3 LA INFORMACION

3.1. El concepto de información

3.1.1. Información y probabilidad

En la Teoría de la Comunicación, el término **información** tiene un significado distinto de su uso corriente; en sentido técnico, la información contenida en un mensaje no tiene nada que ver con el "contenido" o "significado" del mismo: "un mensaje cargado de significado y otro lleno de tonterías pueden tener exactamente la misma información" (Shannon).

La información está en función de la *probabilidad de aparición* de los mensajes: cuanto *menos* probable es la aparición de un determinado mensaje –o parte de un mensaje–, mayor información contiene y recíprocamente. La información en sentido técnico es la medida de su imprevisibilidad.

Recordemos que la fuente selecciona el mensaje entre un conjunto de mensajes posibles. En el caso más sencillo, la fuente puede seleccionar solamente entre un conjunto de mensajes definidos (p. ej. un dado lanzado una vez). Otras veces, –y esto es más frecuente– la fuente selecciona sucesivamente una serie de símbolos del conjunto de que dispone y esta secuencia constituye el mensaje (p. ej. un número de lotería, un póquer de ases, una melodía, una frase escrita). En algunas fuentes, los mensajes no se componen de símbolos discretos sino que constituyen un "continuo" o función variable (p. ej. fluctuaciones de un estado físico, la palabra hablada, una imagen en movimiento...).

La información contenida en el mensaje depende de entre cuántos mensajes ha sido seleccionado, es decir del número total de mensajes posibles de la fuente y de la probabilidad respectiva que cada uno tiene de ser seleccionado en un momento dado. Es la medida de la *libertad de elección* que tiene la fuente al seleccionar el mensaje.

Imaginemos una fuente F1 que consiste en una bombilla A que puede estar encendida o apagada al azar. Esta fuente puede emitir por tanto *dos* mensajes: bombilla encendida y bombilla apagada. Lo que podemos codificar como: encendida = A = 1 y apagada = 0.

F1: A = 1, apagada = 0

Imaginemos ahora una fuente F2 que dispone de dos símbolos: A y B. Por ejemplo, un soporte con dos bombillas cada una de las cuales que puede estar encendida o apagada. Esta fuente puede emitir cuatro mensajes posibles: A, B, AB, y todas apagadas.

F2: A = 10, B = 01, AB = 11, todas apagadas = 00;

Imaginemos ahora otra fuente F3 que dispone de tres símbolos: A, B y C (soporte con tres bombillas). Esta fuente puede emitir *ocho* mensajes: A, B, C, AB, AC, BC, ABC y todas apagadas.

F3:	A = 100,	B = 010,	C = 001
	AB = 110,	AC = 101,	BC = 011
	ABC = 111,	todas apagadas = 000	

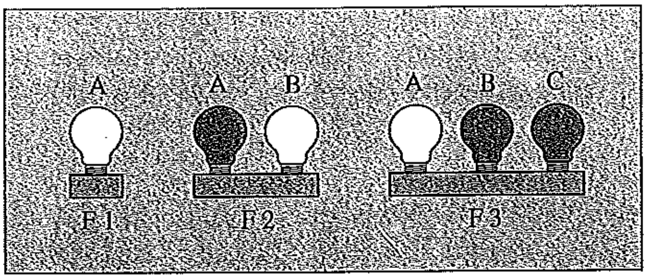


Fig. 7

Si cada fuente selecciona los mensajes estrictamente al azar, los mensajes de la fuente F2 contienen *doble* información que los de F1. Los de F3 contienen *el triple*.

La información mide la capacidad de elección de la fuente y está claro que F3 tiene tres veces más *libertad de elección* que F1 puesto que elige entre ocho opciones igualmente posibles (si, para mandar una felicitación, tengo ocho tarjetas del mismo precio donde elegir, mi libertad de elección es mayor que si sólo me sacan dos).

Un sistema de dos posiciones puede por tanto transmitir o almacenar dos mensajes posibles ("relé on-off", encendido/apagado, magnetizado / no magnetizado...). Si aumentamos los relés, la cantidad de mensajes posibles crece en progresión geométrica.

