

ALGUNAS CAPACIDADES COGNITIVAS GENERALES QUE SUBYACEN AL LENGUAJE*

Javier Valenzuela Manzanares
(Universidad de Murcia)

1. Introducción

La pregunta de cuáles son las características que nos hacen específicamente humanos, es decir, cuáles son los rasgos que nos distinguen de otras especies inteligentes, especialmente de nuestros parientes más próximos evolutivamente, los primates, sigue concitando un enorme interés tanto en legos como en especialistas. Una de las características distintivas suelen aparecer en primer lugar es nuestra capacidad lingüística. El lenguaje humano es claramente algo especial y sin parangón en la naturaleza. Nuestra capacidad de comunicación lingüística posee un nivel de eficacia y flexibilidad no encontrado en ningún otro lenguaje animal; ningún otro medio de comunicación es tan potente, versátil y complejo como el nuestro. Al mismo tiempo, el lenguaje tiene una profunda influencia en nuestro sistema conceptual y en nuestra misma esencia, de manera que una explicación del ser humano que no tenga en cuenta su facultad lingüística siempre será incompleta. En este sentido, es perfectamente comprensible que nos preguntemos cuáles son los mecanismos biológicos que subyacen al funcionamiento del lenguaje y que permiten a nuestra especie distinguirse del resto. Este área de investigación ha resultado ser extraordinariamente compleja y ha generado una enorme controversia en círculos académicos, puesto que las propuestas son de signo muy diverso e incluso opuesto. Por un lado, existe una corriente de investigación que opina que, al ser un rasgo específicamente humano, el lenguaje debe tener una base genética muy definida dedicada de manera específica a crear las capacidades que explican el lenguaje (e.g., Bickerton, 1990; Chomsky, 1975; Jackendoff, 2002; Jenkins, 2000; Pinker, 2009). Estas capacidades, especialmente diseñadas para adquirir el

* Este trabajo es resultado del proyecto de investigación 19482/PI/14 financiado por la Fundación Séneca-Agencia de Ciencia y Tecnología de la Región de Murcia en el marco del PCTIRM 2011-2014, así como del proyecto de investigación FFI2013-45553-C3-1-P financiado por el Ministerio de Economía y Competitividad.

lenguaje y codificadas de manera innata en nuestro código genético, permitirían que cualquier humano, nazca en la sociedad que nazca, sea capaz de adquirir y utilizar sin problemas la variante de comunicación lingüística (esto es, el idioma) que se utilice en esa sociedad. Este tipo de suposición de partida ha generado una "agenda de investigación" que busca explicar entonces el fenómeno lingüístico desde estas premisas. Por ejemplo, este tipo de acercamientos intenta, por un lado, buscar cuáles son las características de funcionamiento comunes a todos los idiomas del mundo (puesto que ahí estarán los mecanismos "centrales" cuyo funcionamiento debe ser explicado) y por otro, averiguar qué aspectos de la capacidad lingüística deben ser explicados de manera exclusivamente lingüística. Es decir, un objetivo es localizar aspectos cuyo aprendizaje no pueda ser realizado por animales de otra especie y soportados por capacidades dedicadas únicamente y exclusivamente al lenguaje. Por estas razones, se caracteriza a este área de investigación como partidarios del "innatismo lingüístico" y de la "especificidad del dominio" (mecanismos cuyo dominio de aplicación es "específico" del dominio del lenguaje).

Otro grupo de autores optan por un acercamiento distinto. Estos autores no buscan la explicación del fenómeno lingüístico en mecanismos específicos del lenguaje, sino en capacidades distintas, de tipo general, y avanzan que la especificidad lingüística se puede deber a la combinación de estas capacidades no lingüísticas de una manera especial en el ser humano. En este sentido, la idea es que puede que el lenguaje sea una "máquina nueva hecha con partes viejas". Desde este punto de vista, el lenguaje no surge de capacidades específicamente diseñadas para solucionar el problema lingüístico, sino que emerge de una nueva combinación de capacidades ya pre-existentes, aunque potenciadas en la especie humana, de manera que la diferencia entre lenguajes animales y lenguajes humanos es más cuantitativa que cualitativa. De la mejora puntual de determinadas capacidades cognitivas, muchas de ellas compartidas con otros animales, y destinadas a otros fines, y por medio de una nueva combinación, emerge el lenguaje como nueva capacidad, esta sí, ya específica de la especie. La palabra clave en esta explicación es la noción de "emergencia", en la que la combinación de elementos logra propiedades que no están presentes en ninguno de los elementos individuales.

Tal y como se plantea en este trabajo, la pregunta sobre si existen unos mecanismos específicamente lingüísticos, o el lenguaje surge de la combinación de capacidades de aprendizaje generales, es una cuestión puramente empírica. Ambas opciones corresponden a dos agendas de investigación paralelas, que siguen rutas

distintas y exploran distintas áreas, utilizando métodos distintos. En el presente trabajo, vamos a examinar algunas de esas capacidades cognitivas de tipo general que se han postulado como posibles candidatos para explicar el funcionamiento del lenguaje humano.

