

JUAN A. NUÑO

## SISTEMA DE INFORMACION Y LENGUAJE NATURAL

Puede entenderse 'lenguaje' de la manera más abstracta y funcional como "señal" o "sistema de signos", pero también muy específicamente como "la facultad que posee el hombre de poder comunicar sus pensamientos" (Lázaro Carreter, *Diccionario de Términos Filológicos*). Ello no quiere decir que el lenguaje (=sistema de signos) pueda ser visto "como (al modo de) un sistema de información", sino que el lenguaje en tanto señal forma parte del sistema de información, al que en definitiva se reduce.

Pero, además, el lenguaje, en tanto lenguaje hablado o facultad humana específica, no permite su reducción informática; no es únicamente la señal que emite el transmisor, pasa por el canal y decodifica el receptor. Es eso y algo más. Justamente ese "algo más" rompe la continuidad sistémica e incide en la integración del lenguaje hablado en una teoría general informativa.

— I —

Se va a preferir no manejar 'comunicación' en tanto sinónimo de 'información'. Conviene, por el contrario, hacer a ésta subconjunto de aquella. Se va a entender entonces por 'comunicación' el "proceso mediante el cual se transmiten señales", mientras que se reserva 'información' para una comunicación significativa, es decir, "transmisión de un tipo especial de señal denominada mensaje". En otras palabras: información es comunicación con significado.

Sabido es que cualquier sistema de comunicación descansa en operaciones probabilísticas. A partir del procedimiento ideado por Ehrenfest, Markov desarrolla su concepto de probabilidades encadenadas o "cadenas de Markov", que sirve para establecer que la probabilidad presente lo es en función de los sucesos pasados, algo que dicho matemáticamente tiene un valor, pero que si se declara con intenciones filosóficas (Hegel, por ejemplo) adquiere otro sentido. No menos sabido es el gran partido que la lingüística le ha sacado a las cadenas de Markov, cuya aplicación podrá ser tan elevada como, por ejemplo, medir las secuencias de vocales y consonantes en la novela de Pushkin, *Oigúin Onéguin*, o tan baja como para predecir que después de la expresión "en el caso" es muy elevada la probabilidad de que aparezcan los términos "de que".

Pero la probabilidad en los sistemas de comunicación no sólo queda reducida al instrumental estadístico proporcionado por las cadenas de Markov, sino que desempeña un papel fundamental, desde el momento en que determina la entropía misma del sistema.

Conocido es que en cualquier sistema energético la entropía es un efecto del intercambio de energía o de calor, una consecuencia de la famosa segunda ley de Clausius. Pero lo notable es que, en los sistemas de comunicación, la entropía, en tanto efecto de aquella ley, se erige en centro del sistema, ya que viene a coincidir con la información misma. Esto es: la entropía de un sistema de comunicación no se mide tan sólo por el ruido del canal, sino ante todo por la propia información transmitida. El desorden, la desorganización, el azar que caracteriza al fenómeno entrópico viene dado por la elevada probabilidad de elección que existe en la formación del mensaje. Dicho con definición ya consagrada: "información es la medida de la libertad con que se elige al seleccionar un mensaje". De donde: a mayor libertad de elección, mayor grado de incertidumbre y, por lo mismo, mayor cantidad de información. Se entiende que es 'incertidumbre' de la buena, es decir, de la que sufre el emisor del mensaje y no el receptor (en tal caso, sería 'ruido'), y la sufre ante

la cantidad de información a transmitir. Trátase de una aplicación de la expresión *embarras du choix*: la incomodidad de tener que elegir.

