

¿Un nuevo camino para la digitalización de la palabra hablada?

Sistema de criptofonía de baja velocidad

JOSÉ I. NORMAND BERGAMÍN
Coronel de Aviación
Criptólogo y Criptoanalista

CON la implantación del teletipo en las comunicaciones, se generalizó la codificación digitalizada de las 26 letras del abecedario con diferentes valores binarios, a las que se añaden 6 funciones de teletipo (espacio, paso de línea, retroceso de carro, etc.) para completar las $2^5 = 32$ com-

binaciones del código CITT n.º 2.

En los primeros años ochenta se logra la transmisión por líneas telefónicas de fotografías, cartas o dibujos en color, entre ordenadores, alcanzándose la digitalización de hasta 256 colores diferentes.

Hasta ahora, para tener un

Sistema de Digitalización de la palabra hablada que consiga en la recepción una buena inteligibilidad y el reconocimiento del interlocutor, se precisa codificar digitalmente a una velocidad de 16.000 bits por segundo y como mínimo a 9.600 bits/seg. con muy buenas líneas de transmisión.

CUADRO 1

LISTADO DE FICHERO EQUIVALE.TXT

AH	A	LLE	YE	HOL	OL
HA	A	XE	SE	HOM	OM
HAL	AL	JEL	GEL	HON	ON
HALL	AL	JEN	GEN	HOR	OR
HAM	AM	JER	GER	HOS	OS
HAN	AN	JES	GES	HOZ	OZ
HAR	AR	KER	QUER	KO	CO
HAS	AS	XEM	SEM	LLO	YO
HAZ	AZ	XEN	SEN	XO	SO
KA	CA	LLEN	YEN	KLO	CLO
XA	SA	LLE	YER	XON	SON
LLA	YA	LLES	YES	LLOR	YOR
XAN	SAN	ZER	CER	LLOS	YOS
XAS	SAS	HI	I	HU	U
LLAN	YAN	Y	I	HUN	UN
LLAR	YAR	HIL	IL	HUR	UR
HE	E	HIM	IM	HUS	US
EH	E	HIN	IN	KU	CU
HEC	EC	HIR	IR	LLU	YU
HEL	EL	HIS	IS	XU	SU
HEM	EM	ZI	CI	KUR	CUR
HEN	EN	JI	GI	PULL	PUL
HEP	EP	KI	QUI	XUL	SUL
HER	ER	XI	SI		
HERT	ERT	JIL	GIL		
HERTZ	ERTZ	JIR	GIR		
HEZ	EZ	JIS	GIS		
ZE	CE	XIS	SIS		
JE	GE	OH	O		
KE	QUE	HO	O		

CUADRO 2

INFORMACIONES RELACIONADAS

- La FONÉTICA estudia los sonidos, estableciendo el repertorio, con arreglo a particularidades y diferencias.
- La FONOLOGÍA estudia los sonidos, organizándolos en Sistema.

FONEMA:

- Sonido simple del lenguaje hablado.
- Unidad fonológica mínima, que en una lengua hablada puede oponerse a otras con contraste significativo.
- Se caracteriza por su función significante, por su capacidad para diferenciar significados.

SONIDO:

- Sensación acústica, producida en el órgano del oído, por el movimiento vibratorio de los cuerpos.
- Los sonidos lingüísticos se dividen en: vocálicos, aquellos que por sí solos constituyen una sílaba y consonánticos, los que carecen de esa propiedad. De estos existen diferentes subdivisiones.

LETRA:

- Cada uno de los signos con que se representan en la escritura los sonidos del habla.

SILABA:

- Grupo fónico elemental, compuesto de uno o varios sonidos que se pronuncian en una sola emisión de voz al hablar.
- La noción de sílaba es más intuitiva que científica.
- Acústicamente, las sílabas se encuentran limitadas por depresiones de la perceptibilidad.

- En español en el hablar normal, dos grupos fónicos de un mismo número de sílabas, duran aproximadamente el mismo tiempo. (En inglés, dos grupos fónicos de diferente número de sílabas, duran a menudo el mismo tiempo.)

PALABRA:

- Conjunto de sonidos o letras, que constituyen una unidad reconocible por los hablantes de una lengua, de una idea.
- Es la mínima secuencia de segmentos dotada de significado y susceptible de ser aislada por pausas.

ELOCUCION:

- Modo de distribuir los conceptos y las palabras a lo largo de un discurso.
- Forma de usar la facultad del lenguaje para expresar las ideas.

De todas formas, el objetivo que ahora se persigue es lograr la digitalización, trabajando a una velocidad entre 300 a 400 bits/seg. Se exige que haya seguridad de entendimiento total del mensaje transmitido, en la recepción, pero no reconocer a la persona que nos habla.

