglas mecánicas y máquinas que siguen reglas mecánicamente. 25

1.2.1 Máquinas que siguen reglas 26

1.2.2 Humanos que siguen reglas 31

2. Comprender lo que otro quiere decir 34

2.1 Comprender una oración es 'saber lo que la oración dice' 39

2.1.1 Observaciones sobre significado como esencia/idea general 39

2.1.2 Observaciones al significado como 'imagen mental' 45

2.1.3 Observaciones a la comprensión entendida como proceso mental o ¿Describir un proceso mental es igual que describir mi habitación? 49

2.1.4 Observaciones a la tesis de la comprensión como 'estado mental' 55

2.2 Ser capaz o tener la habilidad 58

2.3 Formas de vida («Lebensform», «Lebensformen», «Form des Lebens») 64

2.3.1 Formas de vida, comprensión profunda y representación perspicua 65

3. Conclusiones 68

Referencias Bibliográficas 71

Lista de Figuras

Figura 1 *Esquema de Organización de una Máquina Capaz de Ejecutar Algoritmos*……21

Figura 2 *Esquema de Organización de una Máquina Computadora que Ejecutar Algoritmos*……22

Resumen

Título: Máquinas y lenguaje ¿puede una máquina comprender lo que otro dice? Una respuesta desde Wittgenstein[*]

Autor: Dennis Yadira Jaimes Jiménez[**]

Palabras clave: Wittgenstein, Lenguaje, Inteligencia Artificial, Algoritmo, Comprensión

Descripción:

Esta investigación responde a la cuestión: ¿un sistema artificial puede ser usuario del lenguaje natural? Las observaciones que realizamos una indagación gramatical que busca aclarar el malentendido sobre el enunciado: «–El sistema artificial inteligente *x comprende* lo que enuncia y *comprende* lo que otro ha dicho». Para alcanzar este objetivo revisamos dos presupuestos conceptuales en los que se apoyan los estudios del desarrollo del lenguaje natural en los sistemas artificiales inteligentes. Por un lado, se analiza la analogía entre el comportamiento de la mente/cerebro y los procesos algorítmicos computacionales que presenta el modelo referencial-computacional de la mente (MRCM). En específico, la noción de 'algoritmo' como proceso causal estricto y mecánico, no con el fin de demostrar o refutar que nuestros estados mentales funcionen como procesos algorítmicos, dado que nuestra investigación no es científica; más bien revisamos nuestro uso de los algoritmos en varios contextos. Por otro lado, en vista de que la comprensión es el criterio para atribuir el uso del lenguaje natural a un agente, se analiza el uso nos aproximamos al uso gramatical de 'comprensión'. A partir de esto proponemos mostrar el error de la búsqueda de un estado o proceso –sea mental o cerebral– que corresponda a 'comprender', más bien se plantea entender la comprensión como capacidad.

[*] Trabajo de grado.

[**] Facultad de Ciencias humanas. Escuela de filosofía. Director PhD. Dairon Alfonso Rodríguez Ramírez.

Abstract

Title: Machines and language. Can a machine understand natural language? An answer from Wittgenstein[*]

Author: Dennis Yadira Jaimes Jiménez[**]

Key Words: Wittgenstein, Language, Artificial Intelligence, Algorithm, Understanding

Description:

This research answers the question: can an artificial system be a user of natural language? The observations that we make a grammatical inquiry that seeks to clarify the misunderstanding about the statement: "–The intelligent artificial system x understands what it states and understands what another has said." To achieve this objective, we review two conceptual assumptions on which the studies of the development of natural language in intelligent artificial systems are based. On the one hand, the analogy between the behavior of the mind / brain and the computational algorithmic processes presented by the referential-computational model of the mind (MRCM) is analyzed. Specifically, the notion of 'algorithm' as a strict and mechanical causal process, not in order to prove or disprove that our mental states function as algorithmic processes, since our research is not scientific; rather we review our use of the algorithms in various contexts. On the other hand, in view of the fact that comprehension is the criterion for attributing the use of natural language to an agent, the use is analyzed and we approximate the grammatical use of 'comprehension'. From this we propose to show the error of the search for a state or process –be it mental or cerebral– that corresponds to 'understanding', rather it is proposed to understand understanding as capacity.

[*] Bachelor Thesis

[**]Facultad de Ciencias humanas. Escuela de filosofía. Director PhD. Dairon Alfonso Rodríguez Ramírez.

Introducción

Ray Kurzweil[1] relata en una entrevista de 2015 que en Google están creando un modelo capaz de usar el lenguaje igual que los humanos, esto es, no sólo para leer documentos, sino para comprenderlos. Un sistema cualificado para leer miles de millones de páginas web creando una Base de Conocimiento[2] acerca del mundo, a la que podremos dirigirnos con preguntas manteniendo un diálogo inteligente:

> El mensaje de cualquiera de nuestros artículos es información. Queremos que el ordenador lea todo lo que hay disponible en la Web, y todas las páginas de cada uno de los libros existentes, para llegar a ser capaz de tener un diálogo inteligente con el usuario, tanto como para ser capaz de responder sus preguntas… «Google conocerá la respuesta a tu pregunta aun antes de habérsela planteado… Puede llegar a conocernos mejor que nuestra compañera; incluso, mejor que nosotros mismos.»
> « ¿Y qué es más difícil de simular en el cerebro humano?» le preguntaron. «Pienso en la ironía o en el sarcasmo [o en la polisemia en general, que permite los 'juegos de palabras']. Porque implican varios niveles de interpretación. Y esto confunde a las computadoras. Lo que hace el cerebro humano es crear una jerarquía, respecto de su propio pensamiento: ahí está el secreto de cómo pensamos.» (Kurzweil, citado por Garrido, 2017, p. 152)

Las ideas de Kurzwell no son gratuitas desde los primeros años de investigación en el campo de la Inteligencia Artificial (IA) se ha buscado desarrollar sistemas artificiales inteligentes que de alguna manera usen el lenguaje como lo conocemos; por ejemplo, Terry Winograd

[1] Ray Kurzweil (1948-) defensor de la Singularidad, creador del primer lector electrónico para ciegos, del primer escáner digital y del sistema de reconocimiento de voz que dio lugar a SIRI, entre otros. Actualmente trabaja como Director de investigación de Google; sus obras más reconocidas son: *The Singularity is Near* (2005) y *How Create a Mind?* (2012).

[2] BC o KB: Las bases de conocimiento son la evolución lógica de los sistemas de bases de datos tradicionales, en un intento de plasmar no ya cantidades ingentes de datos, sino elementos de conocimiento (normalmente en forma de hechos y reglas) así como la manera en que estos elementos han de ser utilizados. También se les trata de dotar de conocimiento sobre sí mismas, es decir, una KB ha de "saber lo que sabe". Por ejemplo, ante una consulta determinada, una base de datos accederá a los datos almacenados en ella y dará una respuesta afirmativa o negativa, independientemente de que tenga o no la información relevante; en cambio, una KB respondería "sí", "no" o "no lo sé", en el caso de carecer de todos los datos pertinentes a la consulta. Concepto tomado de: Pérez Hernández, (2002).

desarrolló SHRDLU[3], un programa que "comprendía" el lenguaje natural[4]. SHRDLU recibía instrucciones en inglés para que un robot las ejecutara, además podía sostener conversaciones medianamente fluidas incluyendo frases que utilizan nombres compuestos y pronombres ambiguos, siempre y cuando le hablaran de bloques encima de una tabla particular. Sumado a esto, los investigadores podían añadir más vocabulario en cualquier momento a SHRDLU siempre que se indicara el significado (la referencia) de los nuevos términos. Sin embargo, SHRDLU funcionaba como un micromundo, por esto se enfrentó a una serie de reparos que desbordaban su capacidad. Haugeland (2001) nos advierte de tres: (a) no posee ninguna iniciativa motivada o deliberada, (b) no tiene un sentido de continuidad en la conversación y (c) los problemas difíciles de la percepción incluyen una relación entre símbolos y objetos no-simbólicos, pero SHRDLU por completo resulta ser simbólico. La creación de éste programa en los años 70's se basó en la construcción sintáctica apoyada por la IA Simbólica que dejo de lado la característica contextual de la semántica, lo que hace que SHRDLU carezca de sentido de la situación. Por lo tanto, la aplicación de SHRDLU en un contexto no fue posible.

Los problemas que presentaron los sistemas construidos con base en la sintaxis cambiaron la perspectiva de los estudios en diferentes ramas de la IA, por ejemplo, en el Procesamiento de Lenguaje Natural (PLN). Margaret Masternam[5] en Cambridge volcó su indagación al desarrollo del contexto gramatical del lenguaje, lo que dio lugar a la traducción automática y la búsqueda y recuperación de información. Sobre la traducción automática, el grupo de investigación encabezado por Masternam no llevó a cabo la traducción de equivalencia palabra por palabra a modo de diccionario, más bien propuso buscar relaciones entre grupos de palabras dentro del mismo texto en cuestión, lo que permite solucionar la traducción de palabras ambiguas. Por ejemplo, 'banco' podría traducirse como 'bench' o 'bank', dependiendo de las otras palabras usadas en el texto ya sea 'parque' o 'dinero', respectivamente. El interés por el enfoque contextual provoca la creación de más relaciones entre palabras, no sólo la sinonimia, también la antonimia, elementos de clases, nivel de clase compartidos, partes de un conjunto, concurrencia temática, etc. En la actualidad el PLN

[3] Véase: Winograd, (1972).
[4] Véase: Winograd, (1980).
[5] El trabajo de Margaret Masterman se puede leer en: Masterman, (2005).

funciona a partir de vectores de palabras que representan nubes de significado probabilístico que participan en un concepto dado, de manera que "El procesamiento de lenguajes naturales depende más de la fuerza (potencia computacional) que de la maña (análisis gramatical). Las matemáticas –específicamente la estadística– han superado a la lógica y el aprendizaje automático ha desplazado al análisis sintáctico" (Boden, 2017, p. 67).

Otros sistemas usan la búsqueda ponderada –iniciada también por Masterman en 1976– o la minería de datos para desarrollar PLN, a esto se añade un sistema basado en reglas y se obtienen sistemas híbridos que pueden identificar patrones, a veces hasta insospechados para el ser humano, ejemplo de esto es Siri de Apple. Siri es un asistente personal que puede responder rápidamente a preguntas diferentes usando redes de información en internet (Wikipedia, Google Maps, The New York Times, entre otras), además usa Wolfram Alpha (buscador de respuestas que utiliza el razonamiento lógico) para averiguar la respuesta a un amplio rango de preguntas fácticas. Adicionalmente Siri se adapta a la voz y dialecto, los intereses y las preferencias del usuario individual; pese a todo esto es fácil desviar la conversación y hacer parecer fuera de lugar las respuestas de Siri.

En la actualidad, Ben Goertzel y David Hanson líderes en la creación del software incorporado al robot *Sophia,* presentan un sistema artificial inteligente que no solo se preocupa por responder a una conversación sencilla, sino por atender al problema de la corporeidad olvidado por la IA clásica. *Sophia* cuenta con procesamiento de datos visuales y reconocimiento facial, con el fin de imitar gestos y expresiones faciales. De este modo se vincula la IA y la robótica con el fin de crear robots cada vez más humanoides que reconozcan patrones de comportamiento, además de aprender de estos sobre la marcha. Todo esto alimenta la aspiración de crear un sistema artificial inteligente que pueda ser usuario del lenguaje natural como lo plantea Ray Kurzweil, en otras palabras, que tenga sentido expresar: «–El sistema artificial inteligente *x* comprende lo que enuncia y comprende lo que otro ha dicho».

Ahora bien, las observaciones que realizamos en este texto no se detienen en el desarrollo técnico de la IA o la exposición detallada del funcionamiento de los sistemas artificiales inteligentes que se han construido con el propósito de desarrollar lenguaje natural. Tampoco, se preocupan por desentrañar fenómenos tales como la conciencia, la

intencionalidad o la inteligencia general[6] que puedan explicar cómo *funciona* la mente o el cerebro del ser humano cuando usa el lenguaje. En cambio, a la luz de las obras de madurez de Wittgenstein[7] desarrollamos una indagación gramatical que busca aclarar el malentendido que sugiere el enunciado: «–El sistema artificial inteligente *x comprende* lo que enuncia y *comprende* lo que otro ha dicho». Para alcanzar este objetivo revisamos dos presupuestos conceptuales en los que se apoyan los estudios del desarrollo del lenguaje natural en los sistemas artificiales inteligentes. Veamos:

i. La equivalencia entre el funcionamiento de la mente/cerebro y los procesos algorítmicos computacionales.
ii. La comprensión es algún tipo de estado o proceso mental/cerebral que puede ser descrito, y, por tanto, con posibilidad de ser replicado.

En este orden de ideas, en el primer capítulo reflexionamos respecto a la analogía entre el comportamiento de la mente/cerebro y los procesos algorítmicos computacionales que presenta el modelo referencial-computacional de la mente (MRCM). En especial prestamos atención a la noción de 'algoritmo' como proceso causal estricto y mecánico, no con el fin de demostrar o refutar que nuestros estados mentales funcionen como procesos algorítmicos, dado que nuestra investigación no es científica; más bien revisamos nuestro uso de los algoritmos en varios contextos. Sobre esto podemos adelantarnos a decir que: sin importar que el ser humano siga el algoritmo más simplísimo para resolver una tarea no se reduce a un proceso causal mecánico como sucede en una máquina, sino que acaece necesariamente en un contexto normativo.

En el segundo capítulo, en vista de que la comprensión es el criterio para atribuir el uso del lenguaje natural a un agente, nos aproximamos al uso gramatical de 'comprensión'. A partir de esto proponemos mostrar el error de la búsqueda de un estado o proceso –sea mental o cerebral– que corresponda a 'comprender', más bien se plantea entender la comprensión como capacidad. Todo esto bajo la perspectiva de una investigación gramatical que rechaza

[6] Sobre el concepto de 'Inteligencia general' puede se puede leer: Boden, 2017, pp. 29-61.
[7] En especial: *Escritos a máquina* (*EM*), *Gramática filosófica* (*GF*), *Observaciones al fundamento de las matemáticas* (*OFM*) e *Investigaciones Filosóficas* (*IF*).

la imagen esencialista del significado. Si aceptamos la imagen esencialista del significado, admitiremos que la ‘comprensión’ posee una referencia señalable, por ejemplo, la comprensión de una orden se entendería como un proceso de mental/cerebral de interpretación de símbolos que ocasiona la acción de hacer efectiva tal orden. Según esto sería posible reconocer que un sistema artificial cumple con tales criterios, y, por tanto, comprende los símbolos. Respecto a este problema, con anterioridad John Searle se ha preguntado: ¿la capacidad de manipular símbolos acarrea la comprensión de dichos símbolos? Searle apoyándose en su experimento mental *la habitación china* responde negativamente a la posibilidad de que las máquinas puedan comprender. Sin embargo, esta clase de argumento remite a otros conceptos que explican la comprensión –en el caso de Searle la ‘intencionalidad’–, de manera que se construye una teoría para dar respuesta a la cuestión. No obstante, esta clase de argumentos se inclinan a formulaciones tales como:

1. Nada puede ser un agente que comprenda lo que otro dice si no posee *x* (remplace *x* por lo que desee: inteligencia, intencionalidad, conciencia, estados mentales); pero
2. ningún sistema artificial posee *x*; por lo tanto
3. ningún sistema artificial inteligente puede comprender lo que otro dice.

Haugeland (2001) nos ha prevenido sobre este tipo de argumentos que *a priori* descartan lo que se intenta indagar, por esta razón nos situamos en una investigación gramatical. Tal investigación no está dirigida a comprobar fenómenos, más bien revisa la gramática profunda de las expresiones en nuestra práctica del lenguaje, tal revisión aspira a la claridad y desiste de la construcción de las teorías. De modo que, nuestro segundo capítulo se configura desde la diferencia entre dos usos del verbo ‘comprender’, por un lado, la ‘comprensión superficial’ y, por otro lado, la ‘comprensión profunda’, esta última en relación estrecha con las reglas de un juego de lenguaje y las formas de vida.

1. Humanos que *siguen reglas* y máquinas que "siguen reglas"

Para empezar, recordemos la dicotomía cartesiana: el ser humano se compone de *res cogitans* y *res extensa*, la primera substancia dentro del orden de la razón contiene pensamientos, deseos, creencias, etc.; mientras que la segunda, gracias a su naturaleza material contiene mecanismos y objetos que son definidos por el espacio (tienen forma, tamaño y volumen). Luego, los pensamientos no tienen ninguna propiedad material y en la mente nunca podrán existir objetos físicos. A pesar de tal división, de alguna manera se presenta interacción entre *res cogitas* y *res extensa*, la mente (sus estados) causa, por ejemplo, el movimiento y las sensaciones causan experiencias (mentales) del mundo. Descartes, en el *Discurso del método* y las *Meditaciones metafísicas*[8], asegura el conocimiento directo, claro y certero del mundo mental, pero el mundo externo solo es cognoscible de modo indirecto a través de los sentidos. En consecuencia, se gesta la oposición entre lo interno (mente) y lo externo (cuerpo). Esta imagen cartesiana hereda un trabajo exigente para quienes se dedican a la filosofía, la psicología y la ciencia cognitiva, puesto que responder *qué* sucede y *cómo* sucede la relación entre mente y cuerpo se convierte en un paso obligado dentro de sus investigaciones.

La búsqueda de una respuesta a estas preguntas desencadena la construcción de modelos cognitivos que intentan caracterizar la mencionada relación; otros modelos como el conductista burlan el problema negando una parte de la dualidad. El conductismo no encuentra utilidad científica a una vida mental como compañía de la conducta, debido a que no es posible "medir" los fenómenos mentales a partir de datos científicos, en virtud de que cualquier estudio sobre la vida mental inevitablemente comienza a partir de la subjetividad y los datos que desde la introspección se aportan no resultan necesarios para desarrollar "la ciencia de la conducta". Contra la negación de la vida mental surge la tesis funcionalista, que intenta explicar la dicotomía cartesiana por medio de la diferencia computacional entre software y hardware. La tesis general del funcionalismo considera los estados mentales como estados funcionales que se constituyen por relaciones causales entre sí; estos estados

[8] Edición leída: Descartes (2011).

funcionales operan con inputs (información sensorial) para dar como resultado outputs (conducta).

Las ideas defendidas por el funcionalismo se rastrean en el modelo del pensamiento planteado por Allen Newell, Cliff Shaw y Herbert Simon; la piedra angular de dicho modelo es la hipótesis del sistema de símbolos físicos. La hipótesis asegura que un sistema de símbolos físicos cuenta con los medios necesarios y suficientes para realizar actos propios de inteligencia general. En palabras de Newell y Simón (1994) un 'sistema de símbolos físicos' es:

> Un sistema de símbolos físicos consiste en un conjunto de entidades, llamadas símbolos, que son modelos físicos que pueden presentarse como componentes de otro tipo de entidad llamada expresión (o estructura del símbolo). De este modo, la estructura del símbolo se compone de un número de instancias (o signos) de símbolos relacionados de alguna manera física (como cuando un símbolo se encuentra al lado de otro). En cualquier momento, el sistema contendrá un conjunto de estas estructuras de símbolos. Además de estas estructuras, el sistema también contiene un conjunto de procesos que operan en las expresiones para producir otras expresiones: procesos de creación, modificación, reproducción y destrucción. Un sistema de símbolos físico es una máquina que produce en el trascurso del tiempo un conjunto en desarrollo de estructuras de símbolos. Este sistema existe en un mundo de objetos más extenso que el de estas expresiones simbólicas en sí. (Newell y Simon, 1994, p. 121)

Jack Copeland (1996) subraya sobre la hipótesis del sistema de símbolos físicos tres funciones centrales 1) identificar, 2) designar e 3) interpretar:

1) La máquina manipula símbolos (lenguaje máquina).

2) Una expresión o una cadena de símbolos, que hace las veces de nombre, se asocia con una entidad (objeto), esto es la designación. Esta operación "[…] establece una relación de 'acceso' entre el nombre y la entidad: una vez que ha asignado un nombre a, por ejemplo, un registro particular, la máquina localizará automáticamente el registro cuando encuentra su nombre en la instrucción" (Copeland, 1996, p. 124).

3) La máquina manipula los símbolos para hacer efectiva una determinada instrucción.

