

Lingüística Matemática

Participantes: J.A. Bellón Cazabán, E. García Camarero, M. Molóndez Rolla, A. Melizo Moya, M. Parra Pozuelo, J. Seguí de la Riba, V. Sánchez de Zavaleta, R. Trujillo.

Reuniones: Se celebraron los días 21 y 28 de febrero. Las próximas serán los días 14 y 28 de marzo a las 4 de la tarde.

Comunicaciones:

El día 21, R. Trujillo y J.A. Bellón, profesores de las Facultades de Letras de La Laguna y Granada respectivamente, presentaron las siguientes comunicaciones:

La Economía en el Sistema Fonológico

Saussure hizo ver que una lengua representa un conjunto de diferencias de ideas unido a un conjunto de diferencias fónicas. Pero un sistema de signos donde las unidades se diferencien sin más sería costosísimo para la memoria: es necesario que la diferenciación se dé sobre la base de una cierta identidad, de suerte que con un número muy bajo de unidades fónicas distintivas pueda construirse un gran número de signos y que con la combinación de éstos pueda expresarse un número infinito de pensamientos.

A la unidad distintiva mínima llaman los lingüistas fonema. La sustitución de un fonema por otro dentro de un segmento es suficiente para alterar su significación. Por esto el fonema es una unidad funcional. La relación de un fonema con todos los que pueden sustituirlo en un punto cualquiera de la cadena hablada recibe el nombre de oposición fonológica.

Ahora bien; es evidente que el número de signos que puede tener una lengua está en relación con el número de fonemas y con la cantidad de éstos que puedan entrar en la palabra. Una lengua que sólo admitiese como máximo palabras de cuatro fonemas, no podría tener un vocabulario tan extenso como otra que admitiese como máximo seis, suponiendo que ambas poseyeseen sistemas fonológicos con igual número de unidades. Pero es igualmente evidente que la lengua que sólo admite palabras de cuatro fonemas podría aumentar considerablemente su vocabulario ampliando su sistema fonológico. No debe extrañarnos, pues, el hecho de que lenguas en que la longitud media⁸ de la palabra es corta, como el inglés o

⁸ Hablamos de longitud media porque el hecho de que en una lengua las palabras sean normalmente cortas, no impide que pueda haber un número reducido de palabras más o menos largas.

el francés, tengan sistemas fonológicos muy amplios, mientras lenguas de palabras normalmente largas, como el español, tengan sistemas fonológicos reducidos.

Está claro que un sistema fonológico amplio (es decir, costoso) sólo puede ser económico en un sistema léxico de palabras cortas (esto es, de bajo costo). En una lengua donde, por término medio, sean largas las palabras (costo de ejecución elevado), un sistema fonológico demasiado amplio resultaría innecesario, ya que las palabras se diferenciarían entre sí por muchos rasgos a la vez, lo cual supondría una redundancia excesiva. Sabemos, sin embargo, que en todas las lenguas existe este tipo de redundancia, necesaria en ciertos límites, ya que las condiciones del mensaje nunca son óptimas (ruidos, falta de atención - por parte del que escucha, etc.). Esta es la razón por la que una lengua no agota jamás todas las posibilidades de combinación de sus fonemas. Sin embargo, el margen de redundancia tiene también sus límites, y habría que determinar también, por decirlo así, la cantidad de redundancia que admiten las diferentes lenguas, comparando las posibilidades estrictas del sistema con las realizaciones efectivas del mismo (es decir, cuantificando las unidades innecesariamente empleadas, teniendo en cuenta las posibilidades del sistema).

Descontando, pues, este margen de redundancia, parece evidente que la longitud media de la palabra y el número de fonemas de una lengua deben tender a guardar un cierto equilibrio, un condicionamiento mutuo. Y parece igualmente cierto que cuando este equilibrio se ha alterado (desgaste fonético, como en el francés, o introducción de neologismos largos, como en el español renacentista, y también en el francés) debe tender a restablecerse a pesar de todas las presiones que pueda haber en contra (norma culta, etc.). Como hace ver A. Martinet "La energía gastada con fines lingüísticos tiende a ser proporcional a la cantidad de información transmitida".