Se ha partido de obtener el total de las sílabas que forman

las palabras del diccionario de la lengua. Realizadas las estadísticas necesarias, obtenemos el dato de tener menos de $1024 = 2^{10}$ sílabas diferentes, para digitalizar la totalidad de las palabras.

Se establece la aceptación de existir igualdad de sonidos, aunque aparezcan escritos con letras diferentes (cuadro 1). La

mayoría tienen igual sonido (KA = CA, ZE = CE, JIL = GIL, etc.), pero existen otros en que debemos aceptar su igualdad (LLE = YE, XI = SI). No es idéntico sonido "lle" de llegar, que el "ye" de yegua. No es igual el "xi" de xilófono, que el "si" de sifón. Pero lo admitimos, pues no va a impedir el entendimiento de un mensaje. (El sonido de la "x",

CUADRO 3

LISTADO DEL FICHERO *dicciona.doc*

e-rra-la	es-ca-que-ar-se	es-co-llo
e-rre	es-ca-ra-ba-jo	es-com-bro
e-rror	es-ca-ra-mu-za	es-con-der
e-ruc-to	es-ca-ra-pe-la	es-con-di-do
e-ru-dí-to	es-car-bar	es-con-dri-jo
e-rup-ci-ón	es-car-ce-o	es-co-ñar
es-bel-tez	es-car-cha	es-co-pe-la
es-bel-to	es-car-dar	es-co-plo
es-bi-rrro	es-ca-ri-a-dor	es-co-ra
es-bo-zo	es-car-la-la	es-cor-bu-to
es-ca-be-char	es-car-la-li-na	es-co-ri-a
es-ca-be-che	es-car-men-tar	es-cór-pi-ón
es-ca-bro-so	es-car-mi-en-to	es-co-la-do
es-ca-bu-llir-se	es-car-ni-o	es-co-le
es-ca-cha-rrar	es-ca-ro-la	es-co-ti-lla
es-ca-fan-dra	es-car-pa-do	es-co-zor
es-ca-fo-i-des	es-car-pa-du-ra	es-crí-bir
es-ca-la	es-car-pi-a	es-crí-to
es-ca-la-da	es-car-pin	es-crí-tor
es-ca-la-fón	es-ca-sa-men-te	es-crí-tu-rar
es-ca-lar	es-ca-se-ar	es-crú-pu-lo
es-cal-da-do	es-ca-sez	es-cru-pu-lo-so
es-cal-dar	es-ca-li-mar	es-cru-je
es-ca-le-no	es-ca-yo-la	es-cru-li-ni-o
es-ca-le-ra	es-ce-na	es-cu-a-dra
es-cal-fa-do	es-cep-ti-cis-mo	es-cu-a-dri-lla
es-cal-far	es-cép-ti-co	es-cu-a-drón
es-ca-li-na-la	es-cin-dir	es-cu-á-li-do
es-ca-lo-fri-an-te	es-ci-si-ón	es-cu-a-lo
es-ca-lo-fri-o	es-cla-re-cer	es-cu-cha
es-ca-lo-nar	es-cla-re-ci-do	es-cu-char
es-ca-lo-pe	es-cla-vi-na	es-cu-dar
es-cal-pe-lo	es-cla-vi-tud	es-cu-de-ro
es-ca-ma	es-cla-vi-zar	es-cu-di-lla
es-ca-mar	es-cla-vo	es-cu-do
es-ca-mo-le-ar	es-cle-ro-sis	es-cu-dri-ñar
es-cam-par	es-cle-ró-li-ca	es-cu-e-la
es-ca-ci-a-dor	es-clu-sa	es-cu-e-to
es-can-ci-ar	es-co-ba	es-cul-pir
es-can-da-li-zar	es-co-bi-lla	es-cul-tor
es-cán-da-lo	es-co-ce-du-ro	es-cul-tu-ra
es-can-da-llo	es-co-cer	es-cu-pir
es-can-dí-na-vo	es-co-cés	es-cu-rre-pla-tos
es-cá-ner	es-co-fi-na	es-cu-rrí-di-zo
es-ca-ño	es-co-ger	es-cu-rrir
es-ca-pa-da	es-co-lar	es-drú-ju-lo
es-ca-par	es-co-la-ri-dad	e-se
es-ca-pa-ra-le	es-co-li-o	e-sen-ci-a
es-ca-pe	es-col-la	e-sen-ci-al
es-ca-pu-la-ri-o	es-col-lar	