Sumado a 1), 2) y 3), la hipótesis del sistema de símbolos físicos se configura en virtud de la particularidad que todo símbolo posee, a saber, la 'realizabilidad múltiple'. Esto es, un símbolo puede realizarse físicamente de cuantas maneras se desee, lo que supone la posibilidad de que dos máquinas ejecuten el mismo proceso de manipulación simbólica, aunque en su nivel físico estén fabricadas con diferentes materiales o se comporten de manera distinta. Por ejemplo, en el desarrollo de los computadores se han usado diferentes métodos para realizar los símbolos, pero desde la programación poco importa cómo se realizan físicamente, un computador bien podría estar construido con engranajes y piñones −como fue construida la máquina analítica de Charles Babbage− o como nuestros computadores electrónicos. Cabe subrayar que, el fenómeno de la realizabilidad múltiple, explicado por Turing (1950) en *Computing Machinery and Intelligence*, insinúa un nuevo camino para entender la mente/el cerebro y el pensamiento del ser humano: el cerebro se presenta como una realización orgánica de un sistema universal de símbolos, luego, es posible entender los procesos de pensamiento humano como similares a los procesos de computacionales.

Herbert Simon y Allen Newell (1962) intentaron explicar la conducta a través de procesos de información, es decir, la especificación de programas que producen el comportamiento. Dichos programas se componen de procesos de información elemental. Todo esto consecuencia de la descripción del método para el estudio del cómo el ser humano da solución a problemas, junto con otros procesos mentales. El método al que Simon y Newell apuntan consiste en la creación de una teoría de los procesos centrales en forma de un programa o clase de programas, que demuestra la suficiencia de la teoría para producir conducta de resolución de problemas realizada en un computador, y la teoría es comprobada sobre procesos humanos por medio de una comparación entre la descripción de los procesos que ejecuta el programa y el protocolo de un sujeto humano para resolver, por ejemplo, un mismo problema.

Ahora bien, retornemos a las ideas del funcionalismo, Hilary Putnam en su obra *Minds and Machines* (1975) aunque no comparte que la experiencia consciente sea idéntica a un proceso cerebral, describe los estados mentales como estados funcionales, esto último gracias a la noción de 'realizabilidad múltiple'. El ejemplo que constantemente se repite en Putnam para explicar dicha noción es el 'estado mental del dolor'. Según la teoría de la identidad de tipos el estado del dolor es un estado neurofisiológico, lo que exige que el mismo

estado neurofisiológico del dolor que presentan por ejemplo otras especies de mamíferos es igual que el estado neurofisiológico de dolor en el ser humano. De modo que existe un tipo físico-químico común para toda la variedad neurofisiológica de las diferentes especies y que está relacionado con el mismo estado mental, en este caso el dolor. Sin embargo, Putnam señala que si esto es así, entonces no podríamos reconocer, por ejemplo, en un extraterrestre el dolor, puesto que quizá no comparte nuestra composición físico-química. Putnam da un paso más y replantea la concordancia entre estados mentales y estados neurofisiológicos, lo que sugiere la correspondencia de estados mentales con estados computacionales.

Veamos lo anterior con un ejemplo: el computador posee un conjunto de programas y el hardware actúa de acuerdo con una serie de órdenes en sus programas, esto significa que atraviesa una serie de estados físicos hasta que realiza una tarea asignada. Sin embargo, otra máquina física diferente puede ejecutar la misma tarea, es decir, esta última atraviesa estados físicos diferentes, pero idénticos estados computacionales. Los estados computacionales no se definen por los estados físicos, sino por un conjunto de relaciones causales. Así, un estado mental no adquiere su identidad consecuencia de un estado neurofisiológico, sino por un conjunto de relaciones lógicas que mantiene con otros elementos del sistema. De esta forma, se erige la tesis principal del funcionalismo: los estados mentales son estados funcionales. Cabe aclarar que, existen diferentes matices de la tesis funcionalista, por su parte Putnam usó el modelo computacional teórico desarrollado por Alan Turing, y con este modelo señaló al ser humano como un autómata probabilístico. La realizabilidad múltiple abre la posibilidad funcionalista de construir mentes artificiales conceptualizando inteligencias no humanas, dado que, el funcionalista intenta explicar cómo un estado mental puede lograrse en diferentes estados cerebrales gracias a la plasticidad del cerebro[9].

Anterior al modelo funcionalista pero acorde con este, Kenneth Craik (1943) defendió que los seres humanos convierten los hechos externos en modelos internos y razonan mediante la manipulación de esta representaciones simbólicas, estos pueden volver a convertir los símbolos resultantes en acciones o reconocer una correspondencia entre estos y los hechos externos. Craik, aunque no vio en vida el ordenador digital, defendió que el sustrato físico del cerebro es menos pertinente que el modelo en que funciona. A partir de

[9] Otras obras relacionadas con el funcionalismo son las escritas por Jerry Fodor: *The Language of Thought* (1975) y *Psychological Explanation* (1968).

estos antecedentes, dentro del campo de la computación se ara la posibilidad de que la máquina sostenga un lenguaje comprensivo con los usuarios, debido a que los procesos o estados mentales en tanto estados funcionales que permiten el uso del lenguaje natural, pueden ocurrir en un sistema artificial.

Pese a la aceptación del modelo computacional de Newell y Simon y la tesis funcionalista, la analogía entre los procesos computacionales y los procesos mentales ha sido rebatida. Señalemos dos problemas con los que se encuentra dicha analogía: primero, el conocimiento y las representaciones del mundo se consideran como datos de entrada que deben ser matematizados con el fin de que puedan ser estructurados en algoritmos. Sin embargo, la posibilidad de representar el mundo, el conocimiento y la vida del ser humano mediante datos formalizados que puedan manejar los sistemas inteligentes resulta ser una idea controversial.

Segundo, el mismo Hilary Putnam (1994) identifica por un lado, el problema de simbolizar la inferencia inductiva mediante la lógica matemática[10] y, por otro lado, el aprendizaje del lenguaje. La inferencia inductiva se aplica, por ejemplo, a la capacidad de reconocer semejanzas entre cosas que no solo se apoya en patrones de representación, sino que descansan en lo que Wittgenstein llama 'aire de familia'; esto indica que, no hace falta que todos los miembros de una categoría tengan algo en común. Veamos. Cuando nos referimos a «Un hombre blanco» en contraste con «Una hoja blanca», el ser humano necesita algo más que un diccionario para comprender estas 'inducciones en conflicto', necesita conocimiento previo y una lógica tan compleja que dé cabida a la identificación de la semejanza:

> […] pero las semejanzas no son en absoluto elementos constantes del estímulo físico, ni sencillos patrones de información recibida por los sentidos. Por esta razón, el éxito que han tenido ciertos programas informáticos detectando patrones no resuelve el problema de la *semejanza* tal como se plantea a la hora de aprender un lenguaje natural. Lo que hace a los cuchillos semejantes, por ejemplo, no es que todos sean iguales (no lo son), sino que están todos hechos para cortar o ser clavados; y un sistema que pueda reconocer una semejanza

[10] Véase también para el desarrollo del problema de la lógica y la axiomatización de la realidad junto con el problema marco: McCarthy (1999).

> relevante en los cuchillos tiene que ser capaz de atribuir propósitos a los agentes. (Putnam, 1994, pp. 41-42)

La más afamada refutación a la analogía entre el ser humano y la máquina es la que sostiene John Searle (1984) y su experimento mental de la habitación china. Consecuencia de dicho argumento, Searle reconoce que las máquinas pueden manipular símbolos, pero no comprenden dichos símbolos. Entonces, aunque el modelo computacional de Newell y Simon realmente describiera "lo que sucede" en nuestro cerebro, la computación simbólica no posee 'intencionalidad'. Searle da el nombre de Inteligencia Artificial Fuerte al modelo computacional de la mente que puede llevar a cabo procesos mentales y, tal modelo, pretende explicar la misma capacidad humana realizar procesos mentales. No obstante, si el sistema artificial pudiera realizar procesos mentales, así mismo, debería llevar a cabo la comprensión de los símbolos. Pero, la comprensión humana es consecuencia de la intencionalidad, porque de esta se desprende nuestra facultad de significar. El significado de los símbolos en programación proviene de los programadores o de los usuarios, quienes ingresan inputs e interpretan outputs. Entonces, quien comprende es el observador, no la máquina. De los sistemas artificiales únicamente se puede predicar una 'intencionalidad derivada':

> Las manipulaciones formales de símbolos en sí mismos no poseen intencionalidad; no significan nada; ni siquiera son manipuladores de símbolos, pues los símbolos no simbolizan nada. En la jerga lingüística, sólo tiene una sintaxis, pero no tienen semántica. Dicha intencionalidad que los computadores parecen tener están únicamente en las mentes de quienes los programan y aquellos que los utilizan, los que envían los inputs y los interpretan. (Searle, 2001, p. 422)

El esfuerzo de refutación de Searle presenta problemas que ya han sido señalados por Margaret Boden (1994) y Jack Copeland (1996). Podemos resaltar de estas objeciones el carácter escurridizo de la 'intencionalidad'. De ahí que Searle necesite de otros argumentos para rebatir a la IAF, por ejemplo, la neuroproteína, a saber, el cerebro está hecho de un tipo especial de material y posee procesos químicos, que ellos y solo ellos generan lo que llamamos 'qualia'.

Ahora bien, hemos mencionado de modo superficial las raíces que nutren la analogía de nuestro interés: i. La equivalencia entre el funcionamiento de la mente/cerebro y los procesos algorítmicos computacionales. El modelo que comparte dicha equivalencia y que ha tenido mayor acogida por parte de la comunidad académica y científica es el modelo computacional-referencial de la mente (MCRM). En este modelo el pensamiento se entiende en términos de estructuras representacionales mentales sobre las que operan procesos computacionales. Quien está familiarizado con la programación conoce que los programas que integran los sistemas artificiales inteligentes son, en última instancia, estructuras de datos y algoritmos (procedimientos mecánicos que operan en distintos niveles de datos), es así que, el MCRM presenta el paralelo, por un lado, entre estructuras de datos y representaciones mentales y, por otro lado, entre algoritmos y procesos mentales. Para comprender de qué se trata el MCRM tomemos prestadas las palabras de Thagard (2008):

> Objetivo de la explicación: ¿Por qué las personan tienen una clase particular de *conducta inteligente*?
> Modelo explicativo: Las personas tienen representaciones mentales.
> Las personas tienen *procesos* algorítmicos que operan sobre esas *representaciones*.
> Los *procesos* aplicados a las *representaciones* generan la *conducta.* (p. 41)

Si aceptamos dar continuidad a la analogía, entonces la dinámica entre datos y algoritmos sucede en la mente[11]. De esta manera el MCRM se presenta como la superación del conductismo y es apoyado dentro de las ciencias y la psicología porque ofrece una respuesta específica a: ¿*qué* sucede en nuestra mente/cerebro? y ¿*cómo* sucede? A saber, dinámicas entre estructuras de datos y algoritmos[12]. Cabe aclarar que, a partir de la estructura general del MCRM se presentan diferentes enfoques de investigación que presentan variaciones en lo que se designa por 'representaciones'; por ejemplo, se ha equiparado las 'representaciones mentales' con: la lógica, las reglas, los conceptos, las analogías, las imágenes y las conexiones neuronales. Sin embargo, aunque las diferencias entre enfoques

[11] Debemos decir que en el campo de las ciencias empíricas se prefiere naturalizar 'la mente' y reemplazarla por 'el cerebro'(Bennett et al., 2008; Crick, 1995; Llinás, 2001; Penrose, 1994).
[12] Para profundizar sobre este modelo algortimico véase: Kitcher, (2001); Schank, (1986).

son importantes, no resultan relevantes para el despliegue de las observaciones de este capítulo, porque nos preocupamos por la otra parte de la analogía como ya veremos.

Si bien reconocemos que la metáfora enriquece la perspectiva sobre cualquier tipo de fenómeno, el MCRM ha sido blanco de diferentes críticas, ya sea porque deja de lado las emociones, la conciencia o el cuerpo encarnado; tales críticas han puesto en entredicho la capacidad del MCRM para representar la totalidad y el funcionamiento de la mente humana, sin embargo, aun cuando esta discusión es importante desborda los objetivos de este trabajo. Asimismo, no evaluamos cada uno de los modelos del MCRM y que son usados en la IA. En cambio, proponemos partir del siguiente razonamiento general:

1) La práctica del lenguaje natural es un tipo de conducta –o al menos se expresa en conductas–.
2) Si la práctica del lenguaje es un tipo de conducta, según el MCRM, entonces ser partícipe de un lenguaje es consecuencia de algún tipo de proceso o un conjunto de procesos mentales.
3) La analogía del MCRM propone entender los procesos mentales como algoritmos.
4) Por lo tanto, una máquina que emplea algoritmos, puede considerarse como un usuario del lenguaje en tanto se apliquen algoritmos similares, iguales o que sugieran los mismos resultados que los procesos mentales humanos, en nuestro caso el uso del lenguaje.

En vista de esto limitamos lo que sigue de este capítulo a solo una parte de la analogía del MRCM, a saber, la concordancia entre algoritmos y procesos mentales, con especial atención en el concepto de ‘algoritmo’ en la medida que se considera un procedimiento mecánico que una máquina puede seguir. El algoritmo es un conjunto de pasos que indican qué hacer a la máquina, lo que podemos asimilar que una máquina “sigue reglas”, de esto que nos preguntemos por el ser humano que sigue reglas. Todo esto sin perder de vista que la comprensión del lenguaje natural resulta ser consecuencia de procesos mentales representados por procesos algorítmicos, lo que sugiere que se pueden replicar en sistemas artificiales inteligentes. Por esto dividimos lo que queda del capítulo así: en 1.1 recogemos la noción de ‘algoritmo’ aceptada por la matemática e informática; después, en 1.2 resaltamos

algunas observaciones sobre las ideas de Turing y la acción mecánica de la máquina que sigue las sub-reglas del algoritmo, a partir de las observaciones de Wittgenstein sobre la diferencia entre seguir una regla mecánica y seguir una regla mecánicamente, y, para cerrar este capítulo, en 1.3 presentamos el fundamento normativo en la acción de seguir una regla dentro de un juego de lenguaje.

1.1 El algoritmo y la máquina

1.1.1 Concepto general de 'algoritmo'

En los años 30's y 40's el concepto matemático de 'algoritmo' ocupo buena parte de las reflexiones de: Kurt Gödel, Alonzo Church, Emil L. Post y Alan Turing, y, aún es núcleo de discusión entre teóricos informáticos y matemáticos (Gurevich, 2012; Moschovakis, 2001). Sin embargo, sobre los diferentes enfoques, por ejemplo, el desarrollo gödeliano de funciones primitivo-recursivas, las máquinas de Turing, los algoritmos de Markov, la tesis Turing-Church, etc., Fernández y Sáez Vacas (1987) recogen los rasgos comunes que permiten formular una caracterización general de algoritmo:

> 1) Hay unas *reglas*, o instrucciones, o prescripciones.
> 2) Tales reglas, instrucciones, etc., especifican una *secuencia* (encadenamiento) *de operaciones* o pasos.
> 3) Aunque de forma muy implícita […] las operaciones han de ser llevadas a cabo por un *agente ejecutor*, máquina o ser vivo, por sí o a través de otros agentes. Agente que es el destinatario de las instrucciones.
> 4) Más implícitamente aún, pero contenido en las definiciones, se establece que la secuencia de operaciones tiene una *duración*, que podrá ser tan larga como se quiera, pero ha de ser *finita* (p. 312).

En concordancia con esta caracterización advertimos algunas particularidades del algoritmo: primero, para tal concepto no resulta relevante el agente que diseña las reglas o instrucciones que hacen parte del algoritmo. Segundo, 1) y 2) implican que un algoritmo al menos debe poseer una secuencia de más de un paso. Veamos: «Prohibido fumar» o «PARE» son solo reglas pero no algoritmos; en cambio, «Apague la luz de la habitación, salga de la habitación y cierre la puerta» es una secuencia de tres pasos, de esto que se considere un algoritmo. Esto

último supone que los algoritmos deben ser discretos, a saber, los pasos encadenados deben con claridad ser diferenciados en el tiempo, dado que, un paso depende de la ejecución del anterior[13], entre tanto la ejecución de un algoritmo está dentro de una duración de tiempo medible. Tercero, el algoritmo puede ser ejecutado tanto por un ser humano como por una máquina. Y, por último, si el algoritmo tiene como propósito resolver un problema (ejecutar una tarea), por ejemplo, ganar una partida ajedrez, encontrar el producto de una multiplicación, salir de la habitación, etc., entonces, el algoritmo es finito, de lo contrario no se alcanzaría el cumplimiento de una tarea. Por esta razón, aunque el algoritmo se divida en subalgoritmos –tantos como se quieran–, estos son cadenas últimas que deben tener un final.

Sumado a esto, Durán (2018) llama la atención sobre otras dos características −acordes a las descritas por Fernández y Sáenz−: 5) el algoritmo es un proceso *preciso* y 6) tiene una *aplicación amplia*. Sobre 5) podemos decir que, el algoritmo como proceso preciso asegura que se pueda enseñar a otro sujeto *cómo* actuar en cada paso del proceso, de manera que pueda ser repetido y terminado. Acorde con esto, las reglas del algoritmo deben definir exactamente la acción a realizar, sin que haya lugar a la ambigüedad interpretativa o que dependa de la voluntad del agente que lo lleve a cabo. Entonces, el proceso de resolución de un problema se construye de modo que se puede cumplir exactamente, por ejemplo, por una persona que incluso no tenga el mínimo concepto del problema a resolver, únicamente se exige que esta persona sea capaz de cumplir operaciones simples y atienda escrupulosamente al proceso descrito por el algoritmo. Por último, respecto a 6), un algoritmo no se diseña para un problema particular, sino una clase de problemas para los que los datos de entrada y los resultados finales pertenecen respectivamente a conjuntos específicos. Agreguemos a esta característica que, para la solución de un problema pueden crearse diferentes algoritmos que conducen a la solución.

En definitiva, el algoritmo es una concatenación de reglas que si se siguen correctamente da como resultado la solución esperada y el agente que repite el proceso algorítmico no necesita comprender el sentido de este, es decir, repite de modo *mecánico* en el sentido de que el ejecutor se limita a *entender* las reglas y realizarlas, sin necesidad de pensar sobre ellas. En consecuencia, hasta lo aquí expuesto, podríamos decir que tanto

[13] En términos computacionales es posible que se ejecuten varios pasos a la par, esto se denomina ‘computación paralela’, sin embargo, también remite a un proceso causal.

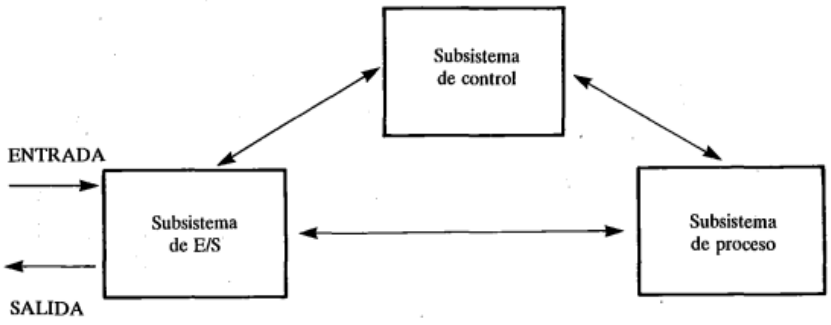
humanos como máquinas pueden ejecutar algoritmos para cumplir con tareas, por ejemplo: calcular.