Antes de iniciar la revisión es necesario dejar claro un asunto, que como todos los que rodean a esta cuestión, puede ser controvertido. Existe una tendencia muy generalizada a asociar "especificidad del dominio" con "innatismo", por un lado, y las "capacidades de dominio general" con "no-innato" o "aprendido". A pesar de que esta asociación es muy frecuente, estos conceptos no tienen por qué estar asociados de manera necesaria, y de hecho, se pueden considerar como perfectamente ortogonales. Una capacidad de dominio general puede tener una base innata (y de hecho, probablemente lo tendrá, puesto que *todo* tiene una base genética e innata) y una capacidad de dominio específica puede ser aprendida y exhibir la modularidad cerebral que se atribuye a las capacidades de tipo específico. Este último sería el caso de las capacidades de lectura y escritura, que son perfectamente específicas, pero también aprendidas. Este trabajo, por lo tanto, no se dedica a intentar contestar la pregunta "¿Es el lenguaje innato?" (que, por otro lado, formulada de esta manera probablemente sea excesivamente vaga y poco productiva). Es evidente que existe una base genética e innata subyacente a absolutamente todas nuestras capacidades y comportamientos: nuestras capacidades de colaboración, la construcción de estructuras sociales complejas, nuestro tipo de alimentación, nuestro gusto por la estética, las creencias religiosas, la existencia de juegos incluso en la edad adulta, y un larguísimo etcétera. Todos estos comportamientos deben tener una base genética, puesto que de otra manera, no serían desarrollados por nuestra especie. Lo que se explora no es esto, sino si existen mecanismos *exclusivamente* dedicados al procesamiento lingüístico, indicadores de una conducta específica que sólo puede ser explicada por su codificación genética directa en nuestro ADN.

2. Algunas capacidades generales que subyacen al lenguaje

2.1. Aprendizaje estadístico

Uno de las capacidades más reseñables de la mente humana es nuestra habilidad para encontrar correlaciones entre distintos aspectos de nuestra experiencia. El cerebro humano es una potente máquina de reconocimiento de patrones: detectamos de manera

automática regularidades en nuestro entorno que nos permiten estructurar el espacio de experiencias, predecir el comportamiento del mundo y facilitar de esta manera nuestra supervivencia. De manera más específica, nuestras capacidades de aprendizaje estadístico nos permiten encontrar el orden subyacente a una determinada estructura secuencial, encontrar correlaciones entre distintas dimensiones de los estímulos que experimentamos o detectar las asociaciones existentes entre distintos tipos de información. Somos además especialmente sensibles a parámetros como la frecuencia (o si se quiere, la probabilidad) con la que algo ocurre o con la que dos estímulos co-ocurren. Existe en la actualidad una gran cantidad de evidencia que demuestra este tipo de habilidades de dominio general se utilizan para solucionar algunos de los problemas subyacentes al procesamiento lingüístico. Revisamos algunos de ellos.

2.1.1. Un primer ejemplo: la segmentación de palabras

¿Por dónde se empieza cuando alguien quiere aprender una lengua? La respuesta más inmediata diría que hay que comenzar aprendiendo el significado de las palabras, es decir, memorizando cuál es la asociación entre un determinado sonido (e.g., [mesa]) y el significado convencionalmente asignado a ese sonido ("un tipo de mueble"). Sin embargo, establecer esta asociación es un problema más complicado de lo que podríamos pensar. Cuando no conocemos una lengua, es prácticamente imposible saber qué partes del habla se corresponden con palabras. Pensemos en qué ocurre cuando escuchamos a alguien hablando un idioma extraño: con toda seguridad no seremos capaces de precisar con exactitud cuántas palabras está diciendo. Y esta es precisamente la situación a la que se enfrentan los niños cuando tienen que adquirir su lengua materna: el habla que escuchan a su alrededor corresponde a un continuo sonoro del que es extremadamente complejo extraer las unidades que lo componen. Los niños no escuchan *no toques eso que quema*, sino algo como "*notoquesesoquequema*": las palabras se escuchan sin separación alguna entre ellas. Además de otras complicaciones adicionales: no solo no se pronuncian todas las palabras sin interrupción alguna, sino que con frecuencia, los sonidos finales de una palabra se solapan con los de la siguiente, como por ejemplo en "*¿comostashoy?*" o "*¿taninvitaolcumple?*" y se influyen entre sí (por ejemplo, cuando una palabra termina en "n" y la siguiente palabra empieza por una consonante bilabial, esa "n" se transforma en "m", como en *un beso (umbeso)* o *San Pedro ("Sampetro")*, un fenómeno conocido con el nombre de "asimilación"). Así que

nos encontramos una situación en la que los niños deben ser capaces de aislar las palabras y reconocerlas como unidades para poder empezar a aprender su significado, pero si no saben las palabras de antemano, no son capaces de aislarlas del habla continua que escuchan a su alrededor: es un círculo vicioso del que resulta difícil escapar. La respuesta a este dilema está en nuestras habilidades estadísticas.

Los niños son sensibles a la frecuencia con la que dos sílabas aparecen una después de la otra en el lenguaje (lo que se conoce con los términos técnicos de "probabilidades de transición" o "probabilidades secuenciales"). Se puede comprobar muy fácilmente con un ejemplo. Pongamos la sílaba del español "pla". Si escuchamos este sonido, ¿qué otro sonido puede venir después? La respuesta en este caso es: muchísimos. Sin pretender ser ni exacto ni exhaustivo, una búsqueda en un diccionario cualquiera nos da aproximadamente unas 1700 palabras que empiezan por la sílaba "pla". Sin embargo, si tras escuchar esta sílaba, escuchamos "ta", el número de opciones para continuar se reduce drásticamente. Es evidente que no hay tantas palabras que empiecen por "pla-ta" (de miles pasamos a decenas). ¿Qué ocurre si después de "pla-ta", escuchamos la sílaba "nos"? ¿Cuál será el sonido que vendrá después? Esta vez, las posibilidades de continuación vuelven a aumentar drásticamente: después de "plá-ta-nos" puede venir cualquier sílaba (por poner ejemplos con las primeras cinco letras del abecedario, "plátanos amarillos/buenos/canarios/de Canarias/españoles"); en la realidad, cualquier sonido puede seguir a esta combinación de sílabas). Y es precisamente este súbito aumento en las posibilidades de combinación lo que indica que hemos alcanzado el final de una unidad, de una palabra (véase Figura 1).