Para poder transmitir –o almacenar– cualquier tipo de mensaje, basta por tanto prever un sistema con un número suficiente de *unidades de dos posiciones* y someter el mensaje a una *codificación binaria*.

Los ordenadores digitales se basan en este principio[1].

Si la fuente no selecciona al azar sino que unas opciones tienen más probabilidad de ser seleccionadas que otras, las primeras tendrán menor información. Si resulta que F1 se enciende una vez de cada diez y las nueve restantes corresponden a "apagada", el mensaje "encendida" contendrá más información que "apagada".

Tal es el caso del embalse. Su estado más probable es que el agua esté por debajo del punto de peligro. La información "nivel de agua en A" es por tanto elevada.

Otros ejemplos:

Un dado es una fuente de información con seis mensajes posibles de idéntica probabilidad de aparición (1/6). Si lanzo un dado y sale un cinco, este cinco contiene menos información que cualquier carta extraída al azar de una baraja.

La baraja, en efecto, considerada como fuente de mensajes (cartas en este caso), dispone de cuarenta mensajes posibles. Supongamos que sale el Rey de Oros. Su probabilidad de aparición era de 1/40.

Un número de lotería, la combinación de la Bono Loto, el resultado de las quinielas, contendrían una cantidad de información muchísimo más elevada.

3.1.2. Información en un segmento del mensaje

Si en lugar de considerar un mensaje completo consideramos un segmento del mismo, la información que éste contiene dependerá asimismo de las opciones que podrían aparecer en su lugar, y de la previsibilidad estadística de las mismas.

Imaginemos una fuente que construye sus mensajes seleccionando sucesivamente una serie de símbolos tomados de un inventario del que dispone. Dos casos son posibles:

a) La probabilidad de aparición de cada símbolo es *independiente* del símbolo que ha aparecido anteriormente (código de fuente independiente de contexto).

Tal es el caso de un bombo de lotería. El bombo selecciona cada vez entre 10 símbolos (bolas numeradas de 0 a 9). Supongamos que emite sucesivamente 4, 0, 3, 2, 7, y construye el mensaje 40327.

Cada uno de estos símbolos tenía una probabilidad de aparición de 1/10, con independencia del que hubiera aparecido anteriormente.

La información que cada uno contiene es la misma pues en cada caso la fuente ha podido elegir entre diez opciones posibles.

b) La probabilidad de aparición de un símbolo o segmento de mensaje *depende de elecciones anteriores.*

Si el mensaje es un texto escrito en español, y la primera letra que aparece es una T, la probabilidad de que aparezca a continuación otra consonante (salvo R) es nula: el emisor sólo ha podido seleccionar una vocal o una R. Del mismo modo, si la primera palabra que aparece es "un", la probabilidad de que la siguiente sea un sustantivo o un adjetivo es muy alta. En la secuencia" café *con* leche" la preposición "con" es escasamente informativa pues sólo puede alternar con "sin". En "tonto *de* remate", la preposición carece de información. Después de la secuencia "colmillo de..." la probabilidad de "elefante" es altísima y la de "alcachofa", muy baja.

No dice el oyente nada nuevo ¿digo lo [illegible]?

1 La memoria de un ordenador se compone de varios miles o millones (según su capacidad) de unidades o bloques de ocho elementos cada uno de los cuales puede estar activado o no. En el origen eran realmente válvulas (de ahí el enorme tamaño de los ordenadores primitivos), luego fueron núcleos de ferrita, chips, finalmente microscópicas unidades de silicio hasta llegar a la miniaturización actual.

La función poética se basa precisamente en construir el discurso con secuencias inesperadas, es decir altamente informativas (Vid. infra. 7.2 y T. 15).

3.1.3. La información en sentido semántico.

La información en sentido técnico estricto es, como se ha visto, un concepto matemático. En el plano semántico, la utilización del término se refiere a la *previsibilidad de aparición de determinados contenidos semánticos* (con independencia de la forma que revistan).

En este caso, evidentemente, la cuantificación exacta no es posible puesto que no se pueden hacer análisis estadísticos que permitan calcular su probabilidad matemática de aparición.