1.1.2 Máquinas que ejecutan algoritmos

Jean-Luc Charbert (1998) afirma que: "Algorithms are simply a set of step by step instructions, to be carried out quite mechanically, so as to achieve some desired result (p. 1). Sin embargo, esta definición es amplia para los algoritmos que se pueden implementar en la computadora, por esto que la ciencia de la computación límite lo considerado por 'algoritmo computacional', como veremos más adelante. Según el apartado anterior, el algoritmo es abstracto en la medida que se construye para resolver diferentes problemas siempre y cuando pertenezcan a la misma clase, por ejemplo, un algoritmo creado para solucionar una multiplicación de dos números enteros que constan de dos dígitos puede ser aplicado tanto a 20 x 20 como a 79 x 67, sin embargo, para la aplicación del algoritmo es necesario *datos de entrada*, en nuestro ejemplo, hace faltan números. Además, después de aplicar un algoritmo se obtiene un resultado, esto se define como los *datos de salida*. Teniendo esto en cuenta tomemos prestado el diagrama de Fernández y Sáez (1987, p. 315) para estructurar una máquina capaz de ejecutar algoritmos:

Figura 1

Esquema de Organización de una Máquina Capaz de Ejecutar Algoritmos

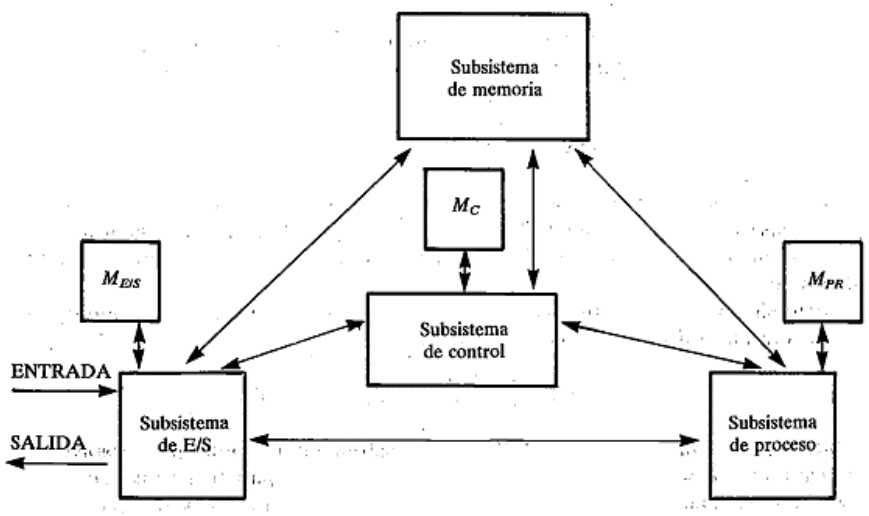


Nota. La figura representa el proceso de una máquina que capaz de ejecutar algoritmos. Tomado de *Fundamentos de la informática* (p. 315) por Fernández y Sáez, 1987, Alianza Editorial.

El subsistema de E/S es donde comunican las órdenes a la máquina y los datos de entrada, y, a su vez la máquina genera resultados después de la aplicación del algoritmo. El subsistema de proceso es el espacio en el que se ejecutan las órdenes que manipulan los datos y dan lugar a la realización del algoritmo. Por último, "el subsistema de control "donde se consigue el secuenciamiento adecuado para la ejecución de las órdenes y donde se generan las señales de control […]" (Fernández & Sáez Vacas, 1987, p. 317) necesarias para el funcionamiento de los otros dos subsistemas. Sumado a esto, el algoritmo entendido como una secuencia de pasos genera estados nuevos del algoritmo, en otras palabras, cada vez que se ejecuta una instrucción del algoritmo se genera un nuevo estado causado por el anterior; esto implica que la máquina debe memorizar el estado del algoritmo: "los subsistemas podrán llevar asociados elementos de memoria (asociados siempre, como se sabe, a todo circuito secuencial) para memorizar las informaciones de estado que corresponden a la misión del subsistema" (Fernández & Sáez Vacas, 1987, p. 317). En las computadoras resulta más económico y conveniente agrupar los mencionados elementos de memorización en un subsistema específico:

Figura 2

Esquema de Organización de una Máquina Computadora que Ejecutar Algoritmos



Nota. La figura representa el proceso de una computadora capaz de ejecutar algoritmos. Tomado de *Fundamentos de la informática* (p. 317) por Fernández y Sáez, 1987, Alianza Editorial.

En concordancia con el diagrama de Fernández y Sáenz, Trajtenbrot (1977) resume lo que sucede en una computadora digital así:

a) Dispositivo de memoria: graba todos los datos necesarios, incluyendo el programa.
b) Unidad aritmética: el tratamiento de los datos que se introducen en ella para obtener el resultado necesario (por ejemplo, la suma dos números) se realiza por medio de la transformación en un dispositivo electrónico de las señales eléctricas de entrada que representan los datos iniciales en señales eléctricas que representan los datos de salida. Los datos de entrada llegan a la unidad aritmética de las células de memoria, en donde estaban almacenadas, y la señal de salida parte de ella a la célula donde se va a conservar.
c) Unidad de control: el *dispositivo de control* en cada etapa del funcionamiento de la máquina crea las condiciones para la realización de la operación siguiente del proceso. Al hacer esto actúa como una central telefónica automática y une a los «abonados» (dispositivos y células de la computadora) que participan con cada una de las operaciones. Diciéndolo metafóricamente, la unidad de control mira el programa y en concordancia con él manda las órdenes sobre el funcionamiento de las unidades de la computadora que tienen que garantizar la operación siguiente. (pp. 46-47)

La descripción de Trajtenbrot guarda concordancia con la imagen de la computadora digital que presenta Turing (1950) en *Computing Machinery and Intelligence* §4. En su artículo Turing propone el concepto de 'computabilidad'. La tesis Church-Turing[14] defiende que un sistema real o abstracto es calculable algorítmicamente, en otras palabras, existe un algoritmo (conocido o desconocido) para calcular la conducta futura. Así que, es posible describir de modo correcto las salidas de un sistema a partir de las descripciones correctas de las entradas que recibe, y, además, que el procedimiento funcione con todas las entradas posibles que pertenezcan a la misma clase. Con Turing podemos decir que, toda tarea ejecutable de manera efectiva es computable, y con Church, toda función numérica es recursiva. Ahora bien, en la actualidad un computador procesa un conjunto dado de

[14] El desarrollo de las ideas de Alan Turing en Cambridge son paralelas a los trabajos de Alonzo Church en Estados Unidos, aunque se publicaron con meses de diferencia, más tarde se concluyó que son equivalentes, por esto que se les llame: tesis Church-Turing. Mientras que Turing expreso sus ideas en términos (que luego se han denominado) Máquinas de Turing, Church lo hizo términos del cálculo lambda (Church, 1936).

programas, que, a su vez son sistemas calculables algorítmicamente. De esto que podamos decir que, un sistema calculable algorítmicamente puede ser simulado mediante un sistema de símbolos universal (lenguaje máquina).

Para desarrollo de la noción de 'computabilidad' noción Turing presenta el paralelo entre una computadora humana y la idea de la computadora digital; la primera sigue reglas fijas y no puede desviarse de estas para realizar la acción de calcular, veamos:

a) Almacenamiento de información: los datos y las reglas del algoritmo, además el lugar donde se consignan estos, por ejemplo, una hoja de papel o la memoria.
b) El tratamiento de información: se presupone que el calculador puede realizar los pasos del algoritmo, es decir, seguir instrucciones.
c) El control del proceso: la toma de decisión para realizar una determinada etapa dada del proceso y para preparar su cumplimiento se ejecuta por el mismo calculador de acuerdo con la instrucción.

Antes de continuar debemos aclarar que, no podemos asegurar que el objetivo de Turing es describir el funcionamiento del pensamiento en ser humano (Copeland, 2011). No obstante, desde la lectura de artículo *Computing Machinery and Intelligence*, se indentifica que Turing pasa desde un argumento matemático a la afirmación de una tesis cognoscitiva, germen que conserva el MRCM –esto lo revisamos en el apartado 1.2 –

Según la descripción de Trajtenbrot y Turing, podríamos decir que el programa introducido en la computadora guía el funcionamiento de la misma. No obstante, Durán (2018) nos llama la atención sobre la diferencia entre 'algoritmo' y 'programa computacional'. En contra de Cantwell Smith (1985), quien asume que el programa computacional es un conjunto de instrucciones que tienen lugar en una computadora física, Durán (2018) diferencia entre "[…] un proceso paso a paso entendido como una fórmula sintáctica (el algoritmo) de un proceso paso a paso a paso que pone la máquina física en los estados causales apropiados (el proceso de computación)" (p.218). Entonces, la noción de 'programa computacional' se divide en dos: el algoritmo y el proceso computacional. De esto que, se caracterice el algoritmo como elemento abstracto, pues consiste en una secuencia simbólica que no da cuenta de las relaciones causales, porque no dependen de una relación

causal con el contexto espacio-temporal; por otro lado, los algoritmos son formales en cuanto se rigen por las leyes lógico-matemáticas, las cuales establecen las directrices para manipular sistemáticamente la secuencia de símbolos que lo componen. De modo que, el algoritmo de alguna manera empuja al proceso computacional gracias a que la computadora, por medio de la manipulación del lenguaje máquina puede ejecutar el algoritmo como se espera[15].

De acuerdo con Durán (2018) reescribimos la noción de algoritmo, esta vez de algoritmo computacional:

> 1. Un algoritmo es definido como un conjunto finito y organizado de instrucciones, que debe satisfacer cierto conjunto de condiciones con la intención de proveer soluciones a un problema;
> 2. Un algoritmo debe ser capaz de ser escrito en un determinado lenguaje;
> 3. El algoritmo es un procedimiento que es llevado a cabo paso a paso;
> 4. La acción de cada paso está estrictamente determinada por el algoritmo, la entrada de datos y los resultados obtenidos en pasos previos;
> 5. Cualesquiera que sean los datos de entrada, la ejecución del algoritmo debe terminar después de un número finito de pasos;
> 6. El comportamiento del algoritmo es físicamente instanciado durante la implementación en la computadora. (p. 221)

1.2 Humanos que siguen reglas mecánicas y máquinas que siguen reglas mecánicamente

Gödel (1981), seguido de Church y Turing, demuestra que matemáticamente resulta imposible que todo problema sea soluble por medio de la aplicación de un algoritmo. Si bien, esta idea es fundamento fuerte para la construcción de una crítica a la IA que busca crear sistemas artificiales que usen el lenguaje natural, en estas líneas no nos preocupa desarrollar tal discusión, en tanto nuestra pregunta no versa sobre la posibilidad de crear sistemas artificiales que funcionen por medio de algoritmos tan complejos que puedan o no usar el lenguaje natural o si el mundo en su totalidad puede ser axiomatizado y representado para

[15] Con esta diferencia hecha por Durán retornamos al problema cartesiano *cómo* se relaciona la mente y el cuerpo, en otras palabras, como se comunica el algoritmo y el proceso de computación.

que un sistema artificial pueda manejarlo. Nuestra propuesta no refiere a un problema técnico o a un problema de capacidad computacional. Ya que, la disputa que toma este camino cae en falacias a la posteridad o en falacias a la ignorancia. Nuestro punto de partida para la discusión presta su atención al siguiente razonamiento:

> Podemos entender el algoritmo como un conjunto de reglas que deben ser seguidas para procurar el resultado deseado.
> En una máquina física se ejecutan algoritmos. Específicamente, un sistema artificial inteligente aplica algoritmos –algunos de estos no son reconocidos por los mismos creadores del sistema–
> Entonces, un sistema artificial inteligente ha de seguir reglas, para cumplir el cometido para el que fue diseñada.

Hasta este punto del razonamiento no hay objeciones, el problema surge cuando el MRCM asume que: si una máquina entrega datos de salida correctos cuando aplica un algoritmo, por ejemplo, para dar solución a una multiplicación de dos números que contienen dos cifras, y un ser humano también llega al mismo resultado, entonces el proceso algorítmico que permite realizar el cálculo en la máquina resulta ser idéntico o similar a lo que acaece en el ser humano, en otras palabras, que la máquina *sigue reglas* del mismo modo que el ser humano *sigue reglas.*

1.2.1 Máquinas que siguen reglas

Turing (1950) ejemplifica la analogía entre el ser humano y la computadora digital mediante la acción calcular, sin embargo, es un análisis conceptual reductivo cuya importancia epistemológica reside en la demostración de que los algoritmos pueden definirse como sistemas complejos de "sub-reglas sin sentido", cada una de las cuales puede aplicarse mecánicamente. Turing pasa de un desarrollo matemático sobre la noción de computabilidad a una imagen epistemológica de cómo funciona la mente humana cuando calcula. Según Turing el cálculo tiene su origen en fenómenos preconscientes y ocultos a la introspección que solo pueden determinarse a través de los resultados observables del sujeto. Dado que

conocemos el funcionamiento de procesos físicos, se busca una descripción a los fenómenos preconscientes y ocultos, a saber, como procesos formales.

Teniendo en cuenta la anterior afirmación, tomemos prestado un ejemplo de Wittgenstein quien en medio de sus reflexiones acerca de la comprensión imagina una 'máquina de lectura' que ejecuta una 'lectura mecánica':

> ¿Qué sucede, pues, cuando él, por ejemplo, lee el periódico? – Sus ojos resbalan –como decimos– a lo largo de las palabras impresas, él las pronuncia – o las dice sólo para sus adentros; en particular lee ciertas palabras captando su forma impresa como un todo, otras cuando su ojo ha captado las primeras sílabas, otras aún las lee sílaba por sílaba y alguna que otra quizá letra por letra no la dice en voz alta ni tampoco para sus adentros, pero después es capaz de repetir la oración literal o aproximadamente. –Puede poner atención en lo que lee o también –como podríamos decir– funcionar como una **máquina de leer: quiero decir leer, en voz alta y correctamente son poner atención en lo que lee**: quizá mientras su atención se fija en algo enteramente diferente (de modo que no es capaz de decir lo que ha leído si se le pregunta inmediatamente después) (*IF*, §156).

Aquí la máquina de lectura lee mecánicamente –igual que en el caso de Turing decimos que una persona calcula mecánicamente–, sin embargo, en el ejemplo se diferencia al menos tres usos del verbo 'leer':

A) el uso mecánico de la máquina que lee símbolos,
B) el uso en el que se expresa que un ser humano lee sin interés de comprender y, por último,
C) el uso que expresa que un ser humano comprende lo que lee.

B) y C) en nuestro contexto constantemente pueden confundirse: podemos imaginar estar en presencia de un expositor del cual dudamos si comprende o no lo que lee en sus diapositivas. No obstante, la analogía que emplea el MRCM se establece entre A) y C), a saber, la correspondencia entre la descripción del uso causal y el uso comprensivo. Preguntémonos sobre B) y C): ¿cuál es el criterio para afirmar que un expositor solo "lee mecánicamente" o que "lee lo que comprende"?, ¿el criterio que usamos para señalar dicha diferencia es alguna

configuración neuronal?, ¿el criterio es una respuesta empírica que responde al por qué y cuándo se lee compresivamente?

Cuando afirmamos que la comprensión es algo sucede antes de leer, es decir que, frente a leer hay un diferencia entre la posesión de tal habilidad y la aplicación de la misma (*IF*, §149), nos vemos obligados a decir que leer o calcular es un estado o un proceso para el que se construyen respuestas como procesos algorítmicos que causa la conducta observada (*CA*, p. 117). La diferencia entre 'leer' y 'leer mecánicamente' es algún tipo de estado mental preconsciente o alguna configuración neuronal oculta de la cual damos cuenta gracias a los resultados que percibimos. El MCRM pretende –partiendo de los datos de entrada y los datos de salida– avistar procesos algorítmicos que corresponden a tales estados preconscientes en el caso de los seres humanos que leen, calculan o usan el lenguaje, sin embargo, este modelo sólo funciona como hipótesis que explica lo que se percibe, el problema de esta "hipótesis" es que no es susceptible a verificación.

Ante estas observaciones los defensores del MCRM pueden responder: un sujeto puede narrar el protocolo que utiliza para encontrar, por ejemplo, la solución de un cálculo, en otras palabras, la construcción de un algoritmo que funciona en su mente. Sin embargo, cuando se pregunta a una persona «¿cómo llegaste a la solución correcta de *x* cálculo?», aunque es posible que tal sujeto responda con claridad el protocolo (las reglas) que paso a paso a seguido, cuando se pregunta a un sujeto por «¿cómo comprende correctamente lo que lee?» la respuesta no resulta ser tan clara.

Aquí el problema de ningún modo es enunciar: «–La máquina calcula» o «–La máquina lee», en algún sentido podemos decir que la máquina está guiada por los símbolos que lee, la cuestión es: ¿un niño, por ejemplo, se guía por las reglas que le han enseñado –para calcular o leer– de manera mecánica aplicando ciertos datos a tales reglas? ¿Acaso 'ser guiado' en el uso que se aplica al ser humano puede reducirse a 'ser guiado' en el sentido de la máquina? El uso de Turing cuando de alguna manera dice que la máquina calcula es ilícito porque es categóricamente diferente el sentido normativo que usamos cuando nos referimos a los seres humanos, Shanker (1998) nos da un ejemplo:

> Reading is indeed a family-resemblance concept: a family which ranges from the infant responding to flash cards to a philosophy class reading the Investigations or a navigator

> reading the stars. But when we speak, for example, of a scanner ‘reading’ a bar code, this is not a still more primitive extension of an infant reacting to signs.
> The mechanist insists that the only difference between the scanner ‘reading’ a bar code and someone reading something mechanica lly lies in the operations of the internal mechanisms guiding their respective actions. (p.57)

Es necesario aclarar que, Wittgenstein de ninguna manera excluye los fenómenos neurofisiológicos o mentales de un ser humano cuando, por ejemplo, calcula o lee pues esto describiría algo que pasa en nuestro cerebro o mente, pero “lo que nos interesa del cálculo o de la lectura no es interno” (*GF*, p.99). En otras palabras, cuando nos preguntamos sobre B) y C): ¿cuál es el criterio para afirmar que un expositor solo “lee mecánicamente” o que “lee lo que comprende”?, ¿el criterio que usamos para señalar dicha diferencia es alguna configuración neuronal?, ¿el criterio es una respuesta empírica que responde al por qué y cuándo se lee compresivamente? De ninguna manera afirmamos que el expositor comprende o no, porque estemos viendo qué sucede en su cerebro o en su mente, más bien afirmamos que comprende o no porque tal acción acaece dentro de un contexto normativo que se comparte entre los asistentes a la exposición y el expositor, es así que los primeros pueden señalar que el segundo comprende o no. Entonces, lo que pase internamente no es el criterio que usamos para decidir si el expositor lee o lee mecánicamente (finge leer), el criterio está en su comportamiento y la situación en la que esto ocurre, es decir al mismo actuar –no a un proceso interno– lo llamamos leer comprensivamente o leer mecánicamente.

Cabe señalar, toda acción del ser humano no es normativa, por ejemplo el reflejo involuntario rotuliano. El golpe en la rodilla ocasiona el estiramiento del tendón y el músculo que se conecta con él. Después, se envía un mensaje a la médula espinal para indicarle que se estiró el músculo, a su vez la médula espinal envía un mensaje al músculo para indicarle que se contraiga. La contracción del músculo nos hace patear. La reacción de patear no está condicionada por una regla, más bien es una acción causada por un encadenamiento de sucesos que empiezan con el golpe en la rodilla. Hay que aceptar que las acciones causales y las acciones normativas no siempre se pueden distinguir, pero de esta indeterminación no se sigue asegurar que todo comportamiento es mecánico. Cuando decimos que un expositor lee “como una máquina”, no apuntamos a un mecanismo interno que suceda en su cerebro o

mente, intentamos señalar su falta de atención, sus errores o la incapacidad de responder preguntas, etc., y, esto lo aseguramos a partir de todo un contexto alrededor de la exposición.

La diferencia entre un mecanismo causal y un contexto normativo se deja entrever, también, cuando se cuestiona por: ¿cuál es la primera palabra que leyó la máquina?, esta pregunta es fácil porque se refiere a un proceso causal sin embargo, en el contexto normativo, nos tenemos que remitir a otras relaciones de las que no se puede señalar su temporalidad. Esta argumentación continúa en el capítulo II, cuando nos preguntamos por 'procesos mentales' correspondientes a la comprensión del lenguaje natural.

Sobre lo anterior los partidarios del MRCM podrían preguntar: ¿acaso, no se podría asumir la posibilidad de trazar un mapa de sus procesos neuronales que suceden cuando un agente está calculando o leyendo? Sin embargo, cada vez que **leo** *Ante la ley,* cuento escrito por Franz Kafka, comprendo algo diferente, ¿los procesos neurofisiológicos son los mismos o difieren entre ellos? ¿Cuándo los demás leen el mismo cuento tienen procesos neurofisiológicos idénticos a los mismos? ¿Cuándo los demás leen el mismo cuento tienen procesos algorítmicos? Entonces, la diferencia entre un proceso mecanicista y un contexto normativo se hace evidente.

Podemos suponer otra objeción en contra de Wittgenstein: las observaciones del filósofo austríaco se reducen a la conducta (datos de salida) para suponer que, por ejemplo, un expositor no lee mecánicamente. Además, en la acción de calcular, tanto la máquina como un ser humano pueden seguir las mismas reglas, recuerden que los protocolos que logran narrar los individuos, sirven de base para la creación de algoritmos que ejecutan los sistemas artificiales. Frente a esta objeción podemos responder: por un lado, el comportamiento, en nuestro ejemplo, del expositor no se reduce a una relación causal como sucede cuando nos golpean la rodilla (ejemplo arriba), sino que, el comportamiento se justifica gracias a las reglas en un juego de lenguaje y el criterio que usan los otros para señalar que el expositor "lee mecánicamente", también, se funda en tales reglas. Por otro lado, si bien el ser humano puede narrar los pasos que realizó para llegar, por ejemplo, al resultado de un cálculo, seguir reglas de un cálculo no se reduce a un proceso causal, como veremos en el siguiente apartado.