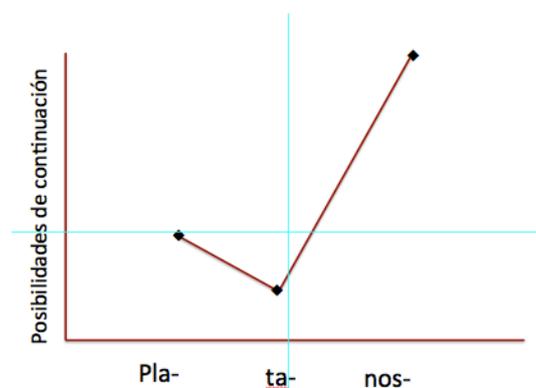


Figura 1. Posibilidades de continuación de sílabas

La psicóloga Jenny Saffran demostró que los bebés son capaces de reconocer la estructura estadística de las secuencias de combinación de sílabas. En su experimento

más conocido (Saffran, Aslin y Newport, 1996), niños de ocho meses escuchaban una secuencia grabada de sílabas sin sentido, realizada sin ningún tipo de entonación (e.g., go-la-bo-pa-do-ti-bi-da-ku-ti-pi-ro-go-la-bo-da-ku-ti, etc.). Esta secuencia estaba construida de tal manera que, por ejemplo, a la sílaba "go" le seguía la sílaba "la" el 100% de las veces, mientras que el sonido "bo" era seguido por la sílaba "da" sólo un 33% de las veces (véase Figura 2).

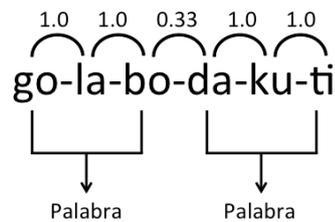


Figura 2. Probabilidades de transición entre sílabas (adaptado del experimento de Saffran et al, 1996)

Durante dos minutos, los niños escuchaban estas secuencias; en una segunda fase, escuchaban combinaciones de estos sonidos para comprobar si eran capaces de reconocer cuáles de esas combinaciones formaban "palabras". Los niños eran capaces de reconocer cuáles eran las combinaciones de sílabas sin sentido que habían ocurrido como unidades invariables en los estímulos anteriores (esto es, como palabras), proporcionando una prueba absolutamente sólida de la sensibilidad de los niños a la información distribucional y estadística presente en las secuencias de sonidos del habla y una elegante solución al problema de la segmentación de palabras.

2.1.2. Sintaxis

La noción de categoría gramatical es enormemente importante en la descripción lingüística, puesto que prácticamente todas las teorías la utilizan como "unidad" para describir la sintaxis de las lenguas. Claramente, para poder describir cómo ordenamos las palabras es necesario invocar este tipo de categorías, que permiten que en vez de tener que expresar reglas sintácticas al nivel de las palabras individuales (e.g., *el, la, perro, mesa, ladra, a, en, su, el, salón, dueño...*), podamos establecer el orden a un nivel superior, haciendo el problema más tratable: por ejemplo, expresamos una regla, que se aplica a una gran cantidad de casos, por medio de la cual los Determinantes preceden a los Nombres. A su vez, la combinación Determinante + Nombre puede formar una

unidad de orden superior (un Sintagma Nominal), que precede a la categoría Verbo cuando es sujeto y la sigue cuando es objeto directo, etc. Es así como se puede lograr una descripción sintáctica de las lenguas relativamente viable y sencilla. Sin embargo, todo lo concerniente al origen y la naturaleza de las categorías gramaticales es también un asunto muy controvertido, puesto que por un lado, en ellas confluye información de muy diverso tipo (fonológico, morfológico, sintáctico y semántico); las categorías gramaticales pueden ser descritas con criterios fonológicos, morfológicos, sintácticos o semánticos. Sin embargo, para algunas teorías, ninguna de las explicaciones basadas en estas características parciales es la central, y han señalado que todas ellas tienden a tener huecos (e.g., Gleitman, 1990). Para las teorías partidarias de explicaciones de dominio específico, la manera de solucionar el problema de la naturaleza de las categorías gramaticales es asumir que estamos preparados de manera innata para reconocerlas, incluso con una exposición a ejemplares muy baja (Chomsky, 1965). Algunas variantes asumen una definición semántica inicial (que equipara de manera aproximada a los nombres a objetos y a los verbos a acciones y estados, por ejemplo), que se generaliza a partir de un proceso conocido como "bootstrapping semántico" (Pinker 1984). En los últimos tiempos, sin embargo, la contribución de cada una de las diferentes fuentes de posible información para el establecimiento de categorías gramaticales ha sido revisada y ampliada por medio de cuidadosos análisis estadísticos, que han relevado fuentes de información en el input que reciben los niños hasta ahora pasadas por alto. Pasamos a una breve revisión de algunas de estas fuentes.

