

Frequencies of Japanese 100 Basic Morae and 3 Special Sounds

	Phonemic Comb.	Mora	Cons/semi	Vowel	Katakana	Token Frequency	Type Frequency
Vowels - /a/, /e/, /i/, /o/ and /u/							
1	V	a		a	ア	6,149,909	18,898
2	V	i		i	イ	43,985,426	67,643
3	V	u		u	ウ	39,052,254	84,624
4	V	e		e	エ	4,767,153	14,103
5	V	o		o	オ	14,053,377	22,195
/p/, /t/ and /k/ (voiceless, stop)							
6	CV	pa	p	a	パ	984,460	2937
7	CV	pi	p	i	ピ	192,847	1343
8	CV	pu	p	u	プ	445,557	3480
9	CV	pe	p	e	ペ	182,551	1165
10	CV	po	p	o	ポ	507,561	2183
11	CV	ta	t	a	タ	16,442,465	27384
12	CV	ti (cji)	t (cj)	i	チ	5,808,199	19482
13	CV	tu	t	u	ツ	9,612,206	22853
14	CV	te	t	e	テ	11,037,767	12822
15	CV	to	t	o	ト	17,102,180	27131
16	CV	ka	k	a	カ	20,530,065	43982
17	CV	ki	k	i	キ	9,926,212	31183
18	CV	ku	k	u	ク	17,211,261	38359
19	CV	ke	k	e	ケ	6,823,149	14873
20	CV	ko	k	o	コ	12,154,255	27551
/s/ and /h/ (voiceless, fricative)							
21	CV	sa	s	a	サ	7,932,086	22624
22	CV	si	s	i	シ	16,518,016	40585
23	CV	su	s	u	ス	9,271,734	22763
24	CV	se	s	e	セ	8,127,590	16836
25	CV	so	s	o	ソ	4,303,939	8857
26	CV	ha	h	a	ハ	12,594,714	12919
27	CV	hi	h	i	ヒ	3,554,252	11507
28	CV	hu	h	u	フ	3,067,681	11707
29	CV	he	h	e	ヘ	1,254,559	2633
30	CV	ho	h	o	ホ	4,064,902	8336
/b/ and /g/ (voice, stop)							
31	CV	ba	b	a	バ	2,749,654	10011
32	CV	bi	b	i	ビ	1,177,824	5569
33	CV	bu	b	u	ブ	2,758