Hechos tan extendidos como el yeísmo o el seseo hacen pensar, por lo que se refiere al español, que el número de fonemas y la longitud media de la palabra están en desequilibrio; esto es, que son las condiciones mismas de la economía del sistema las que determinan tales alteraciones en su estructura. No nos parece insensato pensar que las condiciones de equilibrio puedan calcularse a partir de los factores que hemos considerado más arriba.

Si, por otra parte, tenemos en cuenta que el español renacentista perdió una serie de diferencias fonológicas precisamente en la época en que el caudal léxico se incrementó enormemente con neologismos casi siempre largos, cabe pensar sensatamente que la correlación entre la longitud media de las palabras y el número de fonemas ha sido un factor dinámico que ha condicionado en una cierta medida determinados cambios lingüísticos, y que nos permitirá prever, al menos, las posibles tendencias evolutivas en el futuro.

Como cabe suponer que las condiciones que determinan la redundancia se dan igualmente y en la misma proporción en las distintas lenguas, parece claro que un alarga-

miento o acortamiento de la palabra media hará aumentar o disminuir la redundancia más allá de sus límites normales, y, en consecuencia, el número de unidades diferenciales existentes podrá resultar corto o excesivo para las necesidades de la comunicación.

En el caso de resultar excesivo es evidente que el sistema no será económico, ya que contendrá y obligará a usar más unidades de las que la comunicación precisa (hará gastar demasiada energía), con lo cual es previsible la reacción, tratando de restablecer el equilibrio.

En el caso de resultar corto, la posibilidad de confusión crecerá, con lo que tenderá también a aumentar el gasto de energía necesario para que la comunicación sea suficiente. Esto podría hacer que una diferencia fónica accidental, adquiriese de pronto carácter funcional, valor fonológico.

Sin embargo, parece seguro que la reducción de las diferencias (oposiciones) está limitada a lo que podríamos llamar "puntos débiles" del sistema. Sólo pueden confundirse aquellas unidades que tengan en común varias propiedades distintivas (o alguna esencial) esto es, una cierta similitud acústica o articulatoria (en general, las oposiciones llamadas neutralizables y algunas del tipo J/y , $\text{s}/\text{θ}$, l/r , z/s , $\text{š}/\text{ž}$, etc.). Por eso es muy importante conocer el rendimiento (frecuencia) de las oposiciones entre elementos parcialmente semejantes de un lado y las que se dan entre elementos totalmente alejados dentro del sistema. Las oposiciones de bajo rendimiento deben tender a simplificarse siempre que se den las condiciones expuestas (exceso de fonemas, semejanza fonética, etc.) y siempre que la presión de la norma no lo impida. (nos referimos, claro está, sólo a las condiciones para la simplificación: a los puntos por donde el sistema "puede" ceder).

Si exponemos aquí estos hechos, por lo demás tan simples, es porque pensamos que con la ayuda de los computadores electrónicos puede exclarecerse plenamente lo que hasta ahora es mera intuición.

Un estudio de este tipo nos podría revelar:

- a) si existe una relación efectiva entre la longitud media de la palabra y el número de fonemas que componen el sistema fonológico de una lengua.
- b) qué unidades se comportan verdaderamente como fonemas y cuáles sólo en casos muy excepcionales.
- c) cuál es la longitud media de la palabra en una lengua y si ha sufrido modificaciones.
- d) si tales modificaciones han alterado el equilibrio del sistema fonológico.