En lo referente a la información fonológica, existe evidencia de que las características fonológicas de las palabras son una pista muy importante sobre la categoría gramatical a la que deben ser adscritas. Monaghan et al (2005) ofrecen un resumen de la literatura de los parámetros que distinguen nombres de verbos:

1. Número de sílabas (N>V)
2. Longitud de la palabra en fonemas (N>V)
3. Sonoridad de la consonante final (N<V)
4. Consonantes nasales (N>V)
5. Posición frontal de la vocal acentuada (N<V)
6. Posición frontal son las vocales (N<V)
7. Altura de las vocales (N<V)
8. Posición del acento (N-trocaico, V-yámbico)

9. Los nombres bisilábicos tienden a tener un sólo acento; los verbos tienden a tener además un segundo acento secundario

Existe evidencia experimental de que hablantes ingleses reconocen una palabra como nombre con mayor rapidez cuando contiene vocales de posición posterior (*school*) que frontal (*street*); lo contrario fue hallado para verbos. Ya existen algunos modelos computacionales que son capaces de asignar una palabra a una categoría gramatical utilizando únicamente esta información con un alto grado de eficacia.

Otra de las fuentes de información para reconocer categorías gramaticales es la distribucional (o sintáctica, si se quiere). En su experiencia con el discurso, por ejemplo, los hablantes tienden a encontrar el orden *el perro* y no **perro el*. Además, la palabra *los* tiende a encontrarse delante de nombres plurales masculinos y las preposiciones delante de sintagmas nominales. Un acercamiento puramente distribucional ha sido tradicionalmente rechazado por los partidarios de explicaciones innatas, debido a la idiosincrasia, variabilidad de frecuencia e irregularidad existente en las combinaciones entre palabras. No obstante, en los últimos años y utilizando técnicas de análisis estadístico más sofisticadas, cada vez más autores afirman que el aprendizaje de las categorías gramaticales puede derivarse por completo de patrones de distribución estadística existentes en el input que reciben los hablantes. Por ejemplo, Reeder, Newport y Aslin (2013) realizaron un experimento utilizando lenguajes artificiales en el que mostraban cómo los hablantes son altamente sensibles a los contextos en los que aparece una palabra dada; estos autores fueron capaces de aislar una serie de variables distribucionales utilizadas por los hablantes para asignar a una palabra a una categoría gramatical dada, tales como el *número* de contextos en los que aparece una palabra o la *densidad* o proporción de estos contextos en el input. Estas variables se combinan con otros datos, como como la *frecuencia* relativa de estas pistas: cuando una determinada configuración no se encuentra debido a una baja frecuencia de las palabras que la conforman, el aprendiz puede pensar que es una omisión; cuando una configuración no ocurre a pesar de la alta frecuencia de sus componentes y de lo "esperable" estadísticamente que sería su combinación, el hablante asume que esa configuración no está permitida y debe ser descartada; es lo que se conoce con el nombre de "pre-emption" o "bloqueo" (Goldberg, 2011), que de nuevo ha sido demostrada tanto en estudios de corpus como en estudios de adquisición de primeras lenguas. Hay que decir

que estas estrategias se combinan un pequeño grupo de sesgos de aprendizaje también de tipo general.

2.1.2. Semántica vectorial: extrayendo el significado de las palabras de combinaciones sintácticas

Uno de los acercamientos a la semántica de mayor impacto en los últimos tiempos tiene que ver con nuestra capacidad para extraer información semántica sobre la base de patrones estadísticos de la pura combinación de palabras. El razonamiento básico es que, si dos palabras están cercanas en significado, tenderán a aparecer juntos a grupos de palabras parecidos. Por ejemplo, la palabra "flor" aparece en combinaciones como *oler una flor*, *cortar una flor*, *cultivar una flor*, *comprar una flor*, etc. Si encontramos una nueva palabra, digamos *clavel*, que aparece en los mismos contextos (*oler/cortar/cultivar/comprar un clavel*), podemos suponer que el significado de ambas palabras estará cercano. Asimismo, es razonable pensar que grupos de palabras como *doctor*, *médico*, *enfermera*, *paciente*, *hospital* tenderán a aparecer en contextos parecidos, mientras que, por su parte, palabras como *menú*, *restaurante*, *comida*, *postre* o *camarero* también tenderán a aparecer en contextos similares y distintos de los del primer grupo; lo mismo podríamos decir de *lápiz*, *bolígrafo*, *papel*, *goma* y *sacapuntas* y un larguísimo etcétera. De manera muy aproximada, este es el acercamiento que métodos como el Análisis de Semántica Latente (en inglés, *Latent Semantic Analysis* o LSA), implementan en ordenadores. Estos programas analizan las co-ocurrencias de millones de palabras e intentan extraer significado de esas co-ocurrencias. LSA toma un corpus lingüístico de varios millones de palabras y lo organiza en una gigantesca tabla: las filas son las distintas palabras del corpus y las columnas son los "documentos" (divisiones del corpus en subtextos) en las que una palabra puede aparecer; cada "documento" en los que se divide un corpus puede tener, por ejemplo, entre cinco o diez líneas. En esta gigantesca matriz, que puede tener 60.000 filas/palabras y 140.000 columnas/documentos, cada celda contiene el número de veces en las que una palabra aparece en un documento dado. Cada fila entonces es una concatenación de 140.000 números (lo que llama un "vector"); este vector se simplifica extrayendo la información redundante hasta dejarlo en unas 200-300 dimensiones. Esta concatenación de números se supone que codifica el significado de esa palabra. La comparación entre los vectores de dos palabras determina el grado de cercanía semántica entre ambas: dos palabras

cercanas en significado tenderán a aparecer en documentos similares y por lo tanto, su vector será más similar que dos palabras no similares semánticamente, cuyos contextos de aparición serán muy distintos. La cercanía semántica entre dos palabras se puede calcular por medio de una sencilla operación matemática de comparación dos vectores, lo que da una medida exacta, que oscila entre 1 (para dos palabras exactamente iguales, que aparecen en los mismos contextos) y 0 (para dos palabras cuyos contextos no coinciden en absoluto y por lo tanto se asumen máximamente distantes semánticamente).