Sin embargo, podemos valorar la *probabilidad inductiva* de que ese contenido y no otro sea emitido en una situación o contexto dado.

> Así, si leemos en la prensa que el Papa condena el aborto , que ha habido enfrentamientos en el Líbano o que se han producido incendios forestales, estas noticias son menos "informativas" que si leyéramos que ha aterrizado un OVNI en Madrid o que el Gobierno recomienda el fraude fiscal. En efecto, estos contenidos son mucho menos *previsibles*.

De ahí que asociemos el grado de información de un mensaje a la medida en que es inesperado y a la consiguiente sorpresa que nos produce. En este sentido, el término se halla algo más próximo al significado usual, intuitivo, que reviste en la lengua común.

La información relativa a contenidos afecta asimismo a segmentos de mensajes con relación al contexto:

> Si alguien nos dice que "no es supersticioso"... "*porque la superstición da mala suerte*", la segunda parte de su enunciado es más informativa (menos previsible) que la continuación esperada, a saber, algún argumento de tipo racional contra la superstición.
>
> Si digo que "mi hijo es un alumno brillante"... "*que ha suspendido siete asignaturas*" esta última frase tiene más información que cualquier continuación que versara sobre buenas notas.
>
> Las expresiones irónicas o humorísticas contienen más "información" que los enunciados "en serio" correspondientes.

Obsérvese que, en sentido técnico, el grado de "información" contenido en un mensaje lingüístico no tiene nada que ver con su "veracidad".

Se ha visto INFORMACIÓN como "INFORMACIÓN-PAPA"

3.2. La medida de la información

La información es una magnitud medible en todos los fenómenos susceptibles de ser tratados mediante el cálculo de probabilidades.

En el caso más simple, la fuente tiene dos opciones y nada la "inclina" más a una que a otra. Esta situación define la unidad de Información más comúnmente utilizada y que se denomina **bit** (abreviatura de binary digit).

Un bit es, por tanto, la cantidad de información contenida en un mensaje o suceso cuya *probabilidad de aparición es 1/2*. Es decir, cuando sólo son posibles *dos mensajes* u opciones y ambos tienen igual probabilidad de ser seleccionados por la fuente.

> Ejemplos de mensajes de 1 bit de información: "cara" (o "cruz") cuando lanzamos una moneda al aire; "niño" o "niña" en el nacimiento de un bebé; "izquierda" o "derecha" en una bifurcación desconocida. Si en un relé on-off, "encendido" correspondiera al texto completo del Quijote y "apagado" a la palabra "si", ambos mensajes tendrían asimismo 1 bit.

Como vemos, cada mecanismo de dos estados, ("relé on/off", encendido/apagado, magnetizado/no magnetizado, abierto/cerrado, sí/no, 0/1) transmite (o almacena) un bit de información.

No dice el apartado nada [illegible]

Los mensajes de la fuente F1 (Vid. supra. 3.11) contenían 1 bit; los de F2, 2 bits; los de F3, 3 bits.

1 bit supone 2 mensajes posibles: 0 1					
2 bits suponen 4:	00	01	10	11	
3 bits suponen 8:	000	001	010	011	
	100	101	110	111	
4 bits suponen 16:	0000	0001	0011	0111	1111
		0010	0110	1110	
		0100	1100	1011	
		1000	0101	1101	
			1010		
			1001		

Cada vez que añadamos 1 bit se duplicará el número de mensajes u opciones. Es decir, estamos ante una progresión geométrica. $2^I = N$, siendo I la cantidad de información y N los mensajes posibles.

Ocho relés, es decir 8 bits almacenan o transmiten 256 combinaciones distintas posibles. Es la unidad de medida de la memoria de los ordenadores más sencillos y se denomina comercialmente byte (con y griega). Sus múltiplos son el Kilobyte (K) y el Megabyte (MEGA).