Si alguien decide transmitir el mensaje de que mañana saldrá el sol por el este, el grado de entropía que arrastra esa información es mínimo (no puede decirse que igual a cero, por aquello de que la inducción no legitima la regularidad de los fenómenos naturales). O lo que es igual: no se está informando prácticamente nada. Aumenta la entropía si se selecciona el mensaje de que el candidato X va a ganar en las elecciones L, y llegaría a dispararse la entropía si se decidiera seleccionar como mensaje la noticia de que, al día siguiente de esas elecciones, en honor a la victoria del candidato ganador, el sol saldrá por el oeste. Máxima información (1, en la medida logarítmica), máxima entropía, ya que si, en general, se dispone de  $n$  informaciones, al momento de transmitir las se presentarán  $2^n$  posibilidades o 'estados de información'. Por lo mismo, se define siempre a entropía de un sistema como el logaritmo del número de posibles estados, siendo arbitraria la base del logaritmo (es convencional escoger la base 2 y unidades binarias).

Se observará, en consecuencia, que lo que determina el nivel de entropía es la carga de información. Ahora bien, información es transmisión de un mensaje significativo, no una comunicación cualquiera. Luego: en la medida en que se aumente el nivel de información (posibles estados significativos y, por lo mismo, entropía en el sistema) se estará haciendo a ese sistema más dependiente de su fundamento probabilístico. Por lo mismo: el lenguaje (=sistema de signos) permite la transmisión de información y vincula al sistema de comunicación con la medida estadística de aquella. Y el lenguaje (=facultad de comunicar que tienen los humanos) que más información transmita, más vinculado estará a dicha base probabilística. Quiere decirse con esto que no sólo el lenguaje en tanto señal forma parte de un sistema (probabilístico, azaroso) de comunicación, sino que el tipo de lenguaje más cargado de información (lenguaje humano), por ese mismo hecho, se integra necesariamente en la teoría de sistemas.

## — II —

Pártase del esquema canónico de cualquier sistema de comunicación. A saber: la cadena que va direccionalmente desde una 'fuente de información' a un 'destino', pasando por el 'emisor', el 'canal', el 'receptor', sin olvidar la 'fuente de ruido'. Y por supuesto, la cadena misma en tanto 'señal', sin descuidar el resultado que es el 'mensaje' transmitido. Pues bien, cualquier teórico de la comunicación (y de la información, a mayor abundancia) repartirá sus problemas de transmisión de la información en tres grandes capítulos. Problemas técnicos, que son los del ruido y precisión material del mensaje transmitido; problemas semánticos, a saber, la interpretación del significado (la decodificación) por parte del receptor, todo ello comparado con la intención original del mismo significado en el punto de salida (el emisor, en el proceso codificador). Y por último, problemas influenciales de la información, que se refieren a los efectos deseados y producidos (o no producidos) por la transmisión del mensaje. Esta viene a ser la caracterización promedio de un sistema de comunicación. Compáresela ahora con la que propuso hace veinte años (años sesenta) Jakobson que comienza por distinguir seis elementos necesarios a toda comunicación: emisor, receptor, contacto entre ambos, código común, mensaje y referente del mensaje. No es muy diferente del esquema canónico; no sólo está la relación bipolar (emisor/receptor) y el código por el que se transmite el mensaje, sino que el 'contacto' equivale al 'canal', mientras que el 'referente', sobre el que se entiende y proyecta el mensaje, o bien pudiera hacerse coincidir con la 'fuente de información' o bien es un mecanismo auxiliar de decodificación del mensaje en aras de la precisión del significado. En cualquier caso, lo que interesa destacar no es ese paralelismo, que sería simplemente banal, pues Jakobson no hizo sino aplicar al lenguaje hablado el modelo general comunicacional, sino las funciones que según Jakobson corresponden a cada uno de aquellos seis elementos. Y así, dícese que al 'sujeto' (emisor) le corresponde una 'función emotiva'; al 'proceso', en la medida en que va a ejercer una acción a través del mensaje, le corresponde una

'función conativa'; necesitase también una función que Jakobson denomina 'fática' para poder sostener la comunicación en tanto tal, aun cuando no haya información (por ejemplo, cuando se dice "aló" por teléfono o "uno, dos, tres" o "probando", en un altoparlante); hay asimismo una 'función metalingüística' aplicada al propio código; otra que Jakobson llama 'poética', cuando es la estructura del mensaje la que se analiza, y por último, la 'función referencial', cuando se analiza el mensaje en relación con lo que quiere decir.