Esto puede parecer un acercamiento completamente "ciego" al significado, puesto que el ordenador únicamente reconoce la "forma" de las palabras, sin saber realmente qué significan. Sin embargo, el razonamiento sobre el que se basa resuena con muchas teorías lingüísticas, desde el estructuralismo de Saussure (*La lengua es un sistema de términos interdependientes en el que el valor de cada término resulta solamente de la presencia simultánea de los otros*), hasta el de Firth (*you shall keep a word by the company it keeps*) pasando por Wittgenstein y su acercamiento al significado como uso.

Los acercamientos vectoriales al significado han sido capaces de emular comportamientos lingüísticos muy similares a los humanos. Por ejemplo, se ha comparado el resultado obtenido por estos programas en el examen de inglés americano TOEFL con los resultados obtenidos por estudiantes de segundas lenguas sometidos al mismo examen. En este examen, se proporciona a los estudiantes una palabra y tienen que decir cuál de las cuatro opciones suministradas es la más cercana semánticamente. Los resultados de los humanos y los de la máquina resultaron ser muy similares. Los acercamientos de LSA también son útiles para predecir la cantidad de "facilitación" (priming) entre dos palabras en un test psicolingüístico, y han sido capaces de imitar juicios de categorización, de similaridad, sinónimos, antónimos efectuados por los humanos. Louwersee (2012) utilizó como corpus de entrenamiento la novela de Tolkien *El Señor de los Anillos* y encontró que el programa era capaz de predecir la cercanía geográfica de los pueblos que aparecían mencionados simplemente utilizando un sistema basado en la cercanía en la co-ocurrencia. El paralelismo entre los resultados por estos programas y los obtenidos por los humanos hace pensar que están utilizando un método que es el utilizado por los hablantes humanos. Como prueba se aduce también los resultados obtenidos con niños con ceguera congénita, que nunca han visto los colores: estos niños son capaces de decir que el color "rojo" es más cercano al

"naranja" que al "azul", o que el color "dorado" es más cercano al "amarillo"; sin una experiencia directa, la única manera de que estos niños puedan extraer esta información es por medio de una sensibilidad a los patrones de combinaciones sintácticas de las palabras.

2.1.4. Conclusión parcial

Los ejemplos discutidos hasta ahora muestran claramente como una capacidad de dominio general, nuestra capacidad para extraer patrones de recurrencia probabilística en los estímulos de nuestro entorno, juega un papel relevante en multitud de asuntos directamente relacionados con la estructura de la lengua. No hemos agotado los ejemplos disponibles; por poner un único ejemplo adicional, dentro de la fonética, se mostró que los bebés utilizando información de distribución probabilística para seleccionar cuál es el inventario de fonemas de una lengua. Sonidos como la D fricativa ([ð]) y la oclusiva ([d]) son simplemente variantes alofónicas del mismo fonema en español (como en la palabra *dedo*, en la que la primera "d" es oclusiva, mientras que la segunda es interdental y fricativa), mientras que pertenecen a fonemas distintos en inglés (*then* vs *den*). Los niños son capaces de encontrar también correlaciones entre palabras que escuchan y la presencia de determinados objetos en el entorno, de manera que son capaces de encontrar la proyección adecuada simplemente por observación (Figura 3):

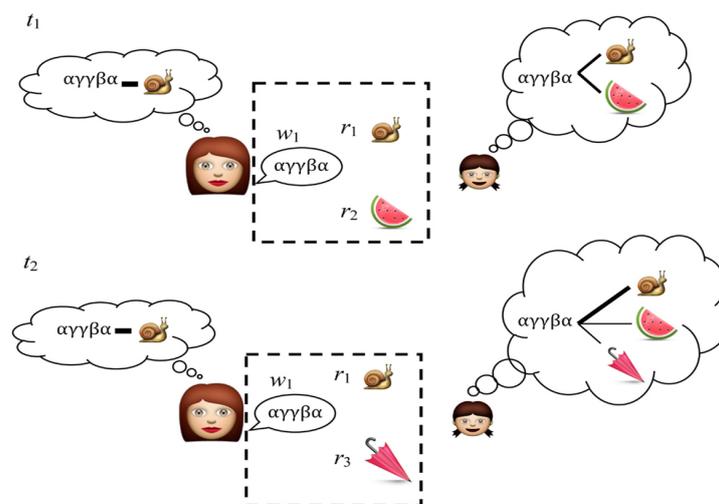


Figura 3. Correlación entre palabras y situaciones

La estrategia de buscar explicaciones al modo en que adquirimos el lenguaje en mecanismos de dominio general resulta de esta manera altamente informativa en el caso de la información de tipo estadístico.