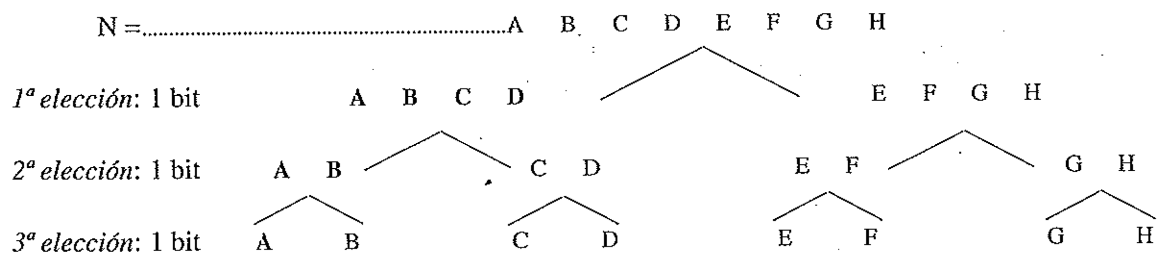
Si lo que conocemos es el número de mensajes posibles de la fuente, es decir N, para obtener la Información hacemos el razonamiento inverso:

Cualquier número de opciones equiprobables puede descomponerse en una sucesión de elecciones binarias. La fijación de cada una de ellas supone por definición 1 bit. Un mensaje tendrá por tanto tantos bits como elecciones binarias han sido necesarias para individualizarlo.

Ejemplo:

Número de opciones posibles: N = 8

Probabilidad de aparición de cada una de ellas P = 1/8.



Total: 3 bits

El suceso A contiene tres bits de información (y lo mismo los sucesos B, C, D, etc... para los que podemos repetir la operación).

Vemos de nuevo que el número de opciones N no es sino 2^3 (2.2.2 = 8). La información I en bits es la potencia a la que hay que elevar 2 para hallar N. Dicho de otro modo: el logaritmo en base 2 de N. Es decir, dado que N = 1/P,

$$\boxed{I = \log_2 1/P}$$

fórmula que permite calcular la cantidad de información en bits de cualquier mensaje.

Imagínense por ejemplo cuatro sucesos A, B, C, D cuyas respectivas probabilidades de aparición fueran:

$P_A = 1/2$ (= 0,5) La información correspondiente es: $I_A = \log_2 1 - (\log_2 0{,}5) = 1$ bit[1]

$P_B = 1/4$ (= 0,25) .... $I_B = \log_2 1 - (\log_2 0{,}25) = 2$ bits

$P_C = 1/8$ (= 0,125) .... $I_C = \log_2 1 - (\log_2 0{,}125) = 3$ bits

$P_D = 3/4$ (= 0,75) .... $I_D = \log_2 1 - (\log_2 0{,}75) = 0{,}5$ bits

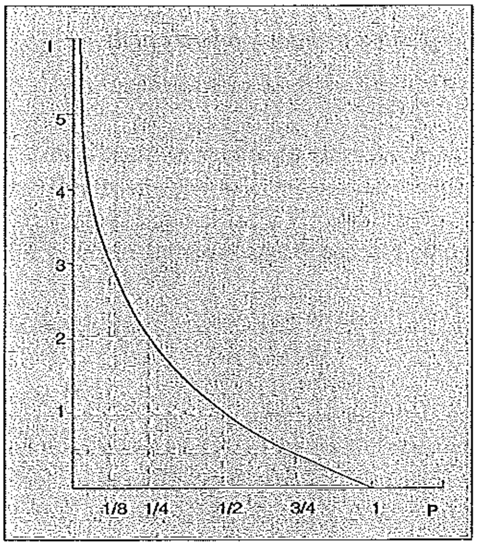


fig. 8. Función logarítmica que establece la relación entre la probabilidad P y la información I.

3.3. La redundancia

La **redundancia** es la parte del mensaje que no contiene información: cuanto más previsible es un mensaje (o un segmento de un mensaje) mayor redundancia contiene. Un mensaje altamente redundante no contiene prácticamente información.

La cantidad de redundancia que posee un mensaje es *complementaria* a su cantidad de información. Si un mensaje contiene un 80% de información, su redundancia será un 20 %.

Tipos de redundancia

a) Redundancia inherente al *código*.