CUADRO 4

LISTADO DEL FICHERO *DICCFSDA.TXT*

A	3616	BLÉN	1	CALF	1
AB	17	BLES	1	CAM	26
ABS	8	BLEZ	1	CAN	97
AC	38	BLI	20	CAP	8
AD	42	BLIN	3	CAR	330
ADS	2	BLO	9	CAS	38
AF	1	BLOC	1	CAZ	6
AG	1	BLOR	1	CE	395
AL	312	BLU	3	CED	1
AM	41	BLUFF	1	CEL	13
AN	192	BO	522	CEN	125
AP	6	BOO	1	CEP	27
AR	701	BOJ	1	CEPS	3
AS	121	BOL	34	CER	152
AT	11	BOM	20	CES	15
AZ	4	BON	15	CET	1
BA	638	BOR	46	CI	1534
BAC	1	BOS	11	CID	1
BAD	1	BOT	1	CIL	5
BAL	49	BOZ	3	CIM	3
BALS	1	BRA	75	CIN	25
BAM	6	BRAL	3	CIR	43
BAN	62	BRAN	10	CIS	25
BANG	1	BRAR	32	CLA	64
BAP	1	BRE	111	CLAC	1
BAR	143	BREL	1	CLAN	2
BAS	29	BREN	1	CLAR	1
BAZ	1	BRI	54	CLE	11
BE	379	BRID	1	CLEC	1
BEC	4	BRIL	3	CLEN	1
BEL	15	BRIN	4	CLI	11
BEN	95	BRIR	4	CLIP	2
BER	166	BRIS	1	CLIS	3
BES	36	BRIZ	2	CLO	8
BEZ	3	BRO	44	CLON	2
BI	657	BRON	10	CLOS	1
BIC	4	BRU	13	CLU	19
BID	2	BRUP	2	CO	1047
BIF	1	BRUS	3	COC	4
BIL	11	BU	152	COG	2
BIN	9	BUL	21	COL	21
BIR	30	BUM	1	COM	175
BIS	33	BUN	9	CON	660
BIT	2	BUS	26	CONS	23
BIZ	10	BUZ	3	COP	1
BLA	22	CA	1078	COR	97
BLAR	9	CAC	1	CORD	1
BLAS	2	CAD	1	COS	24
BLAN	14	CAJ	1	COT	2
BLE	330	CAL	76	COZ	1

cuando no es primera sílaba de una palabra, se mantiene. Ejemplo: léxico.).

En el cuadro 2 damos unas cuantas informaciones, a fin de centrarnos en el entorno de unos mismos conceptos.

Vistas las dificultades que se presentan, se han realizado diferentes trabajos para aportar el mayor número de datos que

puedan ser utilizados. Son los siguientes:

— Determinación de las palabras del diccionario de la lengua, escritas en forma silábica. (Cuadro 3.)

— La estadística realizada, da 943 diferentes sílabas aparecidas, con un total de 60.040.

(El cuadro 4 muestra la 1.ª hoja en la que aparecen ordenadas

alfabéticamente, con indicación del número total de cada sílaba.)

(El cuadro 5, muestra la 1.ª hoja de la estadística, ordenada por frecuencia de aparición de cada sílaba.)

— En el cuadro 6, tenemos ordenadas alfabéticamente, las sílabas consecutivas a cada una. Así de las 3.616 "A" aparecidas,

CUADRO 5

LISTADO DEL FICHERO DICCSF.TXT

A	3616	SE	371	SAR	132
CI	1534	ME	365	JO	129
O	1301	NE	357	TOR	129
TA	1212	PO	345	LU	128
TE	1165	BLE	330	RRE	126
TO	1112	CAR	330	CEN	125
CA	1078	TU	328	LLO	125
CO	1047	FI	313	TRE	125
TI	999	AL	312	RRO	123
DO	990	GE	312	AS	121
RA	883	NAR	275	IM	120
RI	797	ZA	273	TEN	118
ON	778	TRA	271	DAR	114
E	731	GO	252	PAR	114
RE	714	DAD	237	RRA	114
AR	701	LLA	237	ER	112
NA	692	GI	231	FO	112
CON	660	ZAR	230	BRE	111
BI	657	DOR	226	CHA	108
MA	645	MEN	221	NU	106
LA	642	QUI	221	CAN	97
BA	638	QUE	218	COR	97
EN	633	GU	217	BEN	95
RO	611	LAR	212	ÑO	93
SI	611	SU	206	YE	93
LI	610	TRO	202	DIS	91
NO	606	AN	192	OR	87
PA	585	RAR	185	PLA	86
MI	577	PRO	182	ÑA	83
SO	562	PRE	177	TAN	83
IN	559	COM	175	YAR	83
DI	525	BER	166	MAN	82
BO	522	FA	166	RRI	82
ES	520	ZO	165	DEN	80
MO	498	FE	163	CHE	77
NI	484	DU	159	JU	77
PE	463	GRA	157	CAL	76
I	431	JA	153	BRA	75
U	430	BU	152	FU	73
DA	429	CER	152	SIS	73
CU	409	EM	152	LAN	71
TAR	407	GAR	149	NAL	71
SA	398	TRI	149	POR	70
CE	395	PU	146	JAR	68
LO	395	BAR	143	RU	67
BE	379	EX	141	CHO	66
DE	378	MAR	139	CLA	64
PI	378	PER	139	FOR	63
GA	376	MU	138	BAN	62
LE	373	TER	134	LEN	62

CUADRO 5.1.