2.2. Capacidades de categorización

La categorización es una capacidad de dominio general, completamente básica para nuestra supervivencia como especie: las categorías nos permiten simplificar la complejidad del mundo y efectuar generalizaciones rápidas a partir de la ascripción de un ejemplar a una categoría dada. Al asignar un animal concreto y específico a la categoría "gato", inmediatamente accedemos una enorme cantidad de información que tenemos almacenada en su representación categorial, como por ejemplo, cómo se mueve, qué tipo de sonido produce, cuál es su tipo de dieta más usual o cuál es su relación con el ser humano. Los mecanismos de categorización se basan en muchas fuentes: nos podemos basar en pistas perceptuales (agrupando cosas verdes), funcionales (cosas que se pueden comer) o semánticas (cosas que flotan). Para construir categorías combinamos una serie de habilidades; en algunos casos son las capacidades de aprendizaje estadístico que vimos en el apartado anterior; somos capaces de observar la co-ocurrencia de determinados rasgos en un elemento dado: existe una correlación entre raíces, tronco, hojas y ramas que se observa en la categoría "árbol", o plumas, pico y patas en los pájaros. O de la secuencia de subeventos en un evento complejo: el evento global "lavarse los dientes" está compuesto de subeventos como coger el cepillo, ponerle pasta, mojarlo, lavarnos los dientes, enjuagarnos la boca, secarnos con la toalla, enjuagar el cepillo y devolverlo a su sitio. Lo mismo se puede observar en actividades como hacer un café, y un larguísimo etcétera. Igualmente, para asignar un ejemplar a una categoría dada, debemos establecer una comparación entre las características de este ejemplar y las de otros ejemplares de la categoría (o el ejemplar más prototípico de la misma), es decir, debemos ser capaces de mostrar sensibilidad a los parecidos y diferencias entre dos elementos. También debemos ser capaces de realizar un cierto nivel de abstracción, descartando los detalles concretos de un ejemplar que no sean necesarios para su ascripción a la categoría al no recurrir en distintos ejemplares.

De nuevo, se ha mostrado que estos mecanismos, presentes en la capacidad de dominio general de la categorización, juegan un papel muy relevante en la construcción de nuestro conocimiento lingüístico. Por ejemplo, existe una estrategia utilizada para

construir categorías conocida con el nombre de *chunking* o "agrupamiento". Este mecanismo facilita el almacenaje en la memoria de una serie de items; nuestra memoria de corto plazo tiene una capacidad limitada a unos siete elementos (de ahí el conocido artículo del psicólogo George Miller, "el número mágico 7, más o menos 2"). Por ejemplo, podemos recordar una serie de números de manera individual, pero si logramos agruparlos de manera coherente, seremos capaces de recordar una mayor cantidad. Así, la secuencia 1914200014921945 es difícil de recordar si tomamos cada número como una unidad aislada, ya que está compuesta por 20 números. Sin embargo, si dividimos la secuencia en trozos (*chunks*) coherentes, la tarea se ve facilitada: 1914-2000-1492-1945, puesto que reducimos el número de elementos a almacenar a cuatro. En el aprendizaje de secuencias motoras, el agrupamiento juega un papel crucial. Cuando nos enseñan a jugar al tenis, el instructor descompone el movimiento del golpe de revés en una serie aislada de movimientos parciales (cambio de empuñadura, giro de los hombros en la dirección de la pelota, cambio de posición de los pies hasta alinearse de manera perpendicular a la pelota, transferencia del peso del cuerpo, movimiento hacia atrás de la raqueta, posición del codo y realización del golpe, atendiendo a la realización completa del movimiento). Para la ejecución correcta del golpe, esta secuencia de acciones debe automatizarse e integrarse en un todo coherente, de manera que ya no estemos pendientes de cada uno de estos submovimientos y seamos capaces de realizarlos con fluidez y soltura. Lo mismo podría ser aplicable al aprendizaje de un instrumento musical, a la conducción de un coche, etc. Diversos autores han mostrado cómo la estrategia del agrupamiento puede ser altamente exitosa para explicar la adquisición del lenguaje. Los hablantes adultos se basan en un grado mucho mayor del asumido hasta hace mucho en fórmulas pre-construidas, en "grupos" (*chunks*) de lenguaje que se recuperan de la memoria y se montan como piezas de un puzzle (Wray, 2002). De esta manera, se puede ver el conocimiento gramatical como "comportamiento automatizado"; lo que los usuarios de un lenguaje utilizan son patrones automatizados por medio de la repetición, que se esquematizan y generalizan en distintos grados. Las construcciones gramaticales emergen por medio del "agrupamiento" de palabras que coocurren y que forman unidades que se esquematizan y sirven como patrones más generales (haciendo uso de esa capacidad de abstracción o esquematización que mencionábamos era necesaria para la construcción de categorías). Acercamientos lingüísticos como los "basados en el uso" (e.g., Bybee, 2006; Tomasello, 2003) utilizan este tipo de mecanismos para explicar multitud de fenómenos

lingüísticos, desde la gradación de los juicios de gramaticalidad hasta la evolución diacrónica de las lenguas, pasando por los patrones de adquisición de construcciones gramaticales seguidos por los niños al adquirir sus lenguas maternas.

2.3. Capacidades de colaboración y teoría de la mente

Los acercamientos que intentan explicar cómo funciona el lenguaje centrándose en las reglas de combinación de palabras, esto es, la sintaxis, omiten una parte muy significativa de cómo funciona el lenguaje de manera global. No es que el estudio de la sintaxis no sea relevante, ni mucho menos, pero para gran cantidad de autores sí que es únicamente una parte de lo que debiera ser una explicación completa del fenómeno lingüístico. La idea de que el lenguaje funciona debido a que tenemos una serie de palabras (un lexicón), cada una de ellas con un significado asignado, y que la combinación de estas palabras por medio de reglas sintácticas permite construir frases y oraciones que codifican de manera directa el significado que queremos comunicar, se encuentra con problemas muy pronto. Tomemos por ejemplo, el siguiente diálogo:

