EL SIGNO AFIJAL EN LA MUESTRA TEXTUAL
CLAVES PARA ENTENDER EL DESCUBRIMIENTO AUTOMÁTICO DE MORFEMAS

Alfonso Medina Urrea

EL COLEGIO DE MÉXICO

ÍNDICE

Índice de tablas

INTRODUCCIÓN:
LA CADENA HABLADA Y LA CADENA ESCRITA

Cuando escuchamos con atención a alguien que queremos o admiramos o cuyo mensaje nos interesa o nos seduce, solemos identificar qué cosas, personas, ideas o sensaciones nombra, qué dice que hacen o sucede con ellas, dónde ocurre lo que sucede y cómo y cuándo ocurre. Gracias a que compartimos un idioma, esto es, a que conocemos los signos que nuestro interlocutor produce y la manera en que los va estructurando, logramos comprender lo que va diciendo. Si no estamos familiarizados con alguna de sus palabras, intentamos inferir su significado a partir del mensaje, a partir de lo dicho y de lo que suponemos que se va a decir. Por otra parte, si algo ocurre en un orden que no esperábamos, podemos pensar que nuestro interlocutor se equivocó o está tratando de decir algo distinto a lo que anticipábamos y nos esforzamos en darle un nuevo sentido.

Desconocer alguno de los signos en la cadena hablada es una experiencia diferente a la de encontrarnos con un ordenamiento inusual o equivocado de los signos. Una palabra desconocida nos recuerda que no sabemos muchas cosas del mundo y nos invita a inferir su significado o a investigarlo en otro lado, como en un diccionario, en una enciclopedia, en Internet, o preguntándole a alguien. Una estructura sintáctica poco típica o mal formada nos fuerza a reanalizar el ordenamiento de los signos para reinterpretar el mensaje y buscar que tenga sentido. Casi siempre lo logramos, puesto que somos máquinas de descartar ambigüedades y encontrar significados, incluso de inventarlos donde no los hay. Por otra parte, si percibimos que está mal formada, puede ser que en un instante deduzcamos las causas o conjeturemos sobre las consecuencias de esa malformación.

Sabemos que la estructura de lo que oímos o leemos está dada por la secuencia de los signos y por cómo se combinan unos con otros. Hay signos que nos revelan más sobre el mensaje del interlocutor y menos sobre su estructura. También hay signos que dan pistas sobre la estructura y nos hablan menos del mensaje mismo. Conocemos a los primeros como signos de contenido y solemos concentrar nuestra atención en ellos. En un momento dado del discurso, cuando tratamos de adivinar qué signo es el siguiente, los de contenido son los que menos podemos predecir y, una vez que han sido pronunciados, son los que más pueden habernos sorprendido. Es interesante que, como veremos adelante, el tamaño de esa sorpresa crece y decrece con la cantidad de información que intenta transmitir nuestro interlocutor.

En cuanto a los signos con función gramatical, aquellos que nos dan pistas sobre la estructura del mensaje, solemos estar menos pendientes de ellos. Sabemos que ocurren mucho, sus significantes suelen ser breves y comunican relativamente menos información que los de contenido. La que comunican tiende a ser de tipo gramatical. Por eso, los conocemos como partículas o signos gramaticales. Ocurren tanto que ya ni siquiera los vemos conscientemente, aunque sin duda los percibimos en tanto organizan el mensaje. De repente, cuando los encontramos en la posición equivocada o en lugar de otros que inconscientemente esperábamos o, simplemente, cuando no están donde consideramos que deberían estar, se vuelven otra vez visibles.

De esta manera, cada signo gramatical suele ocurrir en ciertos lugares y no en otros. Se aglutina con los demás de maneras determinadas para estructurar los enunciados del discurso: una preposición precede una expresión nominal; un adverbio orbita alrededor de una expresión verbal; un enclítico sigue al verbo y un proclítico lo precede; un afijo se pega a un signo de contenido o a otros signos afijales, etc. Esta morfotáctica es un síntoma de que la lengua es un sistema económico, lo cual se manifiesta en el hecho de que los hablantes, como nuestro interlocutor, puedan nombrar un número potencialmente infinito de cosas,

acciones y situaciones, sin tener que conocer o recordar una cantidad descomunal de signos.