Puede sostenerse que tanto los problemas a que alude el sistema de comunicación como las funciones que Jakobson agregó al sistema especial que es la lengua hablada apuntan a lo mismo: a algo que está más allá del mero esquema de la comunicación. En efecto, vuélvase a los tres problemas informáticos: si se descuentan los llamados problemas técnicos (que son aquellos sobre los que trabaja el ingeniero de sistema para mejorar materialmente las transmisiones), aparecen dos tipos de problemas que bien pudieran calificarse de "filosóficos". Son los problemas semánticos, de significado, y los influenciales o de efectos. Conocida es la maraña de problemas que hay tras la voz 'semántica': significado, sentido, referencia, diacronía, sincronía, sinonimia. Y por si fuera poco, sucede que también aquí hay que atender a las 'intenciones' del emisor del código, justo esa zona de problemas que ciertas escuelas del análisis filosófico del lenguaje dieron en llamar 'fuerza ilocucionaria' del discurso, y que sirven para situar la atención en plena turbulencia de la subjetividad. Pero además, si hay que atender y resolver problemas influenciales, esto es, relativos a la dimensión 'perlocucionaria' del discurso, se amplía el campo de operaciones al conjunto social, saltando de la zona del sicologismo a la de la intersubjetividad. Lo curioso es que otro tanto se encuentra en esas funciones que Jakobson agregara a los elementos de la clasificación del acto lingüístico. El 'referente' o 'función referencial' de Jakobson es justamente el trasunto del problema semántico de toda información: para que un signo tenga sentido, esto es, cumpla su función en el mensaje, tendrá que interpretarse (codificarse/decodificarse) en relación a un contexto o referente. Eso para no volver a mencionar el

subjetivismo al que alude la 'función emotiva' o el pragmatismo de la 'conativa'.

La lección a extraer tanto del trasfondo de ambos esquemas como de su aproximación es doble. Por un lado, aún en el caso más formalizado, matematizado y esquematizado de una teoría de la información, ésta no puede prescindir de un contexto en el que viene a insertarse. No hay transmisión comunicacional en estado puro: tan pronto se transmite la información, ésta depende del significado y el significado desplaza el centro de operaciones a regiones muy poco esquemáticas y nada cuantificables. Las cadenas de Markov servirán para medir el grado de probabilidad en la aparición de las secuencias lingüísticas, pero no sirven para determinar significado alguno.

Puede verse mejor con ayuda de un ejemplo de un orden igualmente termodinámico. Cualquier invento técnico, como por ejemplo, el muy elemental de la noria, es revelador de los límites cuantificativos. El consumo energético de la noria, el trabajo del burro, puede medirse por la cantidad de alimento necesario para mantener constante el peso del animal. Aunque la cadena de Markov mide propiamente probabilidades y, en el ejemplo de la noria lo que se miden son magnitudes, puede establecerse una relación. Pero del mismo modo que las cadenas de Markov ni registran ni califican significados de términos, la medición energética del trabajo empleado en la noria, a fin de convertir una energía en otra, no sirve para medir, establecer ni siquiera describir el significado de la idea de la noria, esto es, el hecho mismo de la concepción del invento, que tuvo que ser logrado mediante algún proceso de información. Pero desde Szilard, en su resolución del problema llamado "demonio de Maxwell", se sabe que información equivale a medición y, por lo mismo, a consumo de energía. Y lo que aquí interesa, toda información exige definición de significados. Luego, no se vacila en asegurar que, desde el momento en que toda información es comunicación significativa, toda teoría de la información tiene como referente (para emplear la expresión de Jakobson que, por lo demás, es de origen matemático) una teoría del signi-

ficado o teoría semántica. Y eso desplaza la informática hacia la filosofía del lenguaje.