Hablante A: ¡¡*La puerta!!*

Hablante B: ¡¡*Estoy en la ducha!!*

Hablante A: ¡¡*Vale!!*

Entender de manera correcta este diálogo implica acceder a una serie de significados que no están contenidos en el significado correspondiente a cada una de las palabras pronunciadas. La primera intervención, *la puerta*, tiene como referente la parte de la casa por la que se accede a la misma; sin embargo, en este contexto, escuchado como un grito en una casa, los receptores lo interpretarían como "están llamando a la puerta, por favor, que alguien vaya a abrir". La contestación del hablante B sólo tiene sentido si se ha entendido de esta manera, ya que hace referencia a una actividad aparentemente no relacionada con la primera intervención (pensemos en un diálogo de forma similar como "*el techo*" y "*me estoy comiendo una paella*"). Lo que se entiende en la segunda intervención sería algo así como "he reconocido que quieres que vaya a abrir la puerta, pero yo no puedo hacerlo porque *estoy en la ducha*". La intervención final, *vale*, asume que el primer hablante ha reconocido la excusa como válida y por tanto podría ser interpretada como "entiendo que al estar en la ducha te pilla mal ir a abrir la puerta y

por lo tanto iré yo a abrirla". No hay manera en que un acercamiento de combinación sintáctica de significados léxicos pueda llegar a este tipo de interpretación. Aunque este ejemplo pueda parecer un caso muy extremo, el lenguaje, en un número muy elevado de casos, funciona de esta manera.

La "simple" combinación sintáctica de los significados de dos palabras es mucho más compleja de lo que pudiera parecer. Examinemos un caso concreto: conocemos el significado de las palabras *azul*, *rojo* y *ojo*, y conocemos que existe una regla de combinación Nombre + Adjetivo. Sin embargo, no está claro porque *ojo azul* señala a una parte del ojo (el iris), mientras que *ojo rojo* señala a otra parte distinta del ojo, la esclerótica. Predecir qué aspecto del significado de una palabra va a ser relevante para la combinación con otra palabra es extremadamente difícil de predecir: "un vaso de plástico" es un vaso hecho de plástico, un "vaso de vino" es un vaso lleno de vino (no un vaso hecho de vino) y un "vaso de postre" no es un vaso hecho de postre, ni un vaso para poner el postre en él, sino un vaso para poner en él lo que sea que se beba mientras se toma el postre.

Una gran cantidad de estos efectos pueden ser explicados si asumimos un acercamiento al lenguaje de tipo *inferencial*, en vez de como codificación. En vez de considerar que el lenguaje codifica los significados directamente, el acercamiento inferencial asume que el lenguaje proporciona una "pista" más, que sirve para que, en conjunción con la comprensión que hablante y oyente tienen de la situación, el oyente pueda recuperar la intención comunicativa del hablante efectuando las inferencias necesarias. Ninguna explicación de cómo funciona el fenómeno lingüístico puede estar completa sin esto: podemos centrarnos en una parte, como la morfología o la sintaxis, pero estas explicaciones no pueden aspirar a explicar cómo tiene lugar realmente el fenómeno lingüístico.

Una de las maneras de cruzar esta separación entre lo que se dice y lo que se entiende es asumir que el lenguaje es una actividad colaborativa, tal y como mantienen autores como Michael Tomasello o Herbert Clark. En este sentido, hablar es un tipo de actividad similar a otras como bailar el vals, boxear o jugar al tenis: se hace con otra persona, de manera que describir la actividad lingüística sin tener en cuenta el papel de los dos interlocutores implicados no permite ver el escenario completo. La recuperación del significado lingüístico tiene sentido en la confluencia entre un hablante, que emite unos sonidos relacionados con una intención comunicativa dada en un contexto concreto, y el esfuerzo de recuperación de su intencionalidad efectuado por otra persona

implicada en el proceso. Es una situación similar aunque de signo contrario a la descrita por el refrán español "no hay peor sordo que el que no quiere oír".

Esta visión del lenguaje como actividad colaborativa se enmarca además en otra de las características que se aducen como específicas del género humano: nuestra gran capacidad de colaboración, basada en nuestras capacidades de teoría de la mente. Autores como los anteriormente mencionados afirman que es precisamente esta capacidad lo que nos distingue de otros primates cercanos evolutivamente a nosotros. La cooperación humana es distinta a la de otras especies, precisamente porque está basada en una lectura muy eficaz de la intencionalidad de los otros: somos los únicos animales capaces de "leer la mente" de la otra persona y efectuar suposiciones muy cercanas de cómo se siente, qué desea y cuál es su propósito. Este tipo de habilidad es básico para la recuperación del significado lingüístico.

Los niños aprenden en primer lugar a establecer un marco atencional conjunto con los adultos. Desarrollan habilidades de seguimiento de la mirada que les permiten identificar a qué le está prestando atención una persona implicada en un intercambio con ellos. Por otro lado, sus habilidades estadísticas de aprendizaje de secuencias les permiten aprender la estructura de las interacciones con los adultos, en muchos casos formadas por actividades basadas en rutinas repetitivas con un orden más o menos fijo y con metas globales y submetas parciales. Es el caso de la rutina del baño, de la comida, de ir al parque, etc. De esta manera, cuando los niños escuchan un objeto lingüístico, lo interpretan siempre en el contexto de estas estructuras, que facilitan la recuperación de la intencionalidad del hablante, puesto se conoce el objetivo y la dirección de la actividad que sirve de contexto a la expresión utilizada.

El niño aprende pues a establecer el signo lingüístico de manera triádica: todo signo lingüístico implica a un emisor, un receptor y el objeto de referencia del signo. La palabra "pato", por lo tanto, es un sonido emitido por un hablante para que un oyente preste atención preferencial a un objeto del mundo presente en el marco atencional compartido por ambos. Existen multitud de experimentos que muestran cómo los niños establecen el significado de las palabras utilizando este tipo de estrategias (una revisión podría ser Tomasello, 2008). Estas estrategias también influyen en el tipo de sintaxis utilizada; Ibbotson, Lieven y Tomasello (2014) mostraron como niños de tres años, por un lado, y adultos por otro, utilizaban la voz activa o la voz pasiva para describir una escena con dos animales implicados en una acción transitiva, dependiendo de cuál de los animales recibiera una atención preferente, evidenciada por la dirección de la mirada

de su interlocutor; este era el animal que escogían como sujeto, fuera agente (en cuyo caso utilizaban la voz activa) o paciente (en cuyo caso utilizaban la voz pasiva).