LISTADO DEL FICHERO TORCSF.TXT

UL	17	YEN	11	MON	6
ZOS	17	ZAN	11	RIS	6
CAL	16	AM	10	RRER	6
NOM	16	BEL	10	RRU	6
ZAS	16	CLU	10	SIM	6
AB	15	CUR	10	SIR	6
BLO	15	FAL	10	SUN	6
CIL	15	PAN	10	TEX	6
ÑOR	15	PES	10	TUR	6
TUS	15	PIS	10	ABS	5
YOR	15	PLO	10	CHAN	5
CHOS	14	RRUM	10	FRO	5
DRA	14	SOM	10	GLO	5
FEC	14	EZ	9	GOL	5
FIR	14	GLAR	9	GUES	5
GRO	14	GRU	9	LIS	5
LIR	14	GUIR	9	LIZ	5
RIR	14	PAS	9	LLON	5
TAD	14	TIM	9	MES	5
DROS	13	TRU	9	MUS	5
FEN	13	ZU	9	PRAR	5
INS	13	BRIR	8	QUES	5
LUZ	13	BRU	8	RAC	5
ÑE	13	CUNS	8	YAN	5
PUL	13	DRO	8	YES	5
ROM	13	FLE	8	AZ	4
COL	12	IG	8	BLAS	4
LON	12	ÑI	8	CEL	4
MIR	12	PEL	8	DIC	4
PEC	12	RRAN	8	FER	4
PLAN	12	SIG	8	FES	4
PRU	12	TRAC	8	FLA	4
SUB	12	BIL	7	FRAS	4
TUM	12	BRUS	7	GIS	4
CLI	11	CONS	7	GLA	4
DIG	11	CUM	7	JAN	4
DIN	11	FLO	7	LAM	4
DUC	11	LIM	7	LIN	4
FRA	11	LOJ	7	LUN	4
GOS	11	MIL	7	MEZ	4
NAN	11	NIR	7	MIN	4
NIS	11	RRAR	7	MUL	4
ÑAR	11	RROR	7	NEN	4
NAS	11	TRIS	7	PLAS	4
REC	11	YER	7	RER	4
RRAS	11	BAL	6	RIT	4
SOL	11	CRIS	6	RRES	4
TIL	11	FAC	6	TIC	4
TRAN	11	FOM	6	UR	4
TRANS	11	GOR	6	YAR	4

vemos que le pueden seguir 320 sílabas diferentes, más la 1.^a que aparece seguida de blanco, que indica que 836 veces la A aparece como final de palabra.

— El cuadro 7, recoge la hoja en que tenemos agrupadas todas las sílabas que utilizan la vocal "A", ordenadas por número de letras y frecuencias.

El resumen de los 5 grupos de las 5 vocales, lo tenemos en el cuadro 8.

Todo esto nos proporciona una verdad: 943 es el número de sílabas diferentes que pueden ser utilizadas para formar las palabras, pero normalmente para escribir y hablar no se utilizan todas las palabras y por ello, queriendo llegar a una ma-

yor realidad, se han repetido las estadísticas partiendo de una novela de un escritor que forma parte de la Real Academia de la Lengua Castellana. (La nueva estadística aparece en los cuadros 5.1, 7.1 y 8.1).

A la vista de los cuadros 8 y 8.1, apreciamos lo siguiente:

— El total de sílabas diferen-

CUADRO 6

LISTADO DEL FICHERO DICCFDOA.TXT

A	836	A COL	1	A FAR	1
A BA	41	A COM	3	A FE	7
A BAD	1	A CON	7	A FEC	7
A BAL	3	A COR	6	A FI	15
A BAN	13	A COS	2	A FIN	2
A BAR	4	A CRA	1	A FIR	3
A BAS	3	A CRE	4	A FLI	1
A BE	32	A CRI	4	A FLIC	1
A BEN	11	A CRO	4	A FLO	2
A BER	8	A CU	27	A FLU	3
A BES	2	A CUS	3	A FO	5
A BI	63	A CHA	7	A FOR	1
A BIL	1	A CHAN	1	A FRAN	1
A BIS	8	A CHE	1	A FREN	2
A BLA	6	A CHI	4	A FRI	2
A BLAR	1	A CHIS	1	A FRO	2
A BLAN	3	A CHO	1	A FRON	1
A BLE	24	A CHU	2	A FU	1
A BLU	1	A DA	16	A FUS	1
A BO	28	A DAN	1	A GA	12
A BOL	1	A DAP	5	A GAR	1
A BOM	2	A DAR	2	A GE	15
A BOR	6	A DE	22	A GEN	5
A BRA	5	A DEL	2	A GI	10
A BRE	5	A DEN	2	A GIL	1
A BRI	4	A DEP	1	A GLO	3
A BRIL	1	A DER	4	A GLU	3
A BRIR	1	A DI	14	A GO	14
A BRO	2	A DIC	1	A GOL	1
A BRON	1	A DIR	2	A GOS	1
A BRU	1	A DO	61	A GRA	13
A BRUP	2	A DOC	1	A GRAN	1
A BU	12	A DON	3	A GRE	6
A BUL	2	A DOP	3	A GRES	1
A BUN	3	A DOR	22	A GRI	10
A CA	24	A DRA	4	A GRO	3
A CAL	1	A DRAN	1	A GRU	3
A CAM	1	A DRAR	1	A GU	29
A CAN	3	A DRE	2	A GUE	3
A CAR	1	A DRI	7	A GUEZ	1
A CE	22	A DRO	2	A GUI	4
A CEL	1	A DRON	1	A I	12
A CEN	5	A DRU	1	A IN	1
A CEP	5	A DU	7	A IS	4
A CER	7	A DUC	2	A JA	2
A CI	45	A DUL	4	A JAN	1
A CIL	1	A DUS	1	A JAR	2
A CLA	3	A E	25	A JO	3
A CLI	1	A FA	6	A JU	1
A CO	26	A FAN	1	A JUS	7

CUADRO 6.1.

LISTADO DEL FICHERO TORCFDOA.TXT

A	1961	A DOR	2	A MU	2
A BA	36	A DOS	7	A NA	4
A BAN	25	A DRA	4	A NAS	1
A BE	5	A DRI	1	A NE	1
A BEN	2	A DRO	5	A NI	12
A BER	27	A DROS	13	A NO	1
A BES	1	A DU	1	A NOS	1
A BI	253	A FA	1	A NU	1
A BIS	4	A FAN	5	A NUN	6
A BLA	27	A FE	1	A ÑA	17
A BLAR	26	A FEC	4	A ÑI	1
A BLAS	2	A FI	3	A ÑO	3
A BLAN	8	A FIR	6	A ÑOS	40
A BLE	14	A FOR	2	A O	77
A BLES	4	A FRI	10	A ON	2
A BLO	6	A GA	4	A PA	16
A BO	5	A GAS	1	A PAR	7
A BOR	1	A GE	12	A PE	10
A BRA	12	A GEN	6	A PER	1
A BRE	2	A GI	1	A PES	2
A BREN	1	A GO	7	A PI	9
A BRI	34	A GRA	3	A PLA	8
A BRIR	5	A GRE	1	A PLAS	4
A BRU	2	A GU	18	A PLO	1
A BU	46	A I	45	A PO	14
A BUN	1	A IS	5	A POR	1
A CA	37	A JA	1	A POS	2
A CE	16	A JAR	2	A PRE	12
A CEN	4	A JUS	1	A PREN	6
A CEP	7	A LA	5	A PRO	7
A CER	91	A LAR	5	A PU	1
A CES	4	A LE	13	A PUN	4
A CI	109	A LEN	2	A QUE	30
A CO	3	A LES	10	A QUEL	22
A COM	9	A LI	25	A QUI	30
A CON	2	A LIS	1	A RA	8
A COR	2	A LO	11	A RAN	1
A COS	5	A LU	4	A RE	5
A CRI	1	A LLA	10	A REN	1
A CU	19	A LLI	22	A RI	12
A DA	21	A MA	9	A RON	4
A DAN	1	A MAN	1	A RRA	2
A DAP	1	A MAR	3	A RRAN	4
A DAS	3	A ME	6	A RRAS	4
A DE	25	A MEN	3	A RRE	26
A DEN	2	A MI	4	A RRI	7
A DER	2	A MIS	5	A RRO	5
A DI	8	A MO	1	A RROZ	1
A DO	40	A MOR	5	A RRU	4
A DON	2	A MOS	7	A SAL	1

tes utilizadas (587) en la novela, es un 37,64% menor.

Es evidente que tendrán que codificarse las 943 sílabas diferentes de las palabras del diccionario, pero normalmente la tercera parte de ellas van a aparecer raras veces. Por ello, en la codificación, las sílabas de mayor frecuencia tendrán menor número de bits.

— La distribución de los totales agrupados por sonidos vocales, varía sensiblemente y es más real la estadística correspondiente a la novela, donde la frecuencia de las letras E y A es muy próxima y las más elevadas.

— Comprobamos que las sílabas formadas por 2 letras o fonemas, representan el 60,4%

del total, que sumadas a los sonidos vocales (1 letra) con el 10,84% y 13,46% en cada caso, nos dan como totales 71,24% y 73,86%. Ello nos indica que solamente un 26% de las sílabas que van a ser utilizadas, estarán compuestas por más de dos fonemas.