En definitiva, la imagen que se utiliza para decir que «la máquina lee» o «la máquina calcula» es la operación gobernada por reglas que se aplican a datos de entrada para obtener datos de salida. Sin embargo, las reglas que siguen los seres humanos para cualquier acción

ya sea calcular o leer son reglas públicas, de ninguna manera, procesos ocultos de los que sea necesario conceptualizar como equivalentes a procesos algorítmicos computacionales.

1.2.2 Humanos que siguen reglas

¿Cuál es la diferencia entre un cálculo ejecutado por una máquina y el realizado por una persona? ¿Cuál es la diferencia entre dispositivo electrónico que lee un documento, con una voz robótica, para un usuario y un sujeto que lee? La diferencia que desde Wittgenstein se plantea es la posibilidad de justificación que puede señalar un ser humano mediante las reglas, por ejemplo, si preguntamos a un sujeto que calcula: ¿cómo llegó a tal resultado?, el sujeto no solo da cuenta de operaciones causales, sino qué tipo de cálculo eligió o a qué leyes de la aritmética se remite tal desarrollo, con qué finalidad hace dicho cálculo, puede reconocer un error y corregirlo.

Shanker (1998) señala la importancia de distinguir entre justificación y explicación causal. Es decir, la aplicación de reglas que determinan si podemos describir a un sujeto que lee *x* o calcula *x* es diferente a la descripción las causas –si existen– de un sujeto o una máquina que dice *x* o escribe *x.* Esto resalta una característica importante para nuestra argumentación: el ser humano realiza cálculo en la práctica, igual que lee o usa el lenguaje en la práctica, es decir el cálculo y el lenguaje se realiza siempre en contextos, y estos contextos son normativos. En efecto, la acción de calcular es una acción normativa, de ninguna manera se reduce a un mecanismo causal, descriptivo o empírico. Mientras que las reglas determinan y justifican el resultado del cálculo, la descripción de las causas podría dar cuenta de lo que efectivamente ocurre, pero no lo pueden determinar normativamente. El concepto de ‘calcular’ tiene una relación con ‘seguir una regla’. Ahora bien, esto puede sugerir que: si el cálculo viene determinado por reglas precisas que deben seguirse según un orden estrictamente diseñado para obtener un resultado, entonces estas reglas pueden llevarse a cabo mecánicamente, un cálculo puramente mecánico que una máquina podría hacer. Por esto que debamos detenernos en la noción de ‘seguir una regla’.

¿De qué se trata es el carácter normativo que subyace a ‘seguir una regla’? Podemos confundir el hecho de ‘seguir una regla mecánica’ con ‘seguir una regla mecánicamente’. Veamos. Las reglas del ajedrez indican a los jugadores cómo jugar, cómo mover, cuándo sus fichas están en peligro, cómo atacar, cuál es el orden de turno, etc. De modo que, muestran

algo más que los movimientos permitidos, si dos fichas están en la misma posición del tablero o un jugador realiza dos movimientos seguidos, decimos que mientras se juegue ajedrez estas acciones resultan ser un sin-sentido. Es así que las reglas, no sólo dicen qué se debe hacer, también son la justificación para la corrección de los sin sentidos en un juego. Además, el carácter normativo imprescindible para la regla también justifica el comportamiento dentro del juego. Si preguntaran al sujeto A por su movimiento de la ficha *x* en el tablero de ajedrez, A puede justificar su accionar a partir de las reglas, quizá responda:

- Porque la ficha *x* tiene la posibilidad de *y* movimiento.
- Porque *x* sirve de carnada para poder eliminar en mi siguiente turno un elemento del contrincante que tiene una ubicación desfavorable para mi estrategia
- Porque puedo dominar más casillas en el tablero, etc. Y estas justificaciones tienen pleno sentido dentro del juego de ajedrez.

De ahí que, las reglas cumplen el papel de corrección y justificación de las acciones dentro de un juego, Wittgenstein escribe en *IF*, §82:

> ¿A qué llamo 'la regla según la cual él procede'? – ¿A la hipótesis que describe satisfactoriamente su uso de la palabra que observamos; o a la regla que él consulta al usar los signos; o la que él nos da la respuesta si le preguntamos por su regla? – ¿Y qué pasa si la observación no permite reconocer claramente ninguna regla y la pregunta no revela ninguna? – Pues él me dio ciertamente una explicación cuando le pregunté qué es lo que entiende por "N, pero estaba dispuesto a retirar y modificar esa explicación. – ¿Cómo debo, pues, determinar la regla desacuerdo con la cual él juega? Él mismo no lo sabe. – O más correctamente: ¿Qué debe aún significar aquí la expresión: "regla según la cual él procede"?

Se debe aclarar que, no todos los movimientos de un juego vienen determinados por las reglas (*IF*, §68 y §84). Retomando, las reglas de, por ejemplo, del ajedrez no describen el juego, más bien lo constituyen porque es jugado conforme a estas. Es decir, las reglas se reflejan en la acción, el desarrollo de un cálculo expresa las mismas reglas aritméticas que se presentan como necesarias para poder calcular. Tales reglas deben ser públicas, de lo contrario el otro jugador no podría reconocer si A efectivamente ha jugado correctamente o no, no podría ni siquiera determinar si sus propias jugadas son correctas. Es así que, las reglas

son el estándar de corrección en la práctica, de esto que sea necesaria la enseñanza de las reglas mediante ejemplos. Por ejemplo, no se aprende a realizar una multiplicación con algunas reglas enunciadas, más bien se aprende realizando multiplicaciones.

Si prestamos atención a la práctica del lenguaje natural, que Wittgenstein describe en sus *IF*, se infiere que el lenguaje es atravesado por reglas gramaticales, esto es la práctica del lenguaje es una práctica normativa, aún más compleja que la acción de calcular o leer. A su vez, la práctica normativa presenta relación con las formas de vida y los demás elementos del juego de lenguaje. No hemos dedicado esfuerzos en la descripción del concepto 'juego de lenguaje' en tanto la academia cuenta con suficiente literatura sobre este tópico.

Terminemos este apartado con algunas líneas de Wittgenstein:

> ¿Calcula la calculadora? Imagina que hubiera aparecido por casualidad una calculadora, y que alguien aprieta casualmente sus botones (o un animal anda sobre ellos) y ella realiza el producto 25 x 20. Quiero decir: es esencial a la matemática que sus signos se usen también en lo civil. Es el uso fuera de la matemática, es decir, el significado de los signos, lo que convierte en matemática el juego de signos (*OFM*, 5,§ 63)

Con Wittgenstein podemos decir que la calculadora no calcula aunque presente los resultados correctos, un cálculo y la actividad de calcular del ser humano esta contextualizado dentro del 'juego de calcular', después de la lectura del segundo capítulo podremos decir que esta contextualizado en las formas de vida humana. Para que una máquina calcule debe comprender todo el trasfondo del juego debe existir un trasfondo, identificar criterios de corrección, justificar la acción del cálculo, y, entender el contexto en el que se aplica y el para qué se aplica tal cálculo. Finalmente, inferimos que el MCRM no confunde el uso mecánico de 'seguir una regla' con el uso normativo que funda la acción del ser humano. El MCRM aunque pretenda dar respuesta a qué sucede en la mente realiza un uso ilícito de la analogía entre máquinas que siguen algoritmos y seres humanos que siguen reglas, aunque estas reglas se establezcan en un algoritmo, el ser humano aplica cualquier algoritmo, aunque sea el más simplísimo y por esto se piense que es una acción mecánica el algoritmo se sigue dentro de un contexto que determina el correcto o no desarrollo del mismo.

2. Comprender lo que otro quiere decir

Hemos de arar a lo largo de todo el lenguaje
La rama dorada de Frazer

Con anterioridad hemos revisado la noción de *algoritmo* y su íntima relación con los sistemas artificiales inteligentes, en específico con la acción de *calcular* y la acción de *seguir una regla.* Consecuencia de este análisis se resaltan dos ideas: primero, el carácter normativo que subyace a la acción de *seguir una regla* y, segundo, la relevancia de *aplicar reglas* y de *comprender* en la acción humana de calcular, puesto que el ser humano cuando calcula puede evaluar dicha acción y en caso necesario corregirla. A partir de estas ideas surge la pregunta: ¿qué es 'comprender'? o, mejor aún ¿qué quiero decir cuando enuncio «−Ahora comprendo»?

En la práctica del lenguaje natural la comprensión del significado de las expresiones resulta ser una particularidad intrínseca, y por esto la IA ha dedicado esfuerzos que persiguen no sólo la resolución de problemas o la respuesta a preguntas sobre un campo específico del conocimiento, tal como lo hace un experto, sino que también ha buscado que los sistemas artificiales sean usuarios del lenguaje natural o al menos simulen serlo. A partir de esto planteamos la cuestión: ¿el sistema artificial *x* comprende lo que se le dice? Para dar respuesta a esto es necesario esclarecer qué es *comprender*. Sin embargo, la formulación que se pregunta por el *qué* resulta ser arriesgada en tanto la discusión se puede dar por terminada al caer en el relativismo conceptual como sucede con el término 'inteligencia': –¡Si un sistema artificial es inteligente o no depende a lo que llames 'inteligencia'! (Boden, 2017; Copeland, 1996; Haugeland, 2001). Sumado a esto, los investigadores en diversas áreas tienen problemas con la pregunta sobre el *qué*, en la medida que la pregunta sugiere un método de investigación que asume, por un lado, –gracias a la teoría esencialista del significado– la comprensión como un fenómeno o un estado de cosas disponible para la descripción. Y, por otro lado, la dicotomía cartesiana (interior y exterior), entre mente y cuerpo.

Contrario a esto, desde las líneas wittgenstenianas la acción de comprender[16] un lenguaje no depende de establecer un referente inequívoco o encontrar una esencia que corresponda a 'comprender'. Con el fin de esclarecer las implicaciones y los alcances de la propuesta wittgensteniana en el ejercicio de dar respuesta a la posibilidad de que sistemas artificiales comprendan lo que enuncian y lo que otro les quiere decir, llevaremos a cabo una indagación de tipo gramatical que permita abordar el problema de manera diferente y transformar la pregunta «¿*qué* es 'comprender'?» por «¿*cómo* se *usa* el término 'comprensión'?»

Sobre esto último podemos identificar en Wittgenstein la distinción de dos usos de 'comprender'[17]. Veamos. Abro el libro *Pedro Páramo* en una página al azar (sin leer con anterioridad las páginas precedentes o conocer el contexto de la historia) y leo la oración:

> «−Es el caballo de Miguel Páramo, que galopa por el camino de la Media Luna.»

Cuando he terminado de leer dicha oración de inmediato enuncio: «–He comprendido la oración». No cabe duda de que el lector también puede asentir y unirse a esta afirmación, lo que sugiere que dicha oración es entendida independiente del resto de la obra, esto en virtud de que podemos ofrecer expresiones análogas o sinónimas; por ejemplo:

> Hay un animal-caballo propiedad de un sujeto, en este caso propiedad de Miguel Páramo, este animal galopa por una zona llamada Media Luna.

En efecto, la oración extraída del libro *Pedro Páramo* de algún modo es comprendida gracias a que se domina el español. Sin embargo, si el lector está familiarizado con el resto de la novela de Juan Rulfo, la oración en cuestión se entiende de modo distinto en tanto que la

[16] Wittgenstein en sus escritos intermedios *CAM, GF* y *EM* dedica parte de sus líneas a la descripción de los usos de 'comprensión', este tópico reaparecerá en *IF*. Cabe aclarar que, gracias al singular método wittgensteniano no se encuentra una descripción completa de la comprensión, más bien un paciente estudio de casos particulares que ponen de manifiesto rasgos gramaticales de 'comprender'.

[17] Para fines de este capítulo la palabra en alemán *Verstehen* se traduce como 'comprender' y 'entender', esto porque en la traducción del año 1988 de *IF* realizada por Alfonso García y Ulises Moulines se usa en gran medida el verbo 'comprender' cada vez que aparece *Verstehen*, así mismo sucede en las traducciones de *CAM, GF* y *EM* y en la literatura sobre Wittgenstein. Sin embargo, en la última edición de *IF* de Ulises Moulines (2012) se traduce con más frecuencia *Verstehen* por 'entender'. De esto que decidamos no hacer distinción.

inscribimos en la totalidad de la obra (*GF*, §5), por ejemplo: se entiende como la descripción de un evento fantasmal que evoca miedo en el personaje a quién se le comunica dicha oración, ya que el personaje conoce con antelación que el caballo y su dueño hace tiempo dejaron esta vida. Estos dos usos del verbo 'comprender' podemos leerlos directamente en *IF*, §531:

> Hablamos de entender una oración en el sentido de que ésta puede ser sustituida por otra que diga lo mismo; pero también en el sentido en que no puede ser sustituida por ninguna otra. (Como tampoco un tema musical puede ser sustituido por otro.)
>
> En el primer caso el pensamiento de la oración es lo que es común a diversas oraciones; en el segundo, se trata de algo que sólo esas palabras, en esa posición, pueden expresar (Entender un poema).[18]

En el primer caso se propone el uso de 'comprensión' como *comprensión superficial*[19], esto se enmarca en la idea de que un mismo sentido puede manifestarse mediante diferentes expresiones, y, así mismo, es posible comprender la locución de otro cuando se formula una oración equivalente. Este uso nos recuerda las consideraciones fregeanas sobre el sentido de una proposición. La perspectiva de Frege sobre el sentido y el significado de una proposición alimentó la idea tradicional sobre el lenguaje como expresión del pensamiento. Frege, en su ensayo *El pensamiento* (1918)[20], propone al pensamiento como elemento fundamental que otorga sentido a los signos, es quien responde por el contenido de una proposición, lo que sugiere que un mismo pensamiento puede expresarse por completo a través de diferentes oraciones. En general, la tesis de Frege se inscribe en la visión esencialista del lenguaje. De ahí que, comprender una proposición resulte ser "ir más allá" del propio lenguaje y descubrir su esencia no-lingüística, descubrir lo que el otro quiere decir, lo que el otro piensa.

En el segundo caso Wittgenstein muestra el uso *profundo*, a saber, lo que quiero decir sólo es posible comunicarlo mediante el uso de ciertos signos, pero esto no se refiere al signo

[18] Además el lector puede consultar: *IF*, §§ 532–533.

[19] 'Comprensión superficial' y 'comprensión profunda' no son conceptos planteados por Wittgenstein, más bien proponemos estos nombres con un fin explicativo y sugiriendo en analogía la diferencia entre 'gramática superficial' y 'gramática profunda', conceptos usados por Wittgenstein en *IF* y *GF* y expuestos por Rojo (1999) en *Wittgenstein: Gramática y metafísica*.

[20] Edición leída: Frege, 2013, pp. 196-225.

en cuanto tal, sino al sentido de dichos signos, y, en este caso, el sentido se da dentro de un juego de lenguaje. En otras palabras, lo que otorga sentido a los signos no es un pensamiento, no es una esencia, no es un objeto o imagen mental, sino el uso de una expresión dentro de un juego de lenguaje en el que convergen actividades lingüísticas y extra-lingüísticas. Entonces, la *comprensión profunda* también dirige su atención al sentido de una expresión, sin embargo, el signo se impregna de sentido gracias a la dinámica del juego de lenguaje al que pertenece su uso, misma dinámica en la que convergen las reglas del juego de lenguaje y sus formas de vida correspondientes.

Con el ánimo de aclarar aún más esta diferencia entre comprensión superficial y profunda pongamos por caso la acción de traducir. Se ordena al sujeto P la traducción de una oración del inglés británico (en el cual no es versado) al español, su idioma nativo. La oración que P debe traducir es:

«– That was bloody deliciuos!

Junto con esta oración se proporciona una clave para descifrar la oración correspondiente en español, por ejemplo, la clave son las equivalencias de las palabras inglesas con las palabras del español. P aplica la correspondencia entre palabras y como resultado presenta la oración en español:

«– ¡Eso fue sangrientamente delicioso!»

Después de esto ¿podemos concluir que P comprende la oración inglesa, dado que ha traducido y ha entendido la oración equivalente en el español? Un británico que termina de cenar y pronuncia esta oración, frente al intento de traducción de P, señalaría que no ha estado cerca de comprender, puesto que la correspondencia entre las palabras no sirve en esta situación como criterio para la comprensión. En el ejercicio de traducción al no-británico se le ha escapado algo: "bloody deliciuos" en este caso particular no funciona para expresar alguna relación con 'sangre' o 'sangriento', más bien es usada para expresar: 'muy delicioso' o 'demasiado delicioso' y esto se logra entrever desde una visión panorámica, o como la llamará Wittgenstein, *representación perspicua*[21].

[21] Para una profundización del término léase: *GF*, §2; Lagerspetz (2020); Padilla Gálvez (2014).

No obstante, a esta última idea wittgensteniana se podría objetar que: si se proporciona los datos suficientes y las reglas necesarias para la correcta traducción, P no erraría en este caso particular. Sin embargo, si esto es así aún tendríamos que responder: ¿la comprensión se hace efectiva en el momento que he traducido correctamente la oración al español, puesto que conozco los signos del español? ¿La comprensión de la oración inglesa es algo que debe ser anterior a la traducción? ¿Se comprende palabra por palabra y gracias a tales elementos se puede comprender la oración en su totalidad? ¿En qué *momento* sucede la comprensión? Incluso si se logra la traducción correcta aún se ubica en un nivel superficial, porque la expresión 'bloody deliciuos' puede jugar correctamente en otra situación o en otro juego de lenguaje sin catalogarse como un sin-sentido, por ejemplo, si leemos la oración en la voz de un personaje de una novela de caníbales o de vampiros. De aquí que se hace necesario girar la atención al juego en general. En vista de esto, hace falta algo más que reglas de equivalencia para que la comprensión profunda deje de ser escurridiza. Bajo la propuesta wittgensteniana, como ya se ha mencionado, hace falta compartir el juego del lenguaje y, con este, las *formas de vida*, en nuestro ejemplo, de un inglés.

En pocas palabras, la comprensión superficial se refiere al entendimiento de una proposición gracias a una relación con otras palabras o entidades de algún tipo, mientras que la comprensión profunda sugiere que el sujeto que comprende comparte el juego de lenguaje y las reglas con quien emite la expresión. Esto a su vez implica que ambos sujetos comparten *formas de vida.* El análisis de la relación entre formas de vida y la comprensión, entendida como comprensión profunda, será el objetivo principal de este capítulo. Para alcanzar este cometido, primero, en 2.1 se procura una aproximación a la relación gramatical entre 'comprender una oración' y '*saber* lo que la oración quiere decir', esto a partir de algunas observaciones sobre la tesis esencialista del lenguaje, y, por otro lado, la idea de la comprensión como un proceso mental/cerebral o un estado mental. Segundo, en 2.2 se explica la propuesta alternativa de Wittgenstein sobre la comprensión entendida como *capacidad-habilidad*. Y, por último, desde la noción de *formas de vida*, en 2.3, se responde a la posibilidad de postular la capacidad de comprender como característica de los sistemas artificiales inteligentes.

2.1 Comprender una oración es ‘saber lo que la oración dice’

Habitualmente aseguramos que hemos comprendido una oración en cuanto *sabemos* su significado, de esto que se familiarice ‘comprender una oración’ con ‘saber lo que la oración dice’. Ahora bien, *saber* el significado de una expresión de ninguna manera se puede entender como *poseer* dicho significado. Si aceptáramos la idea de que el ser humano posee de algún modo los significados, admitiríamos, según Wittgenstein, que la comprensión se alcanzaría en tanto delimitamos con exactitud el significado de una palabra y, a su vez, conservamos las denominaciones en algún lugar. En contra de esta idea, en *IF* §150 se muestra el vínculo entre la gramática de la palabra ‘saber’ y la gramática de palabras tales como ‘poder’ o ‘ser capaz’, así como el vínculo entre la palabra ‘comprender’ con ‘dominar una técnica’. Para esclarecer estos vínculos en principio exponemos algunas observaciones sobre, siguiendo a Wittgenstein, la teoría referencial del significado, porque como ya se ha dicho comprender una expresión remite a comprender su significado.