El procedimiento a seguir podría ser: en líneas muy generales, el siguiente:

1. Clasificar las unidades léxicas (tomando como base ya el Diccionario, ya una serie de textos que hubieran sido previamente reducidos a Vocabulario, con sus índices de frecuencia, etc.) según su extensión en grupos de un fonema, dos fonemas, etc., para obtener así la extensión media y el tipo de palabra dominante. (Naturalmente, el programa deberá prever los casos en que varias letras representan un solo fonema o a la inversa, e interpretarlos debidamente).
2. Comparando todas las palabras de cada lista entre sí (y las de cada lista con su inmediata superior), podríamos obtener nuevas listas de palabras que se diferenciasen por un solo fonema (luego, quizá, por dos, tres, etc.), y determinar, en consecuencia, la frecuencia (rendimiento) de cada pareja opositiva. Por ejemplo, en la lista

[piso]	(piso)	
[biso]	(viso)	p/b
[liso]	(liso)	p/l b/l
[kiso]	(quiso)	p/k b/k l/k
[giso]	(guiso)	p/g b/g l/g k/g

nos interesa saber primero que "piso" se distingue de "viso" como p/b, y de "liso" como p/l, etc.; pero también que "viso" se distingue de "liso" como b/l, etc. Descubriremos así todas las oposiciones que funcionan y su frecuencia en cada posición de la palabra. Hecho esto, centraremos nuestro interés en las parejas cuya base de comparación sea lo más amplia posible (p/b, k/g, etc.), que son seguramente las que pueden ofrecer un mayor flanco a la confusión (siempre que razones de economía lingüística lo aconsejen). Así, pues, para que se den las confusiones que nos interesa investigar, parece ser condición previa, 1) que la oposición tenga poco rendimiento, 2) que los miembros de la oposición posean similitud fonética.

3. Estos mismos procedimientos deberán aplicarse al español de las distintas épocas. Sólo sobre esta base podremos determinar si efectivamente ha variado el tamaño medio de la palabra, si estas oscilaciones coinciden con las alteraciones del sistema fonológico, cuál es el rendimiento de las parejas opositivas en cada época, e, incluso, cuáles son los puntos del sistema fonológico actual en los que puede preverse un cambio, siempre que se haya demostrado previamente que aún el número actual es antieconómico.

R.T.

Para la caracterización estilística de un texto literario con ayuda de un ordenador electrónico

- 1) - La crítica literaria de nuestros días busca una línea de objetividad máxima en todas sus manifestaciones. Esa objetividad debe y puede conseguirse a base de dos premisas fundamentales:
 - a) Análisis exhaustivo del texto
 - b) Estudio profundo del entorno donde dicho texto ha salido a la luz

Si relacionamos convenientemente ambas coordenadas obtendremos sin duda una aproximación fructífera a ese todo que una obra literaria constituye.

- 2) - Pero antes de nada conviene sentar otra premisa. La labor del crítico puede resumirse según nuestro criterio en dos palabras: reunir y exponer

Claro que para reunir y exponer los elementos significativos de una obra de arte es preciso desmontarla y desmenuzar todas las piezas que la componen.

- 3) - De acuerdo con lo anteriormente dicho resulta obvio que para desmenuzar la obra de autores prolíficos, o para enfrentarnos a cualquier producción total (ya sea de una sola persona, ya de un género o de una época) se necesitarán unos poderosos medios auxiliares con los que sustituir la limitación de memoria que todos los humanos poseemos.

Hasta la fecha, estos medios han venido siendo los ficheros, si bien hay que decir que gran parte de los críticos se han permitido el lujo de prescindir incluso de ellos y han montado sus estudios sobre "intuiciones reveladoras" acertando acaso alguna vez y equivocándose casi siempre.

- 4) - Si se concede que para enderezar un panorama crítico nada halagüeño habría que realizar una labor inmensa partiendo de la documentación exhaustiva y del registro de los fenómenos que deseen estudiarse, habrá que convenir en la inmensa ayuda que los ordenadores electrónicos podrán prestarnos en este terreno.
- 5) - Pero hay más; una vez incorporada una obra cualquiera a la memoria del ordenador, no sólo podremos acortar enormemente el trabajo más fastidioso para llegar a un análisis de contenido, sino que tendremos las bases para sucesivas profundizaciones formales y así podremos estudiar convenientemente todas las facetas que sea preciso.