— Los espacios entre palabras o silencios, se han contabi-

CUADRO 7

LISTADO DEL FICHERO DE DICCIONARIO

A	3616	TAN	83	TAD	11	JAZ	1	FRAM	2
TA	1212	YAR	83	YAN	10	LAG	1	FRAS	2
CA	1078	MAN	82	ZAL	10	LAX	1	PRAR	2
RA	883	CAL	76	FAS	9	LAZ	1	TLAN	2
AR	701	BRA	75	ABS	8	MAD	1	TLAS	2
NA	692	LAN	71	CAP	8	NAF	1	TRAZ	2
MA	645	NAL	71	LAC	8	NAM	1	BLAS	1
LA	642	JAR	68	DAS	7	NAZ	1	BANG	1
BA	638	CLA	64	ZAM	7	NAC	1	CALF	1
PA	585	BAN	62	BAM	6	ÑAS	1	CLAC	1
DA	429	RAN	54	CAZ	6	SHA	1	CLAR	1
SA	398	MAL	51	SAS	6	TAG	1	CRAC	1
GA	376	SAL	50	ZAS	6	ZAG	1	CRAN	1
AL	312	BAL	49	DAP	5	CHAR	41	CHAM	1
YA	273	TAL	49	GAM	5	TRAR	35	CHAS	1
ZA	273	RAL	48	JAL	5	BRAR	32	DRAC	1
AN	192	DRA	39	PAC	5	TRAN	31	DRAL	1
FA	166	GAN	39	SAM	5	TRAS	30	DRAN	1
JA	153	CAS	38	TAC	5	RRAR	28	FLAC	1
AS	121	FRA	38	GAZ	4	PLAN	22	FRAZ	1
ÑA	83	SAN	38	JAS	4	FRAN	18	GLAN	1
AD	42	ÑAR	35	ÑAL	4	PLAS	16	GRAL	1
AM	41	PAN	32	PAM	4	TRAC	15	GRAS	1
AC	38	BAS	29	PAZ	4	TRAL	15	GRAZ	1
AB	17	CAM	26	RAC	4	BLAN	14	LLAZ	1
AT	11	NAM	25	RAM	4	RRAN	12	MACH	1
AP	6	RAS	25	DAC	3	RRAS	12	PRAG	1
AZ	4	FLA	24	DAZ	3	GRAN	11	RANG	1
WA	3	FAR	23	JAN	3	BRAN	10	TAND	1
AF	1	PAL	23	MAM	3	BLAR	9	TRANS	32
AG	1	BLA	22	ÑAN	3	DRAR	9	FLASH	1
TAR	407	FAC	22	RAP	3	GRAR	9	FRANC	1
CAR	330	FAN	22	RAZ	3	LLAS	8	GANGS	1
NAR	275	MAS	22	XAL	3	FRAC	7	PLANC	1
TRA	271	FAL	21	ADS	2	RRAL	7		
DAD	237	DAN	18	FAZ	2	GLAR	6		
ZAR	230	TAS	18	JAC	2	PLAR	6		
LAR	212	CRA	17	JAM	2	TRAM	6		
RAR	185	MAG	17	LAP	2	CRAR	4		
GAR	157	GAL	16	MAZ	2	FLAN	4		
GAR	149	PAS	16	RAX	2	BRAL	3		
BAR	143	TAM	16	SAZ	2	CHAN	3		
MAR	139	NAS	15	TAZ	2	DRAS	3		
SAR	132	ZAN	15	BAC	1	FRAR	3		
DAR	114	GLA	14	BAO	1	PRAC	3		
PAR	114	DAL	13	BAP	1	BLAS	2		
RRA	114	PRA	13	BAZ	1	CLAN	2		
CHA	108	LAM	12	CAC	1	CHAL	2		
CAN	97	LAS	12	CAD	1	FLAR	2		
PLA	86	GAS	11	CAJ	1	FRAG	2		

CUADRO 7.1.