A partir de la teoría tradicional referencialista de corte esencialista variadas son las propuestas acerca del significado. Para el propósito de estas anotaciones agrupamos dichas propuestas en dos conjuntos: (2.1.1) el significado como idea general y (2.1.2) el significado como una imagen mental que guarda relación con la palabra y el objeto. Junto con esta clasificación, abordamos (2.1.3) la defensa de la comprensión como un tipo de proceso mental/cerebral y (2.1.4) la idea que propone entender la comprensión como un estado mental consciente. El último apartado, (2.1.5) se dedica a escudriñar el segundo supuesto en el que se funda gran parte de la inteligencia artificial: la dicotomía mente-cuerpo.

2.1.1 Observaciones sobre significado como esencia/idea general

Wittgenstein inicia *CM* e *IF* con el ya famoso pasaje de *Confesiones* que resume en líneas generales la propuesta esencialista del lenguaje. San Agustín expone el lenguaje como un sistema de símbolos que se impregna de significado gracias a la correlación con el objeto –de cualquier tipo– al que representa. Leamos directamente las líneas agustinianas *Confesiones*:

> […] al querer expresar todos los sentimientos de mi corazón con gemidos y voces varias y variados movimientos de mi cuerpo para que mí voluntad fuese obedecida, y al no ser capaz de expresarme ni en todo lo que quería ni a todos los que quería. Según aquéllos nombraban alguna cosa, la iba grabando en la memoria; y cuando, según aquella palabra, movían su cuerpo hacia algún objeto, lo veía y retenía que aquel objeto era designado por ellos mediante el sonido que pronunciaban cuando lo querían mostrar.
>
> Que éste era su propósito se percibía por el movimiento del cuerpo, como si de una especie de palabras espontáneas de todas las culturas se tratase, que se hacen con el rostro y con el asentimiento de ojos, y con el movimiento de las demás extremidades, y con el sonido de la voz que indica el estado afectivo al pedir, poseer, rechazar o evitar las cosas. De ese modo, poco a poco iba deduciendo de qué cosas eran signo las palabras colocadas dentro de frases distintas en su debido lugar y oídas muchas veces; y a través de ellas iba ya enunciando mis deseos con una boca instruida en esos signos. Así que intercambié signos para enunciar mis deseos con aquellos entre los que vivía, y me adentré más y más en la tempestuosa compañía de la vida humana, sometido a la autoridad de mis padres y a las indicaciones de mis mayores. (San Agustín, 2010, pp. 130-131)

Para el desarrollo de nuestro capítulo deseo subrayar: 1) San Agustín describe la frustración que invade sus intentos por transmitir sus deseos, y, gracias a su inteligencia da cuenta poco a poco de los sonidos que usan los demás para nombrar objetos, de manera que, halla conexiones nombre-objeto. 2) San Agustín concibe al ser humano con un interior que pretende exteriorizar por medio del lenguaje, en otras palabras, el lenguaje resulta ser el vehículo de sus pensamientos, sentimientos privados y estados mentales[22]. La imagen del lenguaje agustiniana, imagen que también se encuentra en el pensamiento de Russell en *Logical Atomism* o de Wittgenstein en el *Tractatus,* se presenta en pocas palabras así: toda palabra tiene un significado, y este significado es una entidad la cual es reemplazada por la palabra. Todas las tesis que parten de la idea del lenguaje como figura de la referencia se plantean preguntas tales como: ¿qué hay detrás del lenguaje? ¿Cuál es la esencia del lenguaje? A pesar del esfuerzo por responder, esta posición esencialista resulta ser reducida cuando se explora el funcionamiento del lenguaje en la cotidianidad.

[22] Para revisar en específico el planteamiento agustiniano sobre el lenguaje leer: Beuchot (1986); San Agustín, (2003).

Indaguemos brevemente la tesis que defiende el significado de un término como la referencia de cualquier tipo. Si admitimos tal hipótesis, el significado se entiende como una idea general, ya que para diversos particulares usamos el mismo nombre, de modo que debe existir "algo común" entre particulares. De lo que se infiere que 'comprender' una expresión es aprehender algún tipo de entidad común a las referencias particulares correspondientes a tal expresión. En otras palabras, comprender una palabra equivale a tener una idea general contrapuesta a las imágenes particulares que intervienen en el aprendizaje de dicha palabra y, además, en el uso cotidiano de la misma. Si admitiéramos "lo común" como el significado, ¡Eureka! encontraríamos la esencia de *significar* y con ella de *comprender*.

Conforme a esta idea abstraccionista, la acción de conceptualizar se limita a la demarcación ideal, inequívoca y exacta de los significados de las palabras, a saber, una esencia común que sea compartida por los objetos que sean designados por una misma palabra. Wittgenstein contra todo esencialismo, y consecuente con su propuesta sobre el parecido de familia entre 'significado' y 'uso', propone la discusión en otra perspectiva: la investigación gramatical, a saber, el significado de una palabra/expresión dentro de su contexto de uso, porque limitarse a la dimensión conceptual-referencial reduce la característica dinámica del propio lenguaje (*IF,* §23). Decimos "dinámica" porque el significado como *uso* no es absolutamente fijo, sino que pertenece al dominio de las practicas humanas, con más precisión diríamos que, el lenguaje en sí mismo es *praxis*.

El desarrollo de la discusión se puede aclarar con la analogía de "límites borrosos" de los conceptos (*EM,* 60-70), esto es la inexactitud conceptual. Quienes no aceptan esta idea que Wittgenstein explica podrían objetar preguntando: «− ¿Cómo sabes lo que quieres significar con la palabra 'silla' si no has delimitado el concepto? ¡A cualquier cosa podrías llamar 'silla'!». Intentemos, entonces, hallar algo común a los diferentes elementos que llamamos 'silla': podríamos decir que la esencia de la palabra 'silla' es que «¡−Tiene cuatro patas!», no obstante, otro replicaría mostrando imágenes de sillas que no poseen dicha característica; asimismo sucede con el tamaño, el color, el material, y otros accidentes. Quizás entonces el esencialista podría responder, con Aristóteles, el *para qué* es su esencia, para 'silla' es: «−que un sujeto tome asiento». No obstante, el concepto se torna escurridizo, porque existen otros objetos creados para tal fin y no por ello damos el nombre de 'silla' a estos, o, por otro lado, si un artista realiza la escultura de una silla, aunque no tome asiento

nadie, nos atrevemos a señalarla como ‘silla’. Es así como una idea general exacta que representa “algo común” no es intrínseca al mundo, más bien son categorías que formulamos con una finalidad específica.

A pesar de la dificultad para la teoría esencialista del significado aún no se responde: ¿cómo sabemos lo que se quiere decir con *x* si los límites de tal concepto pueden ser borrosos? Con Wittgenstein podemos señalar que la pregunta ya conlleva una trampa, pues sugiere que es necesaria la exactitud conceptual y la posesión de una esencia para que sea comprendida una expresión. Imaginemos que hemos ido de viaje y observamos un paisaje hermoso, cuando llegamos a casa intentamos describir tal paisaje a nuestro hermano (que no ha viajado con nosotros), sin embargo, tenemos la sensación que en nuestro relato se han escapado detalles, nuestra descripción no ha sido exacta. No obstante, no ha sido en vano puesto que nuestro relato ha cumplido una función, por ejemplo, provocar la visita de nuestro hermano a tal lugar. Entonces, la necesidad de la exactitud queda en entredicho. En analogía con el lenguaje, siempre se pueden mostrar ejemplos en los que se usa una expresión o describir contextos de uso, de modo que no es necesario la definición exacta de una esencia para que podamos entablar un diálogo con los otros. Wittgenstein nos recuerda que ‘inexacto’ no quiere decir ‘inutilizable’:

> ‘Inexacto’ es realmente un reproche, y ‘exacto’ un elogio. Pero esto quiere decir: lo inexacto no alcanza su meta tan perfectamente como lo exacto. Todo depende, pues, de lo que llamemos ‘la meta’. ¿Soy inexacto si no doy nuestra distancia del Sol con 1 metro de precisión; y si no le indico al carpintero la anchura de la mesa con 0,001mm de precisión?
>
> No se ha previsto *un* ideal de precisión; no sabemos lo que debemos imaginarnos por ello - a no ser que tú mismo estipules lo que debe llamarse así. Pero te será más difícil acertar con una estipulación; con una que te satisfaga. (*IF*, §88)

Aceptemos por un momento que la función de conceptualizar es delimitar exactamente lo que *es* de lo que *no es*, entonces, el concepto ‘juego’ separa lo que es juego de lo que no es juego. En consecuencia, retornaríamos a la búsqueda de un referente exacto, inequívoco y señalable que pueda reemplazar al término y sea compartido por todos los

juegos. Contrario a esto Wittgenstein señala que los límites se crean y destruyen según el fin que perseguimos en un juego de lenguaje:

> ¿Cómo le explicaríamos a alguien qué es un juego? Creo que le describiríamos *juegos* y podríamos añadir a la descripción: "esto y *cosas similares* se llaman 'juegos'". ¿Y acaso sabemos nosotros mismos algo más? ¿Es acaso sólo a los demás a quienes no podemos describir que es un juego? ─Pero esto no es ignorancia. No conocemos los límites porque no hay ninguno trazado. Como hemos dicho, podemos –para una finalidad especial– trazar un límite. ¿Acaso sólo así hacemos utilizables el concepto? ¡De ningún modo! A menos que se trate de esa finalidad especial. Como tampoco haría utilizable la medida de longitud '1 paso' quien diese la definición: 1 paso = 75 cm. Y si quieres decir "Pero anteriormente no era una medida de longitud exacta", entonces respondo: bueno, entonces era inexacta. −Aunque todavía me debes la definición de exactitud. (*IF*, §69)

Sumado a esto, si decidimos perseguir la utopía de la exactitud, en la que nos embarcaríamos gracias a la búsqueda de la esencia del significado, sería necesario poseer un lenguaje ideal con el cual comparar los demás; por ejemplo, para la palabra 'juego' necesitamos de un referente de un juego ideal y exacto, con el cual compararíamos a los demás elementos particulares, con el fin de decidir cuáles particulares pertenecen a dicha denominación y cuáles no. Pero, al parecer, no tenemos un ideal y tampoco nos hace falta para comunicarnos. En cambio, nosotros somos quienes construimos criterios para la formulación de categorías para usos específicos.

Si la dinámica del lenguaje y, por consiguiente, de los conceptos, no interfiere con la posibilidad de comunicabilidad del lenguaje (*EM*, 248-252; *IF* §§23 – 24 y §630), entonces, el elemento común tan ansiado por quienes persiguen la exactitud conceptual no existe o por lo menos no es necesario para comprender una palabra[23]. En lugar de insistir en los rasgos comunes que poseen los objetos, estos tienen una amplia variedad de usos diferentes. En vista de esto, Wittgenstein se interesa por el lenguaje desde la perspectiva pragmática,

[23] La familiaridad entre palabras y expresiones es una manifestación del juego de lenguaje, y esta característica justifica que se denominen particulares bajo el mismo término, no la suposición de una esencia común a ellos, véase: *GF*, 35. Ahora bien, se debe aclarar que Wittgenstein no defiende la libre interpretación conceptual extrapolada que desencadena en el solipsismo lingüístico, contra el lenguaje privado se despliega la relación entre el juego de lenguaje y las reglas de tal juego.

dirigiendo su atención a ejemplos concretos —muchos de ellos imaginados— de personas que usan un lenguaje y actúan en el curso de su vida diaria con diferentes fines, no sólo el de referenciar. Veamos. *IF* comienza con una situación cotidiana en la que se emplea el lenguaje:

> Piensa ahora en este uso del lenguaje: envío a alguien a comprar. Le doy una hoja en la que aparecen los signos: "cinco manzanas rojas". Esta persona lleva la hora al tendero; éste abre el cajón sobre el que aparece el signo "manzanas"; luego busca en una tabla la palabra "rojo" y frente a ella encuentra una muestra de color; a continuación dice la serie de los número cardinales –supongo que la sabe de memoria– hasta llegar a la palabra "cinco", y por cada numeral roma del cajón una manzana que tiene el color de la muestra. (*IF*, §1)

Con estas líneas Wittgenstein no intenta proponer alguna nueva esencia del lenguaje extraída del uso, sino que, muestra al lenguaje como herramienta dentro de una actividad cotidiana particular, como lo es comprar; de modo que la atención gira en torno a la función de las expresiones, es decir, aquí el lenguaje de ninguna manera tiene la función de referenciar entidades, describir una esencia o manifestar un pensamiento, más bien, se intenta con el lenguaje provocar una respuesta o una acción por parte del interlocutor. De ahí que el foco de atención no se dirige al objeto referenciado por la palabra 'rojo' o 'cinco', sino a las acciones del tendero frente a la solicitud del cliente, es decir, cómo el tendero actúa dentro de determinado contexto, cómo el lenguaje se entrelaza con las acciones extra-lingüísticas. La analogía formulada en *IF* entre expresiones del lenguaje y herramientas acentúa la importancia en la función de las palabras frente al buscado carácter esencial de estas, como bien lo explica McGinn (1997):

> The comparison is one which brings out the everydayness of language, which focuses on the humdrum aspect of its practical role in our lives, and which thereby makes language look less 'gaseous'. The comparison also works against our urge to look for the representational essence of language, for we simply don't feel the same need to explain what makes a tool a tool, or to describe a common essence of tools. What makes a tool a tool is simply that it is used as a tool, each kind of tool in its own specific way. (p. 47)

En vista de este cambio logramos adelantarnos y entrever que no es necesario que todo el que entienda una expresión capte algún tipo de entidad, sino que los individuos sean capaces de usar la expresión en diferentes contextos para propósitos específicos. De esto que se sustituya la búsqueda de la esencia de las cosas y la necesidad de "penetrar en los fenómenos" por una descripción de cada caso, mediante ejemplos. Sin embargo, antes de exponer la noción de comprensión profunda como capacidad de usar un lenguaje revisaremos otras tesis que parten de la idea esencialista aquí expuesta.

2.1.2 Observaciones al significado como 'imagen mental'

Otra teoría esencialista acerca del significado defiende que el lenguaje es manifestación de la vida mental de un sujeto; esta es la tesis mentalista. En líneas generales esta tesis formula la correspondencia palabra-imagen mental, es decir, el significado de las expresiones son las imágenes mentales, y estas a su vez se *deben* guardar en "algún lugar", a saber, la mente −que algunos identifican como el cerebro−. Por consiguiente, 'comprender' es la *posesión* de dichas imágenes mentales. En otras palabras, se comprende el significado de una expresión en cuanto se evoca la imagen mental equivalente, por ejemplo, la imagen de un objeto en el caso de los sustantivos, de una acción en el caso de los verbos, etc. De acuerdo con esto, cuando el sujeto A ordena a B: «−Toma una flor roja», el receptor de la orden comprende tal oración en tanto evoca en su mente una mancha roja, mancha que es sintetizada con la imagen de una flor. Sin embargo, rememorar la imagen mental no es necesario para que B comprenda la orden de A. Wittgenstein defiende esto último mostrando la consecuencia absurda que surge de la idea contraria: si se ordena al sujeto B: «−Imagina una mancha roja», para obedecer esta orden B debe imaginar otra mancha roja con la cual se compara la que he ordenado que imagine, y a su vez debe imaginar otra mancha roja para compararla con la inmediatamente anterior mancha roja, lo que nos llevaría una secuencia *ad infinitum* (*CA*, p. 3.).

Wittgenstein sustituye las imágenes mentales por la observancia de objetos reales situados que juegan en medio de la acción; de nuevo, si se ordena al sujeto B: «−Elige una flor roja», la descripción del comportamiento de B en relación con el significado de 'rojo' no incluye una imagen mental, sino una consulta a una gama de colores que se muestra en una tabla física de colores y su comprensión se evidencia en la acción misma de elegir una flor

entre las demás (esto mismo sucede con el ejemplo de *IF* §1 mencionado en el anterior apartado). Es decir, si efectivamente B elige una flor roja, es prueba de que ha comprendido el término (esta idea es profundizada en 2.2). Por lo tanto, no es necesario postular una entidad que soporte la comprensión, una imagen mental que deba aparecer en posesión de un individuo para después señalar que tal individuo comprende. De esto que, la imagen mental, si bien no se niega su existencia, deja de ser necesaria para comprender.

Para desentrañar la crítica wittgensteniana es preciso formularse dos preguntas: ¿por qué es necesario que un sujeto *demuestre* que ha comprendido? ¿Por qué nos atrevemos a decir que: si el sujeto B toma una flor roja ha entendido la orden de A? En nuestro uso del lenguaje continuamente enunciamos:

«–Ella/Él ha comprendido.»

«–Hemos comprendido.»

«–Usted ha comprendido.»

Damos cuenta de que un tercero ha comprendido una orden, una operación matemática, una situación, una señal de tránsito, etc., es decir, aseguramos que los otros comprenden y con este supuesto vivimos en tanto nos comunicamos. Esto nos lleva a cuestionarnos cómo logramos saber que efectivamente el otro ha comprendido, por esto que sea necesario algún tipo de "prueba". Veamos esto con mayor detalle.

El modelo cartesiano nos ubica en el problema de dar cuenta de los otros. Como sujeto se tiene acceso directo a las imágenes mentales propias, pero de ningún modo a las imágenes mentales del otro, y, si hay algún tipo de conocimiento sobre la vida mental de los demás es de manera indirecta e inferencial. Esta imagen cartesiana limita la certeza únicamente a los predicados psicológicos en primera persona, sin embargo, en el quehacer cotidiano usamos oraciones en las que señalamos la comprensión de los otros. Entonces, desde la postura cartesiana y la tesis mentalista nos vemos ante una encrucijada: ¿qué razones tengo para suponer o enunciar que el otro comprende una expresión sino no tengo acceso a su mente? ¿Qué razones/criterios tengo para sustentar que comprendo correctamente una expresión si sólo tengo acceso a mis imágenes mentales? A pesar de la profundidad que estas preguntas aparentan, Wittgenstein las elude pues muestra los problemas que la idea dicotómica cartesiana presenta, en *IF* leemos:

> §293: […] Supongamos que cada uno tuviera una caja y dentro hubiera algo que llamamos "escarabajo". Nadie puede mirar en la caja de otro; y cada uno dice que él sabe lo que es un escarabajo sólo por la visión de *su* escarabajo. – Aquí podría muy bien ser que cada uno tuviese una cosa distinta en su caja. Incluso podríamos imaginar que una cosa así cambiase continuamente. – ¿Pero y si, no obstante, la palabra "escarabajo" de esa gente tuviese un uso? – Entonces no sería el del nombre de una cosa. La cosa que hay en la caja no pertenece en absoluto al juego de lenguaje; ni siquiera como un *algo*: pues la caja podría incluso estar vacía. –No, se puede 'simplificar la fracción' por la cosa que hay en la caja; se esfuma, sea lo que fuere. Es decir: si se construye la gramática de la expresión de la sensación según el modelo de 'objeto y designación', entonces el objeto cae fuera de la consideración por irrelevante.

> §294: Cuando dices que él ve ante sí una imagen privada, la cual describe, entonces has hecho en todo caso una suposición sobre lo que tiene ante sí. Y esto quiere decir que puedes describirla más de cerca, o que la describes. Si admites que no tienes ninguna idea de qué tipo de cosa podría ser lo que él tiene ante sí – ¿qué te lleva entonces a decir a pesar de todo que él tiene algo ante sí? ¿No es como si yo dijera de alguien: "Él *tiene* algo. Pero no sé si es dinero o deudas o una caja vacía"?

No profundizaremos en la extendida discusión sobre el lenguaje privado[24] que se vincula con la tesis mentalista del significado y en consecuencia de la imagen cartesiana, pero diremos que, si se acepta tal tesis podemos presumir que sólo una persona puede comprender un lenguaje, y, esta persona utilizaría el lenguaje de modo particular, es decir, dicha persona utilizaría reglas privadas para fijar la referencia de sus palabras, por ejemplo. En últimas el lenguaje privado cierra la puerta a la comunicación con el otro. Esto mismo acontece con la idea del significado como imagen mental. La imagen mental 'escarabajo' como fenómeno privado no puede ni compartirse, ni compararse, es decir, solo el sujeto puede conocer su propia imagen mental de 'escarabajo', lo que implica un problema para la

[24] La idea de un lenguaje privado "A nivel semántico postula la posibilidad de entender el significado de una palabra como una entidad identificable con una percepción u objeto mental interno de un sujeto aislado. A nivel epistémico describe la existencia de un conocimiento puramente subjetivo e incomunicable" (Moreno May, 2011, p. 188).

misma verificación de dichas imágenes, porque ni el sujeto que afirma comprender puede conocer si la imagen evocada es la correcta, ni el otro que escucha tal afirmación o que observa el comportamiento de tal sujeto puede asegurar que el otro comprende. A pesar de las consecuencias que se siguen de la figura cartesiana seguimos usando expresiones tales como: «–Ahora comprendo», «–Usted comprende» sin detenernos a preguntar cómo sucede tal comprensión.