Se dijo antes que era doble la lección. Por otro lado, en efecto, al tener que depender de la zona semántica y aún de la pragmática, se está privilegiando, aunque no se reconozca así, a un tipo de lenguaje sobre todos los otros: al lenguaje humano, natural, ordinario. Ya que, por muy perfectos que sean los lenguajes formalizados o notacionales o simbólicos, no tienen otra función comunicacional que la de servir de intermediarios, de eslabones entre el lenguaje común y la realidad estudiada o manipulada. Así como el "fortran" o el "cobol" son simples intermediarios semi-simbolizados entre la pobreza precisa del lenguaje binario y la ambigua riqueza del lenguaje común, cualquier empresa informacional se apoyará en última instancia en el lenguaje natural. En la introducción a los *Elementos de Matemáticas* del famoso cuan inexistente Bourbaki, se lee: "De suyo se entiende que la descripción del lenguaje formalizado se hace en lenguaje corriente, como sucede con la de las reglas del juego del ajedrez". Y sabiendo lo que eso conlleva, agregaban: "No entraremos en la discusión de los problemas psicológicos o metafísicos que supone la validez del empleo del lenguaje común en tales circunstancias...". Es decir, con todo y sus problemas, es imprescindible el empleo del lenguaje ordinario a la hora de describir cualquier otro lenguaje o sistema de comunicación. Algo semejante, desde otra zona de autoridad, decía Jakobson: "Para el género humano, todo sistema de comunicación se correlaciona con el lenguaje y dentro de la red general de la comunicación humana es el lenguaje el que ocupa el lugar predominante" (*Corrientes principales de la ciencia del lenguaje*). Por supuesto, aquí 'lenguaje' es "lenguaje humano ordinario".

Sería erróneo pensar que semejante preeminencia es una consecuencia o histórica o social o de simple conveniencia comunicacional. De eso se trata. De sostener que es inevitable la preeminencia del lenguaje hablado justamente en la medida en que el objetivo comunicacional es la información. O dicho de otro modo: el lenguaje que asegura el máximo nivel de información es el lenguaje común porque es el lenguaje en el cual el grado de libertad de elección del emisor es

máximo; por tanto, la incertidumbre es mayor y la medida de entropía es la más elevada. No se va a entrar aquí en el detalle de la importantísima categoría de 'redundancia' en el análisis de los mensajes, pero no cabe la menor duda de que la redundancia del lenguaje común le permite lograr un elevado nivel de variaciones, de donde surge la muy importante noción pragmática del 'estilo'. Con un ejemplo exagerado: en lenguaje matemático lo que importa es el mensaje transmitido ( $x + y$ ), pero no la forma de decirlo, puesto que las posibilidades de variación (la redundancia) del lenguaje matemático son mínimas. Es decir: en un lenguaje altamente formalizado no se distingue forma de contenido; sólo en el complejo lenguaje humano tiene sentido distinguirlas o confundirlas. En menos palabras: si es cierta la relación informática según la cual información igual a entropía, ello quiere decir que el caso del lenguaje hablado, de su debilidad (máxima incertidumbre selectiva) surge su fuerza (mayor potencial informativo).

La consecuencia postrera puede presentarse en forma de ligera paradoja. En la medida en que un sistema de comunicación genera información genera entropía. De ahí que todo lenguaje se integre en un sistema de información y que en particular el lenguaje humano, por su alta capacidad informativa, pueda ser referido a un sistema comunicacional. Sólo que aquello mismo que le permite ser clasificado en el sistema (la información) sirve para distanciarlo del sistema. La información se logra mediante captación de significado y ésta es una tarea vagamente filosófica (determinación de significados), vagamente psicológica (intenciones del emisor al seleccionar el mensaje) y vagamente sociológica (influencias posibles del mensaje). Al aumentar la información, esto es, al manejar el instrumento más informativo, el lenguaje común, se dificulta al máximo la solución de problemas semánticos, emotivos e inflenciales. De nuevo, repítese la paradoja: aquello mismo que sitúa al lenguaje en el centro del sistema de información (su entropía) sirve para desplazarlo de ese centro, a través de las dimensiones semántica y pragmática. Justamente el "algo más" que posee el lenguaje ordinario.