LISTADO DEL FICHERO -NOVELA-

A	3915	SAR	53	JAN	4	GRAR	1
LA	1545	DAR	52	LAM	4	GRAS	1
RA	1198	GRA	51	YAR	4	TRAL	1
BA	975	MAN	51	CRA	3	TRANS	11
TA	920	MAR	50	RAM	3		
CA	751	CAM	48	XAL	3		
NA	746	GAR	48	FAS	2		
PA	694	SAN	47	ÉAC	2		
SA	636	TAL	47	ÉAN	2		
DA	621	CLA	45	ZAL	2		
MA	549	BAR	40	CAP	1		
AL	483	DAN	40	CAZ	1		
AN	423	LAN	39	DAP	1		
AS	225	BAS	38	GAL	1		
GA	195	BLA	38	GAZ	1		
ZA	185	PLA	37	JAM	1		
AR	163	TAM	35	JAZ	1		
JA	125	GAS	34	MAG	1		
YA	123	SAL	33	NAZ	1		
ÉA	90	JAR	32	RAP	1		
FA	63	ZAR	32	TAC	1		
AD	35	RAL	31	TRAS	67		
AC	25	CAS	29	LLAS	38		
AB	15	FAN	27	GRAN	37		
AM	10	PAL	27	CHAR	29		
AZ	4	PAZ	27	BLAR	27		
AP	2	GAN	24	CHAS	25		
LAS	394	MAL	24	TRAR	23		
MAS	247	NAL	21	BLAN	21		
LLA	233	JAS	20	PLAN	12		
TRA	221	PRA	17	RRAS	11		
BAN	184	CAL	16	TRAN	11		
TAR	183	ZAS	16	GLAR	9		
DAD	160	DRA	14	RRAN	8		
TAN	151	TAD	14	TRAC	8		
DAS	127	FRA	11	RRAR	7		
TAS	125	NAN	11	CHAN	5		
RAN	119	ÉAR	11	PRAR	5		
BRA	92	ÉAS	11	BLAS	4		
RAS	90	ZAN	11	FRAS	4		
NAS	89	FAL	10	PLAS	4		
CHA	79	PAN	10	BRAN	3		
SAS	79	PAS	9	BRAR	3		
CAR	71	BAL	6	DRAS	3		
RAR	67	FAC	6	FRAN	3		
PAR	63	ABS	5	PLAR	3		
LAR	59	RAC	5	FRAR	2		
RRA	56	YAN	5	DRAN	1		
CAN	54	FLA	4	FLAR	1		
NAR	54	GLA	4	FRAZ	1		

lizado en el cuadro 6.1 de la novela y resultan 31.878. Sumados los espacios existentes entre palabras (31.878) con el total de sonidos (66.627) obtenemos los siguientes tantos por cientos:

Silencios.....	32,36%
Sonidos con E.....	18,99%
Sonidos con O.....	13,82%
Sonidos con A.....	19,64%
Sonidos con I.....	10,39%
Sonidos con U.....	4,81%

Que nos indican la enorme importancia de los espacios en los mensajes escritos y de los silencios en las conversaciones.

Como los silencios existen entre las palabras, es posible que los silencios sean igualmente codificados, digitalizados, aunque no es necesario para el objetivo que nos ocupa.

Acuerdo de trabajos

Establecidas diferentes entrevistas con el Dr. Ingeniero de Telecomunicación don Fausto Montoya Vitini, Investigador Científico "Instituto de Teledetección y Telecomunicación" del Consejo Superior de Investigaciones Científicas (C.S.I.C.), se llegó a la firma de un Acuerdo de trabajo en 1988, entre la Subdirección General de Tecnología e Investigación de la D.G.A.M. y el C.S.I.C.

Entre los términos de la colaboración, se planteaba investigar sobre "Nuevas técnicas para la digitalización de la palabra", apoyándose básicamente en el análisis de las sílabas y en un procedimiento de reconstrucción mediante una síntesis del total de sílabas.

En el "Laboratorio de tratamiento de Palabra y Música" del mencionado Instituto, se formó un Equipo por los siguientes colaboradores: doctora Amparo Fúster Sabater, don Rafael Fernández Villalada, don Javier Ponce Martínez y doctor

Javier Sánchez González, bajo la dirección de este último.

Parametrización de Señales

En el caso de codificación de imágenes de diferentes colores, los sensores llegan a determinar la "cantidad" de cada uno de los colores básicos para cada uno de los puntos, formando el compuesto que hay que transmitir.

ción. Esta señal parametrizada debe proporcionar la información necesaria para el reconocimiento, eliminando en lo posible aquella información que no es útil en cada caso.

El parametrizar una determinada señal de voz y representar la evolución de los parámetros calculados a lo largo de la elocución, hace posible realizar una crítica o valoración de los pará-

CUADRO 8							
RESUMEN DE LAS SILABAS (DICCIONARIO ESPAÑOL) AGRUPADAS POR SONIDOS VOCALES Y N.º DE LETRAS							
	1 LT.	2 LT.	3 LT.	4 L.	5 LT.	TOTAL	%
SONIDOS (A) SILABAS DIFERENTES	3.616 1	9.779 30	5.761 132	477 66	36 5	19.669 234	32.76
SONIDOS (E) SILABAS DIFERENTES	731 1	7.144 24	3.982 118	164 44		12.021 187	20.02
SONIDOS (I) SILABAS DIFERENTES	431 1	8.498 23	2.140 116	212 59	1 1	11.282 200	18.79
SONIDOS (O) SILABAS DIFERENTES	1.301 1	8.439 26	3.335 111	158 42	1 1	13.234 181	22.04
SONIDOS (U) SILABAS DIFERENTES	430 1	2.375 24	964 95	62 18	3 3	3.834 141	6.37
TOTAL SONIDOS	6.509	36.235	16.182	1.073	11	60.040	
% DEL TOTAL	10.84	60.35	26.95	1.79	0.07		
TOTAL SILABAS DIF.	5	127	572	229	10	943	

una vez obtenida la codificación con la que ha de reconocerse e identificarse el color que se producirá, gracias a los valores de los parámetros aplicados.