La lengua hablada es diferente a la escrita en muchas maneras, pero la experiencia de escuchar a un interlocutor guarda ciertas similitudes con la de leer un texto. Las palabras gráficas se suceden una tras otra, unas son de contenido y otras gramaticales. De nuevo, las primeras transmiten el mensaje y las segundas lo estructuran. Mucho se pierde en lengua escrita: dejamos de ver los gestos del interlocutor y no hay pistas de la entonación. En cambio, contamos con signos de puntuación que orientan al lector en cuanto a la estructura sintáctica del mensaje y dan ritmo a su lectura. Los espacios separan las palabras gráficas, aunque no separan las bases o lexemas de los afijos. Las comas marcan fronteras entre frases o sintagmas y los puntos separan enunciados simples o complejos. La lengua hablada es fugaz. Si acaso, puede grabarse con un dispositivo de audio y volverse a escuchar o transcribirse. La lengua escrita y transcrita permanece plasmada en algún medio físico. De allí que, como veremos adelante, podamos estudiar mejor con ella las relaciones económicas entre los signos. En particular, veremos que estas relaciones permiten descubrir las fronteras entre bases y afijos y que, si no existieran los espacios entre las letras, tal vez nos permitirían descubrir las fronteras entre las palabras mismas.

Las relaciones económicas y la fuerza de atracción entre los signos

La alusión más antigua al carácter económico de la lengua es probablemente la de Marco Terencio Varrón, quien, algunas décadas antes de nuestra era, se percató de la carga que supone, para la memoria, la carencia de afijos derivativos y flexivos:

> La "declinación" [tanto de nombres y adjetivos como de conjugación verbal] se ha aplicado no sólo a la lengua latina, sino a la de todos los

> hombres, por una razón útil y necesaria. De no haberlo hecho así no podríamos aprender un número tan grande de palabras (ya que las formas naturales en que los vocablos se declinan son infinitas), y aunque las hubiéramos aprendido no podríamos descubrir a partir de ellas qué sistema [sic] las relaciona entre sí[1]. En cambio, ahora sí podemos percibirlo porque se trata de algo semejante, de algo que ha derivado; [...] Dos son, en general, los orígenes de las palabras: la imposición y la flexión. La primera viene a ser la fuente; la segunda, el río. Los hombres quisieron que las formas "impuestas" fueran las menos posibles, con el fin de aprenderlas cuanto antes; y que las "flexionadas" fueran el mayor número posible, para que todos pudiesen emplear aquellas que fuera necesario utilizar[2].

Muchos siglos después, entre los lingüistas mecanicistas estadounidenses, para quienes era natural suponer que los fenómenos del lenguaje, sus causas y sus efectos son observables, el trabajo con corpus textuales o transcripciones de lengua hablada se volvió esencial. Sin embargo, como se sugirió arriba, en los corpus no se puede observar todos los fenómenos lingüísticos. Además de que faltan los gestos del interlocutor y las pistas de la entonación, se escapan otras pistas lingüísticas y extralingüísticas, pasando por los procesos mentales más diversos que, si bien no se manifiestan abiertamente en la escritura, pueden dejar en ella huellas de diversas índoles, poco evidentes, incluso para la mirada del experto.

Por otra parte, entre aquellos lingüistas para los que el lenguaje no se puede caracterizar exclusivamente en términos de causas y efectos visibles al investigador, se ha observado que las palabras se relacionan entre sí para combinarse en estructuras más complejas y que la fuerza

[1] Del latín: "nisi enim ita esset factum, neque discere tantum numerum verborum possemus (infinitae enim sunt naturae in quas ea declinantur) neque quae didicissemus, ex his, quae inter se rerum cognatio esset, appareret" [edición de Roland G. Kent (Varro 1938 [47-45 a.C.], 373)].

[2] Varrón, *De lingua Latina*, tr. Manuel-Antonio Marcos Casquero, VIII, 3-5 (1990 [ca. 40 a.C.], 292-295).