Ahora retomemos la idea del significado como imagen mental. Líneas atrás hemos dicho que no se niega la existencia de imágenes mentales, pero de ninguna manera esto quiere decir que las imágenes fundamenten el significado de los términos. Pongamos por caso la imagen mental de la suma (operación matemática), aunque el sujeto imagine tal imagen esto no quiere decir que pueda aplicar la adición en la acción cotidiana o que simplemente pueda sumar. Aunque se acepte que el sujeto tiene la imagen mental adecuada del objeto al que referencia un término, no quiere decir que la comprenda, es decir, que el individuo pueda *demostrar* la comprensión a través de la aplicación de dicha palabra. Las imágenes mentales que tenemos de, por ejemplo un árbol, son privadas y sólo son comprobadas por el sujeto que las evoca, lo cual, como ya lo hemos dicho, calificaría al significado como privado[25]. Luego, no es posible establecer un criterio tanto para el uso correcto como para la comprensión. Bajo la tesis mentalista no es posible indicar que una persona ha comprendido o no una palabra/expresión, ni siquiera si yo he comprendido[26]. De modo que, la comprensión se transforma en un fenómeno oscuro que se nos oculta y debe ser desvelado. Contrario a esto, a partir de la propuesta wittgensteniana no hace falta descubrir una esencia velada, la acción de comprender se señala bajo criterios externos y públicos. Podemos adelantarnos a decir que la acción que demuestra la comprensión es intrínseca a la comprensión misma (esta idea se desarrolla en el apartado 2.2). El señalado carácter público de la comprensión es compartido con las reglas de juego de lenguaje, como se muestra en el parentesco entre ‘comprender’ y ‘dominar una técnica’, y, recuérdese que dominar una técnica es actuar conforme a las reglas en un juego de lenguaje.

---

[25] Wittgenstein desarrolla argumentos relacionados con la imagen mental para explicar por qué el significado tampoco se reduce una imagen mnémica, para el desarrollo de esto leer: *EM,* 182 – 188.

[26] *GF* §34: […] Supóngase que se trata de la palabra “rojo” y que yo digo automáticamente que la comprendo; ahora se me pregunta de nuevo: “¿realmente la comprendes?”. Evoco entonces, a manera de control, una imagen roja. Pero, ¿cómo sé que el color que se me aparece es el correcto? Y, sin embargo, digo ahora, completamente convencido, que la comprendo.

Antes de seguir con el desarrollo de la argumentación es necesario aclarar: Wittgenstein no deja de lado la idea de que una imagen mental sea parte del significado, asimismo tampoco rechaza que la explicación ostensiva o la referencia sean parte de este. Cuando se postula el *uso* emparentado con el *significado* se propone mostrar el error de toda tesis reduccionista sobre lenguaje que busca una esencia: "(…) la idea de *una* esencia del lenguaje podría resumirse de la siguiente manera: La naturaleza (exclusiva) del lenguaje humano no consiste en un lenguaje que se *refiere a entidades*" (Gallardo, 2011, p. 88) sean objetos materiales o ideas mentales. El uso muestra que el lenguaje cumple con un gran número de funciones en la vida humana, dentro de las cuales referir entidades es sólo una de ellas (Gallardo, 2011; Herrera Ibáñez, 1989; Pérez Otero, 2000).

2.1.3 Observaciones a la comprensión entendida como proceso mental o ¿Describir un proceso mental es igual que describir mi habitación?

Paralelo a la idea de la comprensión como posesión de algún tipo de entidad se desarrolla la tesis que defiende la comprensión como un *proceso interno*, para luego describir tal proceso desvelando así la esencia de la comprensión. Esto es, un mecanismo mental o físico que pueda ser señalado con el término de 'comprender' y que acompañe otras acciones, por ejemplo, la acción de leer. Si describiéramos dicho mecanismo explicaríamos el fenómeno de la compresión, y, a su vez esto permitiría su posterior replica en sistemas artificiales inteligentes. Si se parte de esta consideración la comprensión en general es arrancada de la acción misma de comprender dentro de un contexto y acciones particulares, es decir, se considera como una función: C(*x*), donde *x* puede ser reemplazado por: lectura, locución, orden, problema, cálculo, etc., mientras que el proceso de comprensión (C) abstraído se puede aplicar a cualquier caso. Entonces, si halláramos tal descripción del proceso de la comprensión (C), encontraríamos el concepto exacto e inequívoco aplicable a cualquier caso particular.

Cuando enunciamos: «–Oí una oración» y «–Comprendí una oración», pareciera que oír y comprender son procesos que bien pueden suceder al unísono o el primero antes del segundo, como si se ejecutaran dos acciones diferentes. Mientras que oír sucede en algunas partes del cuerpo comprender se interpreta como un inefable *proceso* que tiene lugar en la mente (*GF,* §32.), por el contrario comprender se intenta ubicar en un lugar físico, esto es el

cerebro. A pesar de esto, resulta difícil asegurar que ‘comprender’ sucede en “algún lugar” físico, puesto que funciona como atributo psicológico; en Bennet et al. (2008) podemos leer:

> Desde nuestra perspectiva *es posible* asignar una ubicación a sensaciones como el dolor y el picor. La ubicación de un dolor está donde señala quien lo sufre, en el miembro que se masajea, en la parte del cuerpo que dice que le duele –porque son estas formas de conducta del dolor las que aportan los criterios para la ubicación de éste–. Por el contrario, al pensar, creer, decidir y desear, por ejemplo, no se les puede asignar una ubicación *somática.* La respuesta a las preguntas «¿Dónde pensaste en esto?», «¿Dónde adquirió esta extraña creencia?», «¿Dónde tomó la decisión de casarse? » *nunca* es «En la corteza prefrontal, por supuesto». (p. 177)

Bennet y Hacker dan cuenta de la idea absurda de ubicar la comprensión en el cerebro. No obstante, antes de aceptar esta crítica, que en principio parece superficial, consideremos la idea de que el proceso de comprensión suceda en el cerebro. Si la comprensión sucede en un órgano físico, entonces el proceso de la comprensión se manifiesta de manera que pueda ser estudiado y medido empíricamente, es decir, ha de ser estudiado como un fenómeno. Empero, el resultado de las investigaciones de este tipo no satisface la pregunta por el *qué*, ni el *cómo* del supuesto fenómeno en cuestión, ya que es problemático determinar la configuración neuronal o actividad química que acontece en un persona cuando comprende, ni tampoco verificar que dicha configuración neuronal no está cuando la misma persona no comprende.

Revisemos la gramática de la palabra ‘proceso’ ¿Cuándo hablamos del *proceso* de comprensión y del *proceso* de digestión usamos de la misma manera la palabra ‘proceso’? En respuesta a esta pregunta intuimos que ‘comprender’ sostiene diferencias importantes con ‘digerir’, el verbo psicológico ‘comprender’ se encuentra en la misma situación problemática que el verbo ‘pensar’:

> “El pensamiento ese ser extraño” –pero no nos parece extraño cuando pensamos. El pensamiento no nos parece misterioso mientras pensamos, sino sólo cuando decimos retrospectivamente “¿Cómo fue posible eso?” ¿Cómo es posible que el pensamiento tratara

> de ese *mismísimo* objeto? Nos parece como si con él hubiésemos apresado la realidad. (*IF*, §428)

'Comprender' y 'pensar' se asemejan en cuando prestamos atención a la sensación de extrañeza y nos preguntamos ¿cómo es posible comprender? ¿Cómo sucede la comprensión? El sinsabor de los intentos por responder a estos interrogantes deja entrever el sin-sentido del uso de 'pensar' o 'comprender' como nombres que designan un determinado proceso físico común que se pueda reconocer para discernir exactamente cuando pienso y cuando no, cuando comprendo y cuando no. Los verbos psicológicos no se comportan, ni poseen las mismas particularidades que un verbo únicamente físico, por ejemplo 'digerir'. El proceso de digestión puede ser descrito y, además, señalar el tiempo y los órganos implicados en dicho proceso. Sobre los verbos físicos se identifican tres rasgos: (a) es posible determinar en *qué momento* exacto se realiza la acción y cuando no, así el ser humano puede señalar asertivamente cuando inicia y cuando termina; (b) son acciones que pueden ser interrumpidas y es posible establecer desde qué momento ha sido suspendida la acción y cuando se retoma; y (c) sobre las acciones únicamente físicas es posible realizar una descripción. Otro ejemplo de un verbo físico es 'correr', sobre esta actividad es posible discernir absolutamente cuando se corre y cuando no, así mismo cuando se deja de correr y cuando se realiza de nuevo la actividad, y, por último, el proceso de correr puede ser descrito y definido, ya que podemos describir los movimientos del cuerpo.

Si aceptáramos que 'comprender' es análogo a los verbos físicos debemos preguntarnos ¿cómo es posible determinar el momento en que una persona empieza a comprender? La pregunta se respondería: «−Una persona comprende cuando realiza dicho proceso−», «−Un sujeto comprende cuando *aquel proceso* cerebral acompaña a otra acción», pero de nuevo ¿cuál es ese *proceso*? ¿Cuándo se lleva acabo dicho *proceso*? Wittgenstein señala estos dilemas en analogía con el juego de ajedrez:

> Supongamos que se pregunta: ¿en qué momento sabes cómo jugar ajedrez? ¿Siempre? ¿O solamente mientras dices que puedes hacerlo? ¿O solamente mientras mueves una pieza? ¡Qué raro que saber jugar ajedrez requiera tan poco tiempo y que una partida requiera tanto! (San Agustín: "¿*cuándo* mido un periodo de tiempo"). (*GF,* §12)

Con este ejemplo y la gramática de la palabra 'proceso' podemos seguir preguntando: ¿acaso después de enunciar «–He comprendido las reglas del ajedrez» puedo dejar de comprenderlas por dos horas y después de las dos horas de nuevo he vuelto comprender o que solo las comprendo cuando juego ajedrez?[27] Ahora bien, en relación con una oración no es posible indicar empíricamente el inicio y el fin de la comprensión de una expresión o medir el tiempo que se tarda en entenderla:

> ¿*Cuándo* entendemos una frase? ¿Cuándo la hemos pronunciado entera? ¿O mientras la pronunciamos? ¿Es el comprender un proceso articulado como el decir una frase, y corresponde su articulación a la de la frase? ¿O es inarticulado y acompaña a la frase a la manera en que un golpe de pedal acompaña la melodía? (*GF*, §12)

Las anteriores preguntas podemos aplicarlas a nuestro ejemplo usado en el principio de este capítulo. Cuando se traduce una oración del inglés británico al español ¿podríamos decir que P ha comprendido cuando finaliza la traducción de la oración o cuando P ha enunciado «–Ahora comprendo»? ¿La comprensión es un proceso a la par de la enunciación de una oración o a la par de escuchar/leer una oración traducida? Cuando formulamos estas cuestiones erramos porque presuponemos que la comprensión no se vincula a los casos particulares en los cuales decimos comprender; esto mismo sucede cuando afirmamos que *pensar* es un proceso en paralelo o anterior a enunciar una oración/oír una oración[28].

Bennet y Hacker entienden que la raíz de confundir los verbos de acción y los verbos psicológicos es consecuencia de caer en la *falacia mereológica:* "(…) atribuir a las *partes* constituyentes de un animal atributos lógicamente aplicables sólo al animal como un *todo*" (Bennett et al., 2008, p. 38). La falacia mereológica es cometida por psicólogos, neurocientíficos, filósofos e investigadores en el campo de la IA que conceden los predicados psicológicos únicamente al cerebro, gracias a una trastocada herencia cartesiana. Digo "trastocada" porque se aplica de manera irreflexiva los predicados psicológicos al cerebro, mientras que Descartes atribuía estos fenómenos a la mente inmaterial, a pesar de la diferencia entre mente y cerebro, a veces, estos dos términos se usan como idénticos. Ahora

[27] *GF*, §1: "¿podemos decir también que hemos comprendido media proposición?"
[28] El desarrollo del problema que sugiere el verbo psicológico 'pensar' se puede esclarecer con la lectura de: Gil de Pareja (1988); Hutchinson (2010); Lin (2019).

bien, podemos preguntar a los investigadores que caen en la mencionada falacia: ¿existe eso del pensar, creer, desear, comprender, del cerebro? ¿Los cerebros piensan y razonan como el ser humano lo hace? La aplicación o la negación de dichos predicados psicológicos al cerebro resulta ser un sinsentido, en tanto el cerebro no se *comporta,* por ejemplo de manera indecisa y tampoco toma decisiones o "El cerebro ni ve *ni es ciego,* del mismo modo que los palos y las piedras no están despiertos, *pero tampoco están dormidos*" (Bennett et al., 2008, p. 37). El *ser humano* es quien se *comporta como tal* que puede decirse sobre él que: decide, pregunta, comprende, porque presenta una ratificación de estos predicados en su comportamiento:

> Los neurocientíficos, psicólogos y científicos cognitivos que adoptan estas formas de descripción no lo hacen como resultado de unas *observaciones* que demuestren que el cerebro piensa y razona. Susan Savage-Ramnaugh ha reunido pruebas concluyentes que demuestran que los chimpancés bonobo, debidamente entrenados y enseñados, pueden plantear y responder preguntas, razonar de forma rudimentaria, dar y obedecer órdenes, etc. La prueba está en su comportamiento, en lo que hacen en sus interacciones con nosotros (incluida la forma que tienen de emplear símbolos). (Bennett et al., 2008, p. 35)

Cuando usamos términos en sentido nominal tales como 'proceso' o 'mecanismo' para referirnos a 'comprender' o 'pensar' caemos en callejones sin salida, no lleva a sin-sentidos. Un proceso cerebral que se active o proceda en determinado momento no resulta ser criterio para determinar si alguien comprende la expresión en un juego de lenguaje o no. En definitiva, la idea de un mecanismo físico que suceda en el cerebro no es necesaria ni suficiente para explicar los fenómenos psicológicos. Quienes apuntan a descubrir un proceso que pueda ser llamado comprensión y no encuentran un correlato material para tal proceso hallan una salida al asunto en la pretendida inefabilidad de la mente, en palabras de Wittgenstein:

> *IF*, §36: […] dado que no podemos indicar *una* acción corporal que podamos llamar el señalar la forma (en contraposición, por ejemplo, al color), decimos que a esas palabras les corresponde una actividad *mental.* Allí donde nuestro lenguaje nos hace suponer que hay un cuerpo, y no hay ninguno, allí quisiéramos decir que hay un *espíritu.*

> *IF,* §196: El empleo incomprendido de la palabra explica como expresión de un extraño *proceso*. (Del mismo modo como se piensa el tiempo como un extraño medio y el alma como una extraña esencia).

Ahora bien, la noción de 'mecanismo mental' no refiere a un asunto empírico. La ciencia construye modelos causales y describe las relaciones causa-efecto, de manera que, formula hipótesis que pueden ser verdaderas o falsas según algunos criterios de verdad, pero la mosca cae en la botella cuando intentamos describir un *proceso* mental tal como describiríamos un estado de cosas o una relación causal (*IF,* § 308). Esto es, se aplica el modelo de un proceso físico-mecánico como es digerir al llamado proceso mental de la comprensión.

El trabajo de la filosofía frente al fenómeno mental se defiende bajo el hecho de que, cuando el sujeto B, por ejemplo, comprende una orden se reconoce un cambio, un cambio de estado mental que sucede dentro de tal sujeto. De modo que, la filosofía tradicional retorna a las categorías cartesianas: el ser humano constituido por dos sustancias alma/mente y cuerpo. Sin embargo, preexiste la falta de respuesta a la pregunta que el mismo método tradicional se plantea sobre la comprensión, a saber ¿*qué* es la comprensión?, puesto que se considera responder sobre este fenómeno que sucede en el ser humano a partir de las mencionadas categorías y la respuesta son las mismas categorías. Por otra parte, si la especificidad de la comprensión se sitúa en la mente, la conclusión desde la filosofía tradicional y la psicología empírica propone un camino para abordar el fenómeno en cuestión: el análisis introspectivo. La introspección como método da lugar al sujeto como observador de sí mismo, esto es, el sujeto que "ve" a su interior, que se "ve" asimismo[29] y después lleva al lenguaje lo que ha "descubierto". Estas observaciones sobre psicología desbordan los objetivos de estas líneas, pero son un abrebocas para mostrar la diferencia de un filósofo que construye teorías frente a un filósofo gramatical a quien le resulta irrelevante la teoría mentalista para explicar la 'comprensión'.

Cabe subrayar que nada de lo expuesto hasta aquí niega la existencia de fenómenos mentales en el sentido de que son reducibles a la dimensión empírica, ni tampoco la defensa de algún tipo de materialismo radical o conductismo. La lectura wittgensteniana intenta

[29] Estos problemas se desarrollan en *FPF* y Últimos *Escritos sobre Psicología* o en la segunda parte de *IF*.

develar los "problemas insolubles" en los que se incurre cuando se usa términos psicológicos del mismo modo que expresiones que pertenecen a otro juego de lenguaje, en este caso empírico/científico, y llega a conclusiones a partir de tal equívoco. Tampoco desde Wittgenstein se descartan los intentos de las ciencias por construir modelos explicativos sobre la base de investigación empírica que puedan dar cuenta de la actividad cerebral. Antes bien, la idea que se quiere combatir es que 'comprender' sea una actividad oscura a la que sólo puede acceder el sujeto por medio de la introspección[30] y que la comprensión, diferente de los procesos cerebrales, no puede ser objeto de percepción inmediata (*GF*, §41).

2.1.4 Observaciones a la tesis de la comprensión como 'estado mental'

La acción de comprender resulta ser diversa: comprendemos órdenes, gestos, acontecimientos, proposiciones, oraciones, etc. Ahora bien, prestemos atención a la comprensión de una proposición. Consecuencia de la tesis esencialista del significado es posible inferir la posición que defiende la necesidad de comprender en primer momento el significado de cada uno de los términos individuales de una proposición para lograr la comprensión de la proposición en cuestión, dicho en otras palabras, la compresión emerge cuando se "suman" los significados que componen la totalidad de la proposición. Partiendo de esta idea cuando un sujeto afirma que ha comprendido la oración: «Es el caballo de Miguel Páramo, que galopa por el camino de la Media Luna», primero comprende el pensamiento detrás de "Es", luego el pensamiento detrás de "el", seguido del pensamiento de "caballo" de modo que se continué así con los demás elementos hasta finalizar la proposición, para proseguir a sumar aquellos pensamientos y afirmar que ha comprendido la oración. Sobre esto Tomasini (2006) llama la atención:

> Le preguntamos a alguien que nos diga que fue lo que entendió cuando entendió el significado de la oración 'Cantinflas era mexicano'. Obviamente, supongo, nos dirá algo como: hubo una persona, conocida como 'Cantinflas', nacida en un país llamado 'México' y esa persona ya no vive'. Dejando de lado la respuesta concreta que se obtenga, lo interesante para nosotros es que la explicación de lo que parecía ser un fenómeno de comprensión viene

[30] Sobre el problema que suscita la psicología introspeccioncita y la posibilidad del conocimiento, por ejemplo, del dolor del otro o demás sensaciones que se plantean como privadas para el sujeto que experiencia el mundo, véase: *FPF*, II.

> dada en palabras. O sea, el lenguaje toma el lugar y desempeña el rol de lo que supuestamente pasaba dentro del sujeto y que en la tradición parecía esencial. Empero, nosotros empezamos a entender que la explicación de la comprensión no nos saca del lenguaje, no nos obliga a mirar hacia lo extra-lingüístico. (p. 99)

Tomasini señala dos aspectos en sus reflexiones: por un lado, retrata un golpe claro a la posición fregeana de composicionalidad, a saber, la proposición como un elemento complejo constituido por elementos más simples, cada elemento posee un significado parcial; gracias a los significados parciales la proposición tiene sentido –esta hipótesis sobre la comprensión corresponde a lo que llamamos comprensión superficial líneas atrás–; por otro lado, y más importante para nuestro objetivo, Tomasini resalta la absurda búsqueda de *algo* detrás de las palabras que justifique el significado de un enunciado, *algo* que brinde soporte al lenguaje.

