

Table 11. Frequencies of the Characters in the Spanish 8M Sample.

Character †	CODES		Character class code	Count †	Percent in total sample	Percent in char. class	Occurrence in total sample	Occurrence in char. class
	ASCII	Unicode						
Blank	32	20	S	1,537,250	19.22	79.96	5.20	1.25
e	101	65	L	797,471	9.97	13.56	10.03	7.37
a	97	61	L	688,798	8.61	11.72	11.61	8.54
o	111	6f	L	538,093	6.73	9.15	14.87	10.93
s	115	73	L	435,390	5.44	7.41	18.37	13.50
n	110	6e	L	415,909	5.20	7.07	19.23	14.14
r	114	72	L	366,836	4.59	6.24	21.81	16.03
i	105	69	L	339,434	4.24	5.77	23.57	17.32
l	108	6c	L	304,215	3.80	5.17	26.30	19.33
d	100	64	L	275,983	3.45	4.69	28.99	21.30
t	116	74	L	266,713	3.33	4.54	29.99	22.04

Table 11. (continued)

Character †	CODES		Character class code	Count †	Percent in total sample	Percent in char. class	Occurrence in total sample	Occurrence in char. class
	ASCII	Unicode						
u	117	75	L	261,918	3.27	4.46	30.54	22.45
c	99	63	L	240,447	3.01	4.09	33.27	24.45
.	46	2e	S	206,220	2.58	10.73	38.79	9.32
m	109	6d	L	160,442	2.01	2.73	49.86	36.64
p	112	70	L	150,993	1.89	2.57	52.98	38.94
,	44	2c	S	115,375	1.44	6.00	69.34	16.66
q	113	71	L	80,380	1.00	1.37	99.53	73.14
b	98	62	L	66,575	0.83	1.13	120.17	88.31
h	104	68	L	63,873	0.80	1.09	125.25	92.04
g	103	67	L	62,506	0.78	1.06	127.99	94.06
y	121	79	L	60,662	0.76	1.03	131.88	96.92
v	118	76	L	59,046	0.74	1.00	135.49	99.57
í	237	ed	L	43,228	0.54	0.74	185.07	136.00
ó	243	f3	L	37,917	0.47	0.64	210.99	155.05
f	102	66	L	35,475	0.44	0.60	225.51	165.73
á	225	e1	L	28,816	0.36	0.49	277.62	204.02
j	106	6a	L	23,464	0.29	0.40	340.95	250.56
é	233	e9	L	22,399	0.28	0.38	357.16	262.47
E	69	45	U	21,070	0.26	11.96	379.69	8.36
z	122	7a	L	19,199	0.24	0.33	416.69	306.22
"	34	22	S	16,918	0.21	0.88	472.87	113.64
P	80	50	U	14,880	0.19	8.45	537.63	11.84
A	65	41	U	14,704	0.18	8.35	544.07	11.98
ñ	241	f1	L	14,700	0.18	0.25	544.22	399.94
S	83	53	U	14,598	0.18	8.29	548.02	12.07
C	67	43	U	13,570	0.17	7.70	589.54	12.98
M	77	4d	U	12,438	0.16	7.06	643.19	14.16
¿	191	bf	S	11,963	0.15	0.62	668.73	160.71
?	63	3f	S	11,934	0.15	0.62	670.35	161.10
L	76	4c	U	9,866	0.12	5.60	810.87	17.86
ú	250	fa	L	8,533	0.11	0.15	937.54	688.99
Y	89	59	U	8,173	0.10	4.64	978.83	21.55
x	120	78	L	7,308	0.09	0.12	1,094.69	804.48
N	78	4e	U	7,174	0.09	4.07	1,115.14	24.56
B	66	42	U	6,819	0.09	3.87	1,173.19	25.83
D	68	44	U	5,869	0.07	3.33	1,363.09	30.02
T	84	54	U	5,460	0.07	3.10	1,465.20	32.26
R	82	52	U	5,010	0.06	2.84	1,596.81	35.16
G	71	47	U	4,513	0.06	2.56	1,772.66	39.04
J	74	4a	U	4,458	0.06	2.53	1,794.53	39.52
H	72	48	U	4,416	0.06	2.51	1,811.59	39.89

Table 11. (continued)