3. Conclusión final

Tal y como expresó el premio Nobel de Medicina François Jacob, la evolución no es una ingeniera, ni una inventora, sino una "manipuladora"[†] de los mecanismos con los que se encuentra, experimentando con ellos de manera aleatoria en busca de nuevos funcionamientos que otorgen una ventaja evolutiva a un organismo biológico dentro de su nicho ecológico. Este es el sentido del acercamiento al lenguaje desde la búsqueda de capacidades de dominios generales. Esta "hoja de ruta" investigadora, por lo tanto, está firmemente basada en lo que conocemos sobre el funcionamiento de la evolución darwiniana, minimizando hasta cierto punto los "saltos evolutivos". De hecho, este tipo de acercamientos al lenguaje, desde nuestras capacidades generales, pueden incluso dar la vuelta al problema inicial, sugiriendo que el lenguaje humano tiene la forma que tiene y funciona de la manera que lo hace debido precisamente a las restricciones y características de nuestras capacidades generales: a nuestras capacidades de encontrar regularidades estadísticas, a las restricciones impuestas por nuestros sistemas perceptivos, de memoria o de atención, por ejemplo. La idea es que la interacción entre todas estas restricciones surgidas de capacidades no lingüísticas forman una especie de "cuello de botella" que hace que ciertos aspectos lingüísticos tengan un mayor nivel de aprendibilidad por los niños, y que son precisamente estos los aspectos que se perpetúan y pasan a formar parte de manera recurrente en los diversos códigos lingüísticos que encontramos en el mundo.

Aunque la discusión entre los partidarios de acercamientos "específicos del dominio" y los de acercamientos de "dominio general" todavía no ha alcanzado el consenso que una cuestión tan básica y tan crucial merece, podemos terminar el presente trabajo con dos citas que demuestran que todos los investigadores implicados estamos finalmente de acuerdo en la manera en que este tipo de debates deben ser solucionados:

Linguists and biologists, along with researchers in the relevant branches of psychology and anthropology, can move beyond unproductive theoretical debate to a more collaborative, empirically

[†] La palabra en inglés es "tinkerer" y la frase es "evolution is a tinkerer, not an engineer".

focused and comparative research program aimed at uncovering both shared (homologous or analogous) and unique components of the faculty of language (Hauser, Chomsky & Fitch, 2002: 1578)

Continued research, rather than dogma, is needed in order to render the most significant progress on the question of domain-specificity and domain-generality in language acquisition (Saffran & Thiesen, 2007: 80)

4. Referencias

- BICKERTON, D. (1990). *Species and Language*. University of Chicago Press, Chicago.
- BYBEE, J. (2006). From usage to grammar: the mind's response to repetition. *Language*, Vol. 82, No. 4. (2006), pp. 711-733
- CHOMSKY, N. (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, MA
- CHOMSKY, N. (1975). *Reflections on Language*. Pantheon, New York.
- GOLDBERG, A. (2011). Corpus evidence of the viability of statistical preemption. *Cognitive Linguistics*, (2011) Vol. 22 1: 131-154.
- HAUSER M.D., CHOMSKY, N. Y FITCH, W.T. (2002). The faculty of language: what is it, who has it, and how did it evolve? *Science* 2002 Nov 22;298(5598):1569-79.
- IBBOTSON, P. (2013). The scope of usage-based theory. *Frontiers in Psychology*, 4: 255
- JACKENDOFF, R. (2002). *Foundations of Language*. Oxford Univ. Press, New York
- JACOB, FRANCOIS (1977). Evolution and Tinkering, *Science* 196 (1977), 1161-1166.
- JENKINS, J. (2000). *Biolinguistics*. Cambridge Univ. Press, Cambridge
- PINKER, S. (1984). *Language learnability and language development*. Cambridge: Harvard University Press
- PINKER, S. (1987). The bootstrapping problem in language acquisition. In B. MacWhinney (Ed.), *Mechanisms of language acquisition*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- PINKER, STEVEN (2009). *El instinto del lenguaje: cómo la mente construye el lenguaje*. Madrid, Alianza Editorial.
- REEDER, P.A., NEWPORT, E.L., & ASLIN, R.N. (2013). From shared contexts to syntactic categories: The role of distributional information in learning linguistic form-classes. *Cognitive Psychology*, 66, 30-54.
- SAFFRAN, J.R., ASLIN, R.N., & NEWPORT, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, 274, 1926–1928

- SAFFRAN, J.R., & THIESSEN, E.D. (2007). Domain general learning capacities. En E. Hoff and M. Schatz (Eds.) *Blackwell Handbook of Language Development*, pp. 68-86. Oxford: Blackwell Publishing
- SAUSSURE, F. (1993). *Curso de Lingüística General*. Alianza Editorial, Madrid (Obra original publicada en 1916).
- TOMASELLO, M. (1999). *The cultural origins of human cognition*. Cambridge, MA: Harvard University Press.
- TOMASELLO, M. (2003) *Constructing a language: A usage-based theory of language acquisition*. Cambridge, MA: Harvard University Press.
- TOMASELLO, M. (2008). *Origins of human communication*. Cambridge, MA: The MIT Press.
- WRAY, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press