- 6) - Mis proyectos en este sentido son totalmente personales y aislados. Simples tanteos de un método adecuado, consignación de directrices y exploración de caminos que en el futuro pueden rendir un importante servicio a la crítica literaria. Hasta ahora mi estudio sólo va orientado al contenido. Se trata de encontrar las palabras claves, que se nos revelan por un mayor índice de frecuencia y, conocidas éstas, estudiarlas en sus respectivos contextos. De esta manera podremos adentrarnos en el mundo personal de cualquier autor y estudiar las coordenadas más representativas (el sentido en suma) de su obra.

Otro paso puede ser el estudio de los recursos estilísticos que emplea más frecuentemente (uso de los sentidos, nociones de geografía o historia que aparecen en su obra, alusiones al mundo vegetal o al animal, etc.).

Y, por fin, un tercer paso consistiría en estudiar los temas capitales que a través de nuestro estudio anterior se nos pueden haber revelado como más significativos, agrupándolos por familias de significados y adscribiéndolos a un denominador común como podría ser la muerte, Dios, el tiempo, la guerra, el recuerdo, etc.

- 7) - Ahora bien, entrando en el terreno de los proyectos habría que adjuntar unas limitaciones para este tipo de análisis literario.

- a) - Por una parte, con estos datos podríamos obtener una valoración adecuada de una producción concreta, pero nunca podríamos ponerla en relación con otras obras coetáneas, anteriores o posteriores. Para ello habría que hacer estudios semejantes con autores y géneros de todas las épocas, de modo que poseyéramos la documentación suficiente sobre la que apoyar cualquier conclusión auténticamente importante.
- b) - Un segundo escollo se nos presenta en el hecho de que cualquier producción artística consta de contenido ideológico y de contenido formal y - ambos están tan sistemáticamente interrelacionados que el estudio sin el otro degeneraría en una visión unilateral falsa y desproporcionada. Por lo cual un segundo paso en nuestra labor de búsqueda podría consistir en aplicar el ordenador a un análisis de tipo formal que nos revelara los recursos que el artista empleó en la ejecución de su obra. A la vez este análisis tendría que basarse en la comparación con otros textos contemporáneos y no contemporáneos para poder insertar en la historia la aventura del hombre en su constante búsqueda de la belleza artística.

Habrá que programar:

Ordenación alfabética de todas las formas fichadas especificando a continuación:

- nº de apariciones en total y frecuencia en %
- nº de apariciones en cada una de las secciones y frecuencia en %
- Especificación de página y verso donde aparece cada palabra
- nº de formas de cada una de las secciones y nº de formas en total
- Ordenación de las palabras con el criterio de su índice de frecuencia de aparición en toda la obra y por secciones.
- K. index.

J.A.B.

En la reunión del día 28 García Camarero apuntó las siguientes notas:

Tendiendo a encontrar una formalización de la semántica de un lenguaje, nos ha parecido que puede ser útil observar el comportamiento semántico de un lenguaje artificial, en particular de un lenguaje de ordenador, ya que en éstos el significado de los mensajes es claro y en última instancia pueden referirse como la configuración electrónica en un instante o en un intervalo de tiempo determinado.

El alfabeto básico usado es:

$$\{0, 1\}$$

o cualquier secuencia formada con estos símbolos. El significado elemental de ello, podría decirse simplificando, es "no pasa corriente", "pasa corriente". Circuitos elementales realizan las funciones de los conectivos lingüísticos "o", "y", "no", con el significado común. La presencia de "corriente" en un determinado cable puede significar "poner en marcha una máquina". Todos estos significados son fácilmente codificables y formalizables. Podremos definir un conjunto de significados y ver la función que liga a cada cadena de ceros y unos con su significado.