Character †	CODES		Character class code	Count †	Percent in total sample	Percent in char. class	Occurrence in total sample	Occurrence in char. class
	ASCII	Unicode						
i	161	a1	S	4,327	0.05	0.23	1,848.86	444.33
!	33	21	S	4,298	0.05	0.22	1,861.33	447.33
V	86	56	U	4,286	0.05	2.43	1,866.54	41.10
0	48	30	N	4,226	0.05	19.21	1,893.04	5.21
O	79	4f	U	3,919	0.05	2.22	2,041.34	44.95
I	73	49	U	3,725	0.05	2.11	2,147.65	47.29
l	49	31	N	3,624	0.05	16.47	2,207.51	6.07
U	85	55	U	3,583	0.04	2.03	2,232.77	49.17
F	70	46	U	3,439	0.04	1.95	2,326.26	51.23
:	58	3a	S	3,334	0.04	0.17	2,399.52	576.67
;	59	3b	S	2,855	0.04	0.15	2,802.10	673.42
2	50	32	N	2,813	0.04	12.79	2,843.94	7.82
Q	81	51	U	2,615	0.03	1.48	3,059.27	67.37
)	41	29	S	2,245	0.03	0.12	3,563.47	856.40
(	40	28	S	2,234	0.03	0.12	3,581.02	860.62
5	53	35	N	2,135	0.03	9.70	3,747.07	10.31
—	45	2d	S	2,069	0.03	0.11	3,866.60	929.25
9	57	39	N	1,926	0.02	8.75	4,153.69	11.42
k	107	6b	L	1,889	0.02	0.03	4,235.04	3,112.30
3	51	33	N	1,843	0.02	8.38	4,340.75	11.94
4	52	34	N	1,797	0.02	8.17	4,451.86	12.24
8	56	38	N	1,337	0.02	6.08	5,983.55	16.46
6	54	36	N	1,175	0.01	5.34	6,808.51	18.73
7	55	37	N	1,126	0.01	5.12	7,104.80	19.54
_	95	5f	S	910	0.01	0.05	8,791.21	2,112.76
Z	90	5a	U	449	0.01	0.25	17,817.37	392.35
K	75	4b	U	428	0.01	0.24	18,691.59	411.60
w	119	77	L	388	0.00	0.01	20,618.56	15,152.39
/	47	2f	S	357	0.00	0.02	22,408.96	5,385.47
W	87	57	U	251	0.00	0.14	31,872.51	701.85
É	201	c9	U	245	0.00	0.14	32,653.06	719.04
X	88	58	U	181	0.00	0.10	44,198.90	973.29
'	39	27	S	142	0.00	0.01	56,338.03	13,539.54
ü	252	fc	L	129	0.00	0.00	62,015.50	45,574.64
+	43	2b	S	56	0.00	0.00	142,857.14	34,332.39
%	37	25	S	47	0.00	0.00	170,212.77	40,906.68
♂	186	ba	S	45	0.00	0.00	177,777.78	42,724.76
ç	231	e7	O	31	0.00	34.44	258,064.52	2.90
Ö	210	d2	O	28	0.00	31.11	285,714.29	3.21
=	61	3d	S	19	0.00	0.00	421,052.63	101,190.21

Table 11. (continued)

Character †	CODES		Character class code	Count †	Percent in total sample	Percent in char. class	Occurrence in total sample	Occurrence in char. class
Ñ	209	d1	U	16	0.00	0.01	500,000.00	11,010.31
ì	236	ec	O	16	0.00	17.78	500,000.00	5.62
>	62	3e	S	14	0.00	0.00	571,428.57	137,329.57
Á	193	c1	U	6	0.00	0.00	1,333,333.33	29,360.83
ö	246	f6	O	5	0.00	5.56	1,600,000.00	18.00
è	232	e8	O	3	0.00	3.33	2,666,666.67	30.00
Ú	218	da	U	2	0.00	0.00	4,000,000.00	88,082.50
Ö	214	d6	O	1	0.00	1.11	8,000,000.00	90.00
ä	228	e4	O	1	0.00	1.11	8,000,000.00	90.00
Û	220	dc	U	1	0.00	0.00	8,000,000.00	176,165.00
à	224	e0	O	1	0.00	1.11	8,000,000.00	90.00
Í	205	cd	U	1	0.00	0.00	8,000,000.00	176,165.00
♀	170	aa	S	1	0.00	0.00	8,000,000.00	1,922,614.00
ê	234	ea	O	1	0.00	1.11	8,000,000.00	90.00
î	238	ee	O	1	0.00	1.11	8,000,000.00	90.00
ù	249	f9	O	1	0.00	1.11	8,000,000.00	90.00
ò	242	f2	O	1	0.00	1.11	8,000,000.00	90.00
Unknown	9	9	S	1	0.00	0.00	8,000,000.00	1,922,614.00
Total	†	†	†	8,000,000				