En contra de Tomasini podrían objetar que cuando se declara «–Ahora comprendo» es fácil caer en la tentación de suponer que la comprensión es algún tipo de estado mental, por ejemplo, la posesión de algún tipo de pensamiento o imagen mental, puesto que, de alguna manera se quien comprende "siente" que ha comprendido. En otras palabras, antes de que el sujeto B entienda, por ejemplo, una orden, estaba en un estado mental y cuando comprende supone estar en un estado mental diferente, en un estado de comprensión. A quien defiende esta hipótesis y busca un estado mental denominado 'comprensión' preguntemos: ¿nos pasó lo mismo cuando entendimos que 2 + 2 = 4 que cuando entendemos la orden: «–Elige una flor roja»? ¿Se posee el mismo estado mental cuando comprendo una operación numérica que cuando comprendo una orden?

Aunque se acepte las ideas del segundo Wittgenstein en las que describe el lenguaje como juegos de lenguaje y el significado emparentado con el uso, se podría rebatir a Wittgenstein que la fuente de la capacidad de comprender y usar los signos correctamente es un estado mental. Con Frege podríamos decir que el pensamiento es lo que se "viene a la mente" cuando un sujeto oye o enuncia una expresión lingüística, lo que sugiere que la expresión muestra dicho pensamiento y este estado es la comprensión. Conforme a esto, para que exista comprensión algo debe acaecer ¡–Debe ocurrir un cambio cuando un sujeto comprende! ¡–Ha de ser un cambio de estado mental! Negar que existe tal hecho es:

> […] to claim that our infinite ability to use language cannot be explained in terms of a finite, generative base; it would be to suppose that the difference, between someone who understands a language and someone who does not, does not lie in the presence or absence, respectively, of a finite internal state which is the source of the infinite capacity. (McGinn, 1997, p. 88)

De acuerdo con este razonamiento, el estado mental de la comprensión es el criterio para juzgar si el otro comprende o no y el criterio para el correcto uso de las expresiones. Entonces, el estado mental es quien obliga al usuario del lenguaje a utilizar una expresión a cambio de otra. No obstante, aunque cabe la posibilidad de la existencia de un vínculo entre imágenes mentales, pensamientos u otros elementos que acaezcan en un estado mental con el uso de las expresiones, este vínculo depende de una consonancia entre el uso y, por ejemplo, una imagen en la práctica. Pongamos por caso la imagen de una flecha: → , si veo esta imagen en un camino decido ir hacia la derecha, pero de ninguna manera es porque gratuitamente evoque en mí una flecha o cambie algo dentro de mí que corresponda a la acción de tomar la derecha, sino porque usualmente esta flecha en la práctica se usa para indicar una dirección, sin embargo, no toda vez que vea una flecha y "pase algo dentro de mí" tomaré el camino de la derecha, esto dependerá del lugar al que deseo llegar en cada caso particular. La regularidad en la práctica es la que empuja al sujeto que mira tal signo para que se comporte de un modo determinado, de lo contrario tendría que captar algo más que la regularidad antes del uso de dicho signo.

En el anterior apartado discutimos sobre la comprensión como mecanismo físico y dedicamos pocas palabras a la comprensión como un mecanismo interno mental, atendamos ahora otra idea en relación a este último: la comprensión como estado mental, y agreguemos el adjetivo 'consciente', ya que cuando comprendo, por ejemplo, una proposición, *siento* que la comprendo, es decir, *soy consciente* de que comprendo. Para esto optemos por seguir con el método analógico y comparemos, por ejemplo, los términos 'estar deprimido' y 'comprender'. Si se asume 'estar deprimido como estado mental, de nuevo podemos hacer uso de la falacia mereológica que subraya Bennet y Hacker, dicho estado mental no es localizable en una parte específica del cuerpo, sino que 'estar deprimido' se predica de la

totalidad de un ser humano. En esta característica ‘estar deprimido’ y ‘comprender’ se asemejan. Con respecto al sujeto que padece ‘estar deprimido’ –si aceptamos que se trata de un estado mental consciente– puede determinar la duración de tal estado, sin embargo, como ya hemos dicho, la comprensión resulta ser escurridiza para la medida de su temporalidad, de esto que ‘estar deprimido’ y ‘comprender’ en este punto se separen. Quizá la salida a este embrollo resulte ser que la comprensión es un estado mental *inconsciente*, ya que comprendemos “de golpe” y no podemos dar cuenta de ningún proceso dentro de la figura del mecanismo causal, ni de una descripción de tal estado. No obstante, predicar de la comprensión como un estado inconsciente, sumerge, aún más, en la oscuridad la aspirada explicación y el ejercicio de la filosofía tradicional queda en entredicho. Sobre esta solución al problema construido por la figura esencialista Wittgenstein señala: […] Nada se prestaría más a una confusión aquí que el uso de las palabras “consciente” e “inconsciente” para el contraste entre estado de conciencia y disposición. Pues ese par de palabras oculta una diferencia gramatical (*IF,* §149), esta diferencia se desarrollará en 2.2.

2.2 Ser capaz o tener la habilidad

En este apartado no pretendemos definir la comprensión señalando *algo* que nos ayude a salir del apuro, por el contrario, se explicará la comprensión profunda como un estado de disposición.

Cuando preguntamos a un sujeto cómo ha comprendido, por ejemplo, un problema algebraico, el sujeto puede responder enunciando los pasos que realizó para llegar a la solución, podríamos utilizar su respuesta a favor del argumento como un proceso o un mecanismo, sin embargo, a veces no puede responder de este modo, por ejemplo cuando se comprende una lectura, ¡Simplemente no *sabe* cómo ha comprendido de repente! Respecto a esto, Ericsson y Simon (1980) determinan dicho momento de iluminación como una experiencia inesperada, pero que se ha desarrollado gradualmente, la sorpresa deviene del ocultamiento del proceso, pero sigue siendo un proceso de comprensión sensible a la descripción. Ahora bien, la comprensión “de golpe” no despierta interés en la IA, en la medida que no se pregunta por la naturaleza de la experiencia, por ejemplo, de asombro que B tiene cuando de repente sabe cómo continuar una serie numérica; más bien, la IA quiere

encontrar la descripción que corresponde a los procesos en el pensamiento que causan tal asombro. El término de 'proceso' se considera como una cadena lineal y causal de, para seguir con nuestro ejemplo, operaciones que B realizó para continuar con la serie correctamente, a veces B es consciente de los pasos que realizó, otras veces no. Esto lo describe De Groot (1965) como 'procesos cerebrales imperceptibles' de los cuales el sujeto sólo percibe el resultado final, logrando asombrarse de sí mismo.

Siguiendo a Groot, de golpe B ha tenido un "momento de iluminación" en el que ha reconocido un patrón o B ha construido una fórmula para continuar, y, después se dice así mismo «–Ahora sé cómo continuar». No obstante, sea cual sea el caso, quien sostiene que la comprensión depende de un estado mental o proceso arguye que B comprendió porque "se le ocurrió la fórmula" o "descubrió un patrón" de manera inconsciente o consciente. Sin embargo, por un lado, la conclusión «–B comprende» se infiere a raíz de un comportamiento observable, por ejemplo, B ha podido continuar la secuencia y ha escrito el número correcto después del 29; por otro lado, B se dice así mismo que comprende en cuanto se siente seguro de continuar. El error que señala Wittgenstein en este ejemplo es confundir la gramática de «B comprende» «B sabe cómo seguir» o «B puede seguir» con «B tiene una fórmula que procede de un estado mental consciente/inconsciente que a su vez le permite continuar». Cabe aclarar, la réplica wittgensteniana no apunta a la existencia o no de un estado mental en el que a B se le ocurra la fórmula para continuar con la serie, sino que, aunque posea tal formula puede ser el caso de que no *sea capaz* de aplicarla, que no *pueda* aplicarla, entonces:

> La comprensión del lenguaje, como la de un juego, parece como un **trasfondo** en relación con el cual una oración particular adquiere significado. Pero esta comprensión, el conocimiento del lenguaje, no es un estado de conciencia que acompañe a las oraciones del lenguaje. Ni siquiera en el caso en que una de sus consecuencias fuera un estado de ese tipo. Es más bien la comprensión, como el dominio de un cálculo, algo similar a *poder* multiplicar. (*GF,* §11). (Negrilla mía)

Cuando decimos «−*Sé* lo que quiere decir esta oración», aquí 'saber', como ya se ha mencionado atrás, no es tener posesión de una esencia determinada, más bien es *ser capaz de hacer algo,* dicho en otras palabras, es un *estado de disposición.* En la medida que

relacionamos la comprensión a un estado disposicional las preguntas que hemos formulado una y otra vez desde la tesis esencialista del lenguaje resultan ser un sin-sentido, y dejan de ser un problema. Con el fin de aclarar esta idea proponemos revisar la gramática de 'estar capacitado'. 'Estar capacitado' para poder aplicar *x* palabra por lo general no indica describir una idea general de *x.* En cambio, señala casos particulares en los que se pueda aplicar el término en cuestión:

> Se piensa el significado como algo que se nos *ocurre* al oír una palabra.
> Lo que se nos ocurre al oír una palabra caracteriza en cualquier caso el significado. Pero lo que se me ocurre es un ejemplo, un caso de aplicación de la palabra. Y este ocurrírseme (*vorshweben*) no consiste propiamente en que esté presente una determinada idea siempre que pronuncio u oigo una palabra, sino en que cuando se me pregunta por el significado de la palabra se me *ocurren (einfallen)* aplicaciones para ella. (*GF*, §75)

Sumado a esto, sobre la capacidad podemos identificar dos características: primero, no desaparece en tanto se agota el acto, por ejemplo, cuando afirmo: «−Soy *capaz* de lanzar el balón más de dos metros», y, efectivamente, realizo la acción que corrobora tal enunciado como verdadero, inmediatamente he acabado la acción no significa que ahora no sea capaz de lanzar el balón más de dos metros. La capacidad permanece. En el caso del lenguaje la capacidad del individuo para comprender una expresión permanece en tanto se ha instruido, aprendido y usado una expresión, de modo que, estará preparado para aplicarla en todo momento. Cada vez que en un acto particular el sujeto aplique determinada palabra *demuestra* que efectivamente la ha comprendido y es actualizado su modo de empleo. Segundo, la capacidad es un horizonte que se expande. Por ejemplo, cuando pedimos explicaciones a un niño sobre los números no serán tan complejas como la respuesta de un matemático (Tomasini, 2006). Cabe señalar que, las explicaciones y las descripciones mediante ejemplos de un niño o un matemático se rigen dentro de un marco regulador de aprendizaje y aplicación de usos en un juego de lenguaje. Así, los límites de la capacidad se ensanchan –recordemos las dinámicas del juego de lenguaje que expusimos en 2.1.1–

En consecuencia, el criterio de comprensión no es el acaecimiento de un estado mental, sino las circunstancias particulares; B comprende porque es capaz de continuar la

serie correctamente gracias a que, por ejemplo, ha sido entrenado en problemas algebraicos, de forma que posea una habilidad en reconocer patrones numéricos. En efecto, deja de importar lo que sucede "dentro" de B para prestar atención al trasfondo en el que ha sido instruido y se revela en su conducta. Conforme a esto, en el caso del lenguaje, Wittgenstein vincula la comprensión profunda con una habilidad, una capacidad o el poder de aplicar una expresión en determinadas circunstancias. En vista de esto, en cualquier ocasión que enunciemos «–Ahora comprendo» dependerá de las circunstancias en que ocurra la enunciación: cuando entiendo una formula, resulta que puedo aplicarla; cuando comprendo una palabra, resulta que la puedo usar en oraciones; cuando comprendo una señal de tránsito es porque sé que debo hacer o no con mi vehículo, etc., y esto sucede en tanto he sido instruido para *ser capaz* (*GF*, §81).

Es así como el comportamiento verbal y no-verbal es el criterio de justificación para que decir del otro que entiende o no[31]. Cabe decir que el uso de las expresiones no es azaroso, ni el comportamiento lo es, pues se definen por las reglas del juego del lenguaje aprendidas en el uso práctico de las mismas. La propuesta wittgensteniana muestra que la comprensión profunda del uso de una palabra comienza en la experiencia de la regularidad de las reglas gramaticales aplicadas a los casos particulares. Cuando se asimila tal experiencia como parte de mi lenguaje, puedo decir que sigo una regla. A su vez seguir una regla es únicamente posible en la práctica del lenguaje; por lo tanto, seguir una regla se equipara a dominar una técnica (*IF*, §199) y se logra dominar la técnica cuando se comprenden y emplean en diferentes situaciones y contextos las reglas gramaticales particulares de tal juego:

> (…) la comprensión consiste en que tengo una experiencia determinada –Pero que esta experiencia sea comprensión consiste en que dicha experiencia es una parte de mi lenguaje" (*EM*, p. 37).
>
> Si bien la comprensión de una expresión es su aplicación ¿cuál es la justificación que avala que yo diga: —Ahora comprendo? ¿Qué muestra que el uso de una palabra sea "correcto"?

[31]Quienes leen en clave conductista a Wittgenstein, en general son lecturas sesgadas y aisladas. En contra del conductismo Wittgenstein replica en *IF* §312, allí no niega la existencia de fenómenos mentales, más bien llama la atención al cuidado con el que se debe conocer "más de cerca" un proceso, idea que se entiende con las observaciones de 3.1.3, a saber, el equívoco intento de estudiar los procesos mentales a la luz de otros tipos procesos, por ejemplo los mecánicos, de los que tenemos información.

> Al parecer el criterio de aplicación correcta es la práctica habitual de las reglas de 'uso'. (*IF I*, §199-201)

Las reglas de un juego del lenguaje muestran cómo y cuándo se usan las expresiones, además de estar implícitas en cada uso. Ahora bien, una réplica a esta noción de regla se puede resumir así: no es posible seguir una regla si antes no se ha comprendido dicha regla. No obstante, enunciar explícitamente una regla no es necesario, puede seguirse en tanto se ha observado a otros actuar y mi actividad sea acorde con ella, es decir, en la *praxis* del lenguaje se expresa las reglas y se siguen las reglas:

> ¡Pensemos en qué caso decimos que un juego se juega según una regla definida!
> La regla puede ser un recurso de la instrucción en el juego. Se le comunica al aprendiz y se le da su aplicación. –O es una herramienta del juego mismo. –O: Una regla no encuentras aplicación ni en la instrucción ni en el juego mismo; ni es establecida en un catálogo de reglas. Se aprende el juego observando como juegan otros. Pero decimos que se juega según tales y cuales reglas porque un espectador puede extraer estas reglas de la práctica del juego – como una ley natural que sigue el desarrollo del juego. (*IF*, §54)

Es así que preguntarse por "¿cuánto **tiempo** antes de obedecer la orden *debes* haberla comprendido?" (*GF* I, §8) o ¿cuánto **tiempo** antes de seguir una regla *debes* haberla comprendido?, resulta ser un sin-sentido. Oraciones tales como: «–Debes comprender la orden antes de poder actuar conforme a ella» o «–Debes comprender una regla antes de poder actuar conforme a ella» tiene un sentido que no es metalógico:

> Si "comprender una oración" quiere decir actuar de alguna manera de acuerdo con ella, entonces la comprensión no puede ser la condición para que actuemos de acuerdo con ella. Pero la experiencia bien puede mostrar que la conducta específica de la comprensión es una condición previa para la obediencia de una orden. (*GF*, §8)

Aunque en el capítulo II se ha desplegado *inextenso* la idea de las reglas de juego de lenguaje, es de importancia resaltar que "Tales reglas no anticipan o determinan el juego de lenguaje; no surgen de una convención anterior, sino que son construcciones de la práctica lingüística,

de la práctica del lenguaje" (Hamilton, 2017, p. 48). La *praxis* asegura la normatividad que a su vez es la piedra angular para soportar el carácter público del lenguaje. La lectura de Norman Malcom (1989) retrata el vínculo entre la existencia de reglas y su utilización en una comunidad, además muestra que el significado de las palabras es independiente de cualquier persona particular en un juego de lenguaje:

> To speak a language is to participate in a way of living in which many people are engaged. The language I speak gets its meaning from the common ways of acting and responding of many people. I *take part* in a language in the sense in which I *take part* in game –which in surely one reason why Wittgenstein compared languages to games. Another reason for this comparison in that in both languages and games there are *rules*. To follow the rules for the use of an expression is nothing other than to use the expression as it is ordinarily used–which is to say, as it is used by those many people who take part in the activities in which the expression is embedded. Thus the meaning of the expression is *independent* of me, or of any particular person; and this is why I can use the expression correctly or incorrectly. It has a meaning independent of *my* use of it. And this is why there in no sense in the supposition that a forever-solitary person could know a language, any more than he could buy and sell. (p. 22)

El juego del lenguaje y sus reglas se fundan en las prácticas tanto lingüísticas como extra-lingüísticas con los otros, estas determinan el criterio de corrección en el uso de las expresiones y los criterios conductuales de la comprensión. De ahí que, el intento de engranar un mecanismo, un proceso o determinar un estado mental resulta corto para describir todos los patrones de comportamiento y usos involucrados en las formas de vida. De esto que, el trasfondo, las prácticas (formas de vida) sean clave importante para la comprensión, en palabras de Hacker (2013):

> El dominio de un lenguaje, de la gramática de las palabras, se realiza en la *praxis*, equivale al aprendizaje de un amplio abanico de modos de actuar y de actividad y de modos de reacción y respuesta al habla y a las circunstancias, modos que son constitutivos de una forma de vida. (p. 31)

2.3 Formas de vida («Lebensform», «Lebensformen», «Form des Lebens»)

En objeción al desarrollo de la comprensión entendida como estado disposicional y acto, se puede esgrimir: «−No se han liberado de nuestra crítica, porque la noción de 'comprensión profunda' se ha fundado en el sujeto que *interpreta* los signos/las expresiones, en otras palabras, *interpreta* la regla de uso a su antojo». Con esto se asume la existencia de un elemento mediador entre el signo y la compresión, a saber, la interpretación. Es evidente que a veces damos a los signos una interpretación; no obstante, esto no involucra que toda expresión o signo requiera de una interpretación: cuando oigo el timbre no se da en mi interior un trabajo de interpretación, sencillamente reacciono ante lo que oigo, lo mismo sucede "Si alguien saca un cuchillo y lo dirige hacia mí, no digo: interpreto esto como una amenaza" (*GF*, §9). Conforme a la idea del refutador, si alguien saca un cuchillo mientras camino por la calle y se dirige hacia mí la conexión inferencial/causal sería: cuchillo + sujeto + camina hacia mí = interpreto que está amenazándome, por lo tanto, saldré corriendo. Sin embargo, no sucede así, más bien reacciono inmediatamente de la manera en que he sido adiestrada ante esos signos, es decir, si bien hay un espacio para una posible interpretación, no es azarosa, las reglas del juego de lenguaje guían mi comportamiento a modo de "un poste indicador, solamente, en la medida de que haya un uso constante, una costumbre" (*IF*, §198). De modo que, las acciones del ser humano no están desligadas de las reglas o regularidades, por lo tanto tampoco son abstracciones respecto a la comunidad con la cuál repite, construye y actualiza el juego de lenguaje, ni una interpretación privada. Así, las reglas constituyen un marco normativo para las acciones, con esto no quiero decir que se acepte un determinismo, porque quienes actualizan las reglas son los seres humanos en la práctica, en sus modos de vivir.

Wittgenstein usa la analogía entre el lenguaje y el juego de lenguaje con varios fines, en este apartado quiero resaltar que el juego de lenguaje es intrínsecamente un conjunto de prácticas, *algo que se hace,* en otras palabras, son las acciones que se ven inmiscuidas con las reglas. A estas acciones Wittgenstein da el nombre de: *formas de vida*. El lugar que ocupa las formas de vida ha sido analizado en extenso por diferentes autores[32], y pretender aquí un

[32] Véase: Padilla Gálvez y Gaffal, 2013.

análisis significaría embarcarnos en otro problema que excede nuestras páginas, a causa de esto y el ya mostrado vínculo entre formas de vida (practicas) y comprensión profunda proseguimos a comentar un problema recurrente sobre las formas de vida.

2.3.1 Formas de vida, comprensión profunda y representación perspicua

El poco desarrollo que tiene *Lebensform* dentro de la obra wittgensteniana anima la controversia de su lectura, gracias a esto la discusión ha suscitado digresiones importantes en la literatura wittgensteniana, dejando aún el debate abierto. Sin embargo, los diferentes lectores de Wittgenstein están de acuerdo en el rechazo ante la trivialización del concepto en cuestión, puesto que se hace evidente el hecho de que la vida de las personas y el lenguaje se relacionan. Sin embargo, 'formas de vida' resulta ser un concepto amplio que puede ser entendido como lo que se encuentra en un nivel más profundo del conjunto ilimitado de las actividades humanas. Escuchar y entender a alguien es algo más que oír cómo habla el mismo idioma con palabras que ambos conozcan, ya que, entender evidencia una estructura de formas de vida específicas, por ejemplo: el comportamiento o la instrucción, además que la afirmación de la comprensión profunda que tiene el otro se logra porque el juego del lenguaje compartido es la condición de posibilidad.