Ejemplifiquemos ésto con algunos tipos simples de autómatas, sea E_1 un alfabeto finito que llamaremos alfabeto de entradas, sea Q el conjunto finito de las posibles entradas del autómata; y sea T una función definida en $E_1 \times Q \rightarrow Q$, así el sistema

$$A_1 = \{E_1, Q; T\}$$

será un autómata del que se conocerá su comportamiento interno (T), al incidir en él los

estímulos E_1 . Al carecer de respuesta el significado de cada señal vendrá dado por la modificación en su situación interna, y los símbolos de entrada formarán un lenguaje compuesto por monosemas sin articular al ser todos ellos independientes y no poder formarse cadenas lingüísticas.

Si a este autómata le agregamos un conjunto de símbolos de respuesta, o alfabeto de salida E_2 , vemos que su comportamiento se enriquece al ser capaz no sólo de modificarse internamente sino de reaccionar sobre el medio con una respuesta; en este caso tendríamos

$$A_2 = \{E_1, E_2, Q; T, S\}$$

donde S es una función definida en $E_1 \times Q \rightarrow E_2$

En este caso el significado vendría dado no sólo por la modificación de su estado sino por su respuesta, que en general comprenderá una acción. Pero aún las señales de entrada formarán un lenguaje no articulado. No así si suponemos el siguiente tipo de autómata

$$A_3 = \{E_1, E_2, Q; T, S; q_0, F\}$$

donde q_0 será un estado inicial en el que suponemos se encuentra el autómata antes de recibir un mensaje, y F un subconjunto de Q , de estados finales, es decir de aquellos estados en los cuales el autómata deja de recibir señales de entrada para emitir una respuesta y S una función $E_1 \times F \rightarrow E_2$. En este caso los mensajes en general estarán formados por cadenas de símbolos del alfabeto de entrada, y tendrán sentido, es decir les corresponderá un significado o respuesta, sólo aquellas cadenas que incidiendo en el autómata cuando se encuentra en su estado inicial (q_0) alcanza un estado final ($q_f \in F$) cuando han leído su último símbolo.

Así estamos ante un lenguaje articulado, provisto de una gramática, considerada ésta como el conjunto de todas las cadenas que hacen emitir una respuesta a A_3 , si llamamos L a este lenguaje dos cadenas $x, y \in L$ serán sinónimos si al incidir en A_3 emiten la misma respuesta. Es evidente que la relación de sinonimia, en este sentido, es una relación de equivalencia.

Por otra parte podemos ver que un lenguaje de máquina actúa sobre el universo bien definido, constituido por los elementos de la misma. La unidad de control es precisamente el autómata que al recibir el mensaje formado por una instrucción (compuesta por una secuencia de ceros y unos) emite una respuesta, orden, con significado que consiste en general en encargar a sus otros componentes la realización de una tarea específica.

La forma general de una instrucción máquina en su expresión más sencilla es:

OP	Dirección
----	-----------

en donde OP (operación) tiene un comportamiento verbal, y la dirección sería un complemento sobre el que actúa el verbo directa o indirectamente. Otra variante de una instrucción es:

OP	MOD	Dirección
----	-----	-----------

en donde MOD indica una modificación a realizarse en la dirección
E.G.C.

Generación de formas plásticas

Participantes:
Málaga: M. Barbadillo
Madrid: Alexanco, F. Alvarez Cienfuegos, F. Briones, M. de las Casas Gómez, E. Delgado, M. Fernández Barberá, I. Fernández Flórez, E. García Camarero, A. García Quijada, A. Martín, J. Montero, J. Peña, I. Ramos, G. Searle, J. Seguí, R. Sempere, S. Sevilla Portillo.
Valencia: V. Aguilera Cerní, J. Ma. L. Yturralde

Reuniones: Se celebraron los días 13 y 27 de febrero.
Próximas reuniones 13 y 27 de marzo.