En todo proceso de reconocimiento de las palabras, es prácticamente imprescindible el sustituir la señal muestreada por una representación más compacta y directamente utilizable por los algoritmos de clasifica-

metros elegidos y ver su utilidad en el reconocimiento.

Conversión analógico-digital (A/D)

Básicamente todo sistema de tratamiento de señales analógicas por medios digitales, requiere la conversión de dichas señales en secuencias discretas (con un número finito de mues-

tras tomadas cada cierto tiempo, que puede ser distinto para cada muestra en el caso más general), cuantificadas (con un número de valores posibles para cada muestra) y codificadas de forma que la señal quede representada para que sea utilizable por el ordenador.

Este proceso de muestreo, cuantificación y codificación, recibe el nombre de conversión analógico-digital (A/D).

tal-analógica (D/A) dispondremos nuevamente de la Señal Analógica y oiremos la sílaba de que se trata.

Todo el procesado, conversión analógico-digital, cifrado, transmisión, descifrado y conversión digital-analógica, se realiza normalmente en Transmisión Automática Cifrada, al establecerse la interconexión telefónica entre dos personas.

De todas formas, la mayor

de los fonemas simples vocales (a, e, i, o, u), empiezan las dificultades cuando se trata de diferenciar sílabas compuestas de dos fonemas (pa, la, ta), (pa, pe, po) (1) y la verdadera dificultad está en la identificación de sílabas con mayor número de fonemas y en la variación de una u otras sílabas en el transcurso de la conversación.

Hemos tratado de facilitar las cosas con las diferentes estadísticas y esperamos que representen una ayuda en la codificación, pero pensemos que el tema se mueve entre 10 y 30 parámetros de valores variables.

Por último hay que recalcar que los criterios actuales ya reconocen lo que fue dictaminado por criptólogos y criptoanalistas: "Para lograr la Seguridad 100% en la transmisión de la voz cifrada —imposibilidad de obtener en claro un mensaje— se precisa hacer una suma en módulo 2, bit a bit, entre la información a transmitir y la Serie pseudo-aleatoria generada".

Son descriptables en más o menos tiempo, según los medios y conocimientos aplicados, aquellos Sistemas de Transposición de División de Frecuencias y de Tiempos.

Es imprescindible una digitalización previa al cifrado. Se trata aquí de tomar todas y cada una de las sílabas que forman las palabras, logrando suministrar la información a transmitir, en bits, a mucha menor velocidad.

Con ello en definitiva se salvarían las dificultades que se presentan para la transmisión por líneas físicas y del cifrado en dúplex. ■

(1) Anteriormente (cuadros 8 y 8.1) se detalla que representan el 60,4% las sílabas utilizadas formadas con 2 fonemas.

CUADRO 8.1

RESUMEN DE LAS SILABAS (NOVELA) AGRUPADAS POR SONIDOS VOCALES Y N.º DE LETRAS

	1 LT.	2 LT.	3 LT.	4 L.	5 LT.	TOTAL	%
SONIDOS (A) SILABAS DIFERENTES	3.915 1	10.801 26	4.230 93	386 32	11 1	19.343 153	29.03
SONIDOS (E) SILABAS DIFERENTES	1.581 1	11.268 21	5.655 80	202 17		18.706 119	28.08
SONIDOS (I) SILABAS DIFERENTES	1.479 1	6.878 19	1.797 71	54 18		10.226 109	15.35
SONIDOS (O) SILABAS DIFERENTES	1.378 1	7.981 24	4.144 76	115 15		13.618 116	20.44
SONIDOS (U) SILABAS DIFERENTES	596 1	3.323 21	782 59	33 9		4.734 90	7.10
TOTAL SONIDOS	8.967	40.251	16.608	790	11	66.627	
% DEL TOTAL	13.46	60.40	24.93	1.18	—		
TOTAL SONIDOS DIF.	5	111	379	91	1	587	

Posteriormente en un Sistema de Cifrado, se realiza la mezcla bit a bit, de la Serie conteniendo la Información con la Serie Cifrante.

Realizada la transmisión, se aplica la misma Serie Cifrante mezclándola bit a bit con la Serie recibida y se obtiene así, la Serie digitalizada conteniendo la Información. Haciendo el proceso contrario, conversión digi-

ficultad se presenta en determinar o saber elegir, los parámetros que definan la información que se precisa, para que en la recepción, después de todas las manipulaciones a que sometemos a cada sílaba hablada, sea identificada y reconocida con los fonemas que la componen.

Podemos comprender que es fácil obtener el reconocimiento