Entre los diversos debates generados entorno a estos conceptos co-existen dos interpretaciones que llaman la atención sobre el problema de la naturaleza múltiple o unitaria de las formas de vida[33]. Por un lado, quienes sostienen que a cada juego de lenguaje le pertenecen específicas formas de vida; por lo tanto, si se acepta la existencia de múltiples juegos de lenguaje así mismo se debe aceptar diferentes formas de vida. Por otro lado, una única forma de vida humana entendida como el trasfondo común a todos seres humanos. En sobremedida quienes defienden esta última se fundamentan en *IF,* §206: "el modo de actuar humano común es el sistema de referencia por medio del cual interpretamos un lenguaje extraño", en otras palabras, la forma de vida humana representa algún tipo de naturaleza humana (Glock, 1996; Hacker, 2013). Contrario a esta idea leemos *IF* §23 y distinguimos que, si bien una forma de vida no es el juego de lenguaje, lo acompaña. Sumado a esto, si tenemos en cuenta la expresión "imaginar un juego" es "imaginar una forma de vida", nos

[33] Dos artículos que recogen la discusión son: Lagerspetz (2020) y Giraldo, (2017)

inclinamos a pensar que, si existe multiplicidad de juegos de lenguaje, también encontraremos multiplicidad de formas de vida. Las observaciones que con anterioridad hemos realizado sobre el significado, las reglas del juego de lenguaje y la comprensión, son razones para dudar de que Wittgenstein pensara en la **naturaleza** de alguna cosa entendida como la **esencia**, algo así como una esencia humana.

Ahora bien, por un lado, si acogemos las ideas de Hacker y Glock, es decir que las formas de vida desvelan algún tipo de naturalismo tenemos que preguntar: ¿es posible que los sistemas artificiales inteligentes logren simular la naturaleza humana entendida como la forma de vida humana? Por otro lado, si rechazamos cualquier tipo de naturalismo, es decir que asentimos ante la hipótesis de que existen tantas formas de vida como juegos de lenguaje, unas tan disímiles de otras, tendríamos que simular la complejidad de la vida humana tanto lingüística como extra–lingüística en un sistema artificial inteligente. Sobre esto, pienso que, si la 'naturaleza humana' se entiende como una *disposición* (capacidad de imaginar, de creer, de seguir una regla, de comprender, de ver conexiones) es posible que sobre las dos lecturas que se desarrollan desde el concepto de *formas de vida* se puede entablar el diálogo.

Por otra parte, la compresión profunda de las expresiones en un lenguaje consiste en *poseer* un conocimiento de la gramática profunda y el léxico de un lenguaje, tener la habilidad de usar las expresiones en diversos contextos, aunque estos no sean idénticos, y la capacidad de entrever representaciones perspicuas. Entonces, participar en un juego de lenguaje es algo más que tener una cantidad de información, significados o relaciones causales. De esto que, B aunque pueda traducir correctamente: «– ¡That was bloody deliciuos!» no significa que entienda y comparta las *formas de vida* por ejemplo británicas o, más aún, que entienda a los británicos. B comprende, pero superficialmente:

> Eso se experimenta cuando uno llega a un país extraño con tradiciones completamente extrañas; e incluso cuando se domina la lengua de un país. No se entiende a la gente. (Y no porque uno no sepa lo que se dicen a sí mismos.) No podemos apreciarnos en ellos. (*FPF* 719)

El individuo para comprender el significado de las expresiones se inmiscuye en las formas de vida, actúa dentro de un marco de referencia. En palabras de Wittgenstein, comprender el

significado de una palabra o expresión es "(…) entender su uso en un juego de lenguaje, aprender a aplicarla" (*GF*, §10). La lógica de la teoría esencialista, piedra angular de gran parte de la IA, no presenta la posibilidad de entender el lenguaje y la comprensión de este modo, olvida la capacidad humana de la 'visión perspicua':

> Una fuente principal de nuestra falta de comprensión es que no vemos el uso de nuestras palabras perspicuamente. –A nuestra gramática le falta perspicuidad. –La representación perspicua facilita la comprensión que consiste efectivamente en que nosotros 'veamos las conexiones'. De ahí la importancia de encontrar y de inventar conexiones intermedias" (*IF*, §122).

Por último, retomamos una cita de nuestro primer capítulo:

> ¿Calcula la calculadora? Imagina que hubiera aparecido por casualidad una calculadora, y que alguien aprieta casualmente sus botones (o un animal anda sobre ellos) y ella realiza el producto 25 x 20. Quiero decir: es esencial a la matemática que sus signos se usen también en lo civil. Es el uso fuera de la matemática, es decir, el significado de los signos, lo que convierte en matemática el juego de signos (*OFM*, 5,§ 63)

Para Wittgenstein la calculadora no calcula por más que presente los resultados correctos, un cálculo y la actividad de calcular del ser humano esta contextualizado dentro de las formas de vida humana. Para que una máquina comprenda –atendiendo a la comprensión profunda– debe existir un trasfondo, una función determinada por la vida humana y una capacidad para ver vínculos dentro del juego de lenguaje. Esto no quiere decir que el estudio científico cerebral no sea importante o no haya avanzado sobre el papel que desempeña el cerebro en la acción humana, más bien que 'calcular' o 'comprender' no se reduce a un proceso cerebral o un proceso algorítmico como su causa, son acciones/capacidades complejas dentro de las formas de vida. Es así que, aunque se encuentre un "proceso cerebral" –el cual sirva de modelo para los sistemas artificiales inteligentes– no se infiere que 'comprender' haya sido explicado o entendido.

3. Conclusiones

A lo largo del presente texto se desplegaron observaciones sobre la posibilidad de que sistemas artificiales inteligentes sean usuarios del lenguaje natural, esto a partir de una investigación gramatical y en consonancia con las particularidades del lenguaje que Wittgenstein presenta y describe en sus escritos intermedios. En específico nuestro objetivo principal era: clarificar el enunciado «–El sistema artificial inteligente *x comprende* lo que enuncia y *comprende* lo que otro ha dicho». Al final de nuestro análisis podemos afirmar que, en tal enunciado se teje y expresa una confusión gramatical consecuencia de buena parte de equívocos en el fundamento de las teorías psicológicas y modelos cognitivos –de los que se nutre la IA–. La confusión gramatical radica en la diferencia de usos que puede tener, por un lado, la expresión 'seguir una regla', y, por otro lado, el verbo 'comprender', cuando no diferenciamos tales usos y tampoco prestamos atención a su gramática profunda nos vemos inmiscuidos en afirmaciones que no se pueden verificar o en preguntas sin respuesta. Ahora bien, este apartado tiene el propósito de hacer una breve recapitulación del análisis realizado en las páginas anteriores y sus consecuencias.

En el primer capítulo, consecuencia de la compatibilidad planteada por el MRCM entre procesos mentales y procesos algorítmicos y que persigue la explicación de la relación entre mente y cuerpo, nos enfocamos en la noción del 'algoritmo'. Así, explicamos –entendiendo la noción de 'algoritmo' desde la teoría informática– la brecha que separa la aplicación de un algoritmo en una máquina y la aplicación de un algoritmo por parte de un ser humano, en otras palabras, cómo una máquina *sigue reglas mecánicamente*, pero un ser humano aunque puede seguir reglas mecánicas, su actuar no se reduce a una relación causal porque siempre acaece en un contexto normativo. Esto último permite que el ser humano reconozca errores y pueda corregirse a partir de criterios de corrección, y, además pueda justificar sus acciones a partir de las reglas que comparte dentro de un juego de lenguaje. De esto que, nuestro primer capítulo, se erigió en torno a la revisión de la noción 'seguir una regla'– encontrada en Wittgenstein– respecto a el análisis del concepto de 'cálculo' desarrollado por Turing, a saber, como proceso mental recursivo.

Hemos estructurado la mencionada sección de la monografía de este modo, porque si se acepta que la conducta es causada por procesos mentales, la práctica del lenguaje también

lo será. Y, si los procesos mentales se entienden a partir del MRCM, entonces el uso del lenguaje natural es resultado de un algoritmo. Pese a que aún no se ha creado tal algoritmo, desde este modelo cognitivo se sostendría la posibilidad de que en un futuro pueda ser creado por el ser humano o creado por una máquina. Frente a esto, aunque aceptamos que de alguna manera la máquina *comprende* (otros ejemplos que hemos usado son: la máquina lee o la máquina calcula), resulta ser una inferencia ilícita decir que la máquina comprende como lo hace el ser humano; de modo que la analogía que propone MRCM y que sostiene las investigaciones en la IA se pone en duda.

En el segundo capítulo, proseguimos con el análisis del segundo presupuesto conceptual en el que se fundan las investigaciones en el campo de la IA, la presunción de que la 'comprensión' es un proceso mental/cerebral que sucede dentro del ser humano y causa la comprensión, por ejemplo, de una lectura. Además que, como resultado de la primera sección se reconoce que el ser humano puede tanto corregirse como justificarse, y, esto sucede en cuanto *comprende.* Asimismo, el presupuesto para usar un lenguaje es la comprensión del mismo. De manera que, en nuestro segundo capítulo mostramos la diferencia frente al análisis del lenguaje entre la actitud esencialista y la *visión perspicua* –piedra angular de la investigación gramatical–. Así, presentamos la crítica a la imagen esencialista del lenguaje porque nos lleva a malentendidos cuando se intenta encontrar la referencia del 'proceso de comprensión' y no se distingue el uso de los verbos psicológicos y el uso de verbos únicamente físicos, todo esto atravesado por ejemplo con el que se da inicio a la discusión en el capítulo II: la traducción de un lenguaje a otro. En contraste, con estos dos problemas hemos planteado dos nociones: 'comprensión superficial' y 'comprensión profunda', esta última entendida como capacidad y de ninguna manera como proceso sea mental o cerebral. Sumado a esto, se presenta la relación de la comprensión profunda con otros elementos de la descripción del lenguaje wittgensteniana, a saber, 'juego de lenguaje' y 'formas de vida' que presenta las particularidades del ser humano usuario del lenguaje natural. Esto es, la calculadora no calcula, la máquina de leer no lee, la máquina no comprende, aunque presentes los datos de salida correctos, porque las actividades de calcular, de leer y de comprender están siempre contextualizadas dentro de la forma de vida humana. En definitiva, el argumento a la postergación que versa sobre la aspiración –de algunos investigadores en el

campo de la IA– de que un sistema artificial inteligente sea usuario del lenguaje naturales resultado de un malentendido gramatical.

Referencias Bibliográficas

Bennett, M., Dennet, D., Hacker, P., & Searle, J. (2008). *La naturaleza de la conciencia. Cerebro, mente y lenguaje* (R. Filella, Trad.). Ediciones Paidós Ibérica.

Beuchot, M. (1986). El signo y el significado en San Agustín. *Diánoia*, *32*(32), 13-26.

Boden, M. A. (1994). Escape de la habitación china. En M. A. Boden (Ed.), & G. Feher de la Torre (Trad.), *Filosofía de la Inteligencia Artificial* (pp. 105-120). Fondo de Cultura Económica.

Boden, M. A. (2017). *Inteligencia artificial* (I. Pérez Parra, Trad.). Turner Publicaciones.

Copeland, J. (1996). *Inteligencia artificial. Una introducción filosófica* (J. C. Armero San José, Trad.). Alianza Editorial.

Copeland, J. (2011). Computation. En L. Floridi (Ed.), *The Blackwell Guide to the Philosophy of Computing and Information*. Blackwell Publishing.

Craik, K. (1943). *The Nature of Explanation*. Cambridge University Press.

Crick, F. (1995). *Astonishing Hypothesis: The scientific search for the soul*. Simon & Schuster.

De Groot, A. D. (1965). *Thought and Choice in Chess*. Mounton Publisher.

Descartes, R. (2011). *Descartes. Obras completas* (C. Flórez Miguel, Ed.). Editorial Gredos.

Durán, J. M. (2018). Ciencia de la computación y filosofía: Unidades de análisis del software. *Principia*, *22*(2), 203-227.

Ericsson, K. A., & Simon, H., A. (1980). Verbal Reports as Data. *Psychological Review*, *87*.

Fernández, G., & Sáez Vacas, F. (1987). *Fundamentos de informática*. Alianza Editorial.

Fodor, J. (1968). *Psychological Explanation*. Random House.

Fodor, J. (1975). *The Language of Thought*. Thomas Y. Crowell.

Frege, G. (2013). El pensamiento: Una investigación lógica. En L. M. Valdés Villanueva (Ed. y Trad.), *Ensayos de semántica y filosofía de la lógica* (2da ed., pp. 196-225). Editorial Tecnos.

Gallardo, A. (2011). Wittgenstein: Epistemología y lenguaje. *Praxis: revista del Departamento de Filosofía*, *66*, 87-98.

Garrido, Á. (2017). *Filosofía y computación*. Dykinson.

Gil de Pareja, J. L. (1988). El pensar en la filosofía de la mente de L. Wittgenstein. *Anuario filosófico, 21*(1), 123-132.

Giraldo, A. M. (2017). Formas de vida humana: ¿unidad o multiplicidad? En J. A. Suárez, M. C. Sánchez, & A. M. Giraldo (Eds.), *Wittgenstein-Husserl. Lebensform y Lebenswelt. Philosophy Workshop 2017* (pp. 57-72). Sello Editorial Javeriano.

Glock, H.-J. (1996). *A Wittgenstein Dictionary*. Blackwell.

Gödel, K. (1981). *Obras completas*. Alianza Editorial.

Gurevich, Y. (2012). *What Is an Algorithm?* (p. 42). https://doi.org/10.1007/978-3-642-27660-6_3

Hacker, P. M. S. (2013). Lenguaje, juegos de lenguaje y formas de vida. En J. Padilla Gálvez & M. Gaffal (Eds.), *Formas de vida y juegos de lenguaje* (pp. 29-60). Plaza y Valdés Editores.

Haugeland, J. (2001). *Inteligencia artificial* (I. Tulli De Firmati, Trad.; 3ra ed.). Siglo XXI Editores.

Herrera Ibáñez, A. (1989). La teoría del significado como uso en Wittgenstein. *Revista de filosofía DIÁNOIA, 35*(35), 203-210. https://doi.org/10.22201/iifs.18704913e.1989.35.644

Hutchinson, P. (2010). Thinking and understanding. En K. D. Jolley (Ed.), *Wittgenstein: Key Concepts* (pp. 92-108). Acumen Publishing.

Kitcher, P. (2001). *El avance de la ciencia: Ciencia sin leyenda, objetividad sin ilusiones* (H. Islas & L. Manriquez, Trads.). Universidad Nacional Autónoma de México.

Lagerspetz, O. (2020). Wittgenstein's Forms of Life: A Tool of Perspicuous Representation. *Nordic Wittgenstein Review*. https://doi.org/10.15845/nwr.v0i0.3560

Lin, F. Y. (2019). Wittgenstein on Thinking as a Process or an Activity. *Wittgenstein-Studien*, *10*(1), 73-104. https://doi.org/10.1515/witt-2019-0004

Llinás, R. (2001). *I of the Vortex. From Neurins to Self*. The MIT Press.

Malcolm, N. (1989). Wittgenstein on Language and Rules. *Philosophy*, *64*(247), 5-28.

Masterman, M. (2005). *Language, Cohesion and Form* (Y. Wilks, Ed.). Cambridge University Press.

McCarthy, J. (1999). Lógica matemática en inteligencia artificial. En S. R. Graubard (Ed.), *El nuevo debate sobre la Inteligencia Artificial. Sistemas simbólicos y redes neuronales.* (2da ed., pp. 335-349). Editorial Gedisa.

McGinn, M. (1997). *Wittgenstein and the Philosophical Investigations* (T. Crane & J. Wolff, Eds.). Routledge.

Moreno May, K. (2011). Wittgenstein y la naturalización de la mente. *Tesis Psicológica*, *6*, 183-200.

Moschovakis, Y. (2001). What is an algorithm? En B. Engquist & W. Schmid (Eds.), *Mathematics Unlimited: 2001 and Beyond* (pp. 919-936). Springer.

Newell, A., & Simon, H., A. (1994). La ciencia de la computación como investigación empírica: Símbolos y búsqueda. En M. A. Boden (Ed.), & G. Feher de la Torre

(Trad.), *Filosofía de la Inteligencia Artificial* (pp. 122-152). Fondo de Cultura Económica.

Padilla Gálvez, J. (2014). *Hacia la representación perspicua. Wittgenstein II*. Tirant Humanidades.

Padilla Gálvez, J., & Gaffal, M. (Eds.). (2013). *Formas de vida y juegos de lenguaje*. Plaza y Valdés Editores.

Penrose, R. (1994). *Shadows of the Mind*. Oxford University Press.

Pérez Hernández, M. C. (2002). *Explotación de los córpora textuales informatizados para la creación de bases de datos terminológicas basadas en el conocimiento* (Vol. 2). Universidad de Málaga.

Pérez Otero, M. (2000). El argumento antiintelectualista de Wittgenstein sobre la comprensión del lenguaje. *Theoria: Revista de Teoria, Historia y Fundamentos de la Ciencia*, *15*, 155-169.

Putnam, H. (1975). Minds and Machines. En *Mind, Language and Reality. Philosophical Papers* (Vol. 2, pp. 362-385). Cambridge University Press.

Putnam, H. (1994). *¿Cómo renovar la filosofía?* (C. Laguna, Trad.). Ediciones Cátedra.

Rojo, R. (1999). Wittgenstein: Gramática y metafísica. *Thémata Revista de Filosofía*, *21*, 263-276.

San Agustín. (2003). *El maestro o sobre el lenguaje y otros textos* (A. Domínguez, Trad.). Editorial Trotta.

San Agustín. (2010). *Confesiones* (A. Encuentra Ortega, Trad.). Editorial Gredos.

Schank, R. C. (1986). *Explaniation Patterns: Understanding Mechanically and Creatively*. Psychology Press.

Searle, J. (1984). *Minds, brains and sciencie*. Harvard University Press.

Searle, J. R. (2001). *Mentes, cerebro y ciencia* (L. M. Valdés Villanueva, Trad.; 4ta ed.). Ediciones Cátedra.

Shanker, S. (1998). *Wittgenstein's Remarks On the Foundations Of AI*. Routledge.

Simon, H. A., & Newell, A. (1962). Computer Simulation of Human Thinking and Problem Solving. *Monographs of the Society for Research in Child Development*, *27*(2), 137-150. https://doi.org/10.2307/1165535

Smith, B. C. (1985). The limits of correctness. *ACM SIGCAS Computers and Society*, *14*,*15*(1,2,3,4), 18-26. https://doi.org/10.1145/379486.379512

Thagard, P. (2008). *La mente. Introducción a las ciencias cognitivas* (S. Jawerbaum & J. Barba, Trads.). Katz Editores.

Tomasini Bassols, A. (2006). Comprensión: Filosofía tradicional versus Pensamiento Wittgensteiniano. *Revista de filosofía*, *24*(53), 89-105.

Trajtenbrot, B. A. (1977). *Los algortimos y la resolución automática de problemas* (B. Del Río Salceda, Trad.). Editorial MIR.

Turing, A. M. (1950). Computing Machinery and Intelligence. *Mind*, *59*(236), 433-460.

Winograd, T. (1972). SHRDLU: A System for Dialog. *1972? Lecture 2: pp. 20-48 : ill. and diagrams. includes bibliography*. http://papers.cumincad.org/cgi-bin/works/paper/2b77

Winograd, T. (1980). *What Does It Mean to Understand Language? 4*(Cognitive Science), 209-241.

Wittgenstein, L. (1978). *Observaciones sobre los fundamentos de la matemáticas* (G. H. von Writht, R. Rhees, & G. E. M. Anscombe, Eds.; I. Reguera, Trad.). Alianza Editorial.

Wittgenstein, L. (1992). *Gramática filosófica* (J. Padilla Gálvez, Trad.). Editorial Trotta.

Wittgenstein, L. (2014). *Escrito a máquina. The Big Typescript. TS 213* (J. Padilla Gálvez, Trad.). Editorial Trotta.

Wittgenstein, L. (2017a). *Investigaciones Filosóficas* (U. Moulines, Trad.; 2da ed.). Universidad Nacional Autónoma de México, Instituto de Investigaciones Filosóficas.

Wittgenstein, L. (2017b). *Los cuadernos azul y marrón* (F. Gracia Guillen, Trad.; 5ta ed.). Editorial Tecnos.