La parte del mensaje que no está determinada por la libre elección del emisor sino por las reglas que rigen las posibilidades de combinación de los símbolos del código constituye la redundancia inherente al código.

Cuanto más redundante es un mensaje o un segmento de un mensaje, menos información contiene y, por tanto, más fácil resulta recuperarlo si se pierde o se suprime. Si es totalmente redundante, sólo cabe una opción y la información que contiene es cero.

Al construir un código artificial, es habitual prever cierta proporción de redundancia con el fin de paliar los efectos del ruido (Vid. supra. 1.3).

Las lenguas naturales constituyen códigos con una alta proporción de redundancia.

1 La probabilidad es un número siempre menor que 1, por lo que su log es negativo. La probabilidad máxima posible sería P=1, que supone I = 0. Recuérdese que el log de 1 es 0. Normalmente se manejan logaritmos decimales y se opera la conversión a la base 2. Para ello basta multiplicar por 3,32 ($\log_{10}$ de 2).

Ejemplos de segmentos de mensaje de redundancia máxima (en negrita):

consonantes	*trapo* *alianza* *árbol*
sílabas	*piña tropical* *independiente*
morfemas gramaticales	*niño bueno* *estos coches*
	Me acuerdo de lo que dijiste
unidades léxicas	*Lo consiguió a duras **penas**.*
	*Hecho a trancas y a **barrancas**.*
sintagmas	*Cría cuervos y te **sacarán los ojos**.*

La redundancia de las lenguas naturales hace posible la existencia de los crucigramas. De no existir redundancia, éstos serían imposibles: imagínese un crucigrama de números de teléfono...

La redundancia del inglés es de un 50%. Es decir, la mitad de las letras o de las palabras del que escribe en inglés son de libre elección y el resto está condicionado por la estructura estadística de dicha lengua. La redundancia del castellano es muy superior.

Compárese en efecto:

Las chicas guapas *Los chicos guapos*
The beautiful girls *The beautiful boys*

(en castellano el adjetivo y en el artículo repiten el género y el número del sustantivo).

El llamado "estilo telegráfico" que utilizamos para economizar en el texto de los telegramas supone eliminar parte de la redundancia perteneciente a las reglas gramaticales del código y dejar sólo las palabras informativas imprescindibles: "*Remesa patatas averiada. Exijo devolución dinero. López.*"

b) Redundancia *libremente introducida.*

El emisor humano puede introducir voluntariamente redundancia en el momento de codificar el mensaje para asegurar su perfecta recepción.

En su forma más simple supone repetir el mensaje o parte del mismo bajo otra forma –o con la misma–.

"*!Ven aquí! !Que vengas te digo! ¿Quieres venir de una vez?*"...

Son ejemplos de redundancia libremente introducida por el emisor:

- repetir un número telefónico agrupando los dígitos de modo diferente: 384721 (tres-ocho-cuatro-siete-dos-uno; treinta y ocho-cuarenta y siete-veintiuno).
- escribir en un talón bancario la cantidad en cifras y en letras.
- el subrayado, los titulares o cualquier recurso tipográfico que resalte una parte de un mensaje escrito .
- un dibujo explicativo de una descripción.
- un gráfico que visualiza unos datos.
- un gesto que significa lo mismo que las palabras a las que acompaña.
- decir lo mismo con distintas palabras.
- elevar la voz.

La introducción de redundancia supone utilizar un gasto de energía mayor del estrictamente necesario para transmitir la información. Su utilización tiene por objeto contrarrestar los efectos del ruido. Constituye, como se dijo en 1.3., un margen de seguridad que asegura que el mensaje llegue sin pérdida de información.

En otros casos, la redundancia puede suponer la ausencia de contenidos informativos que transmitir: un texto que repite idéntica idea bajo distintas formas (por ejemplo un examen con abundante "paja" para llenar papel) es más redundante que otro repleto de contenidos originales.

La llamada *función fática* del lenguaje se caracteriza por producir enunciados con bajísima información semántica, es decir, con altísima redundancia (Vid. infra. 7.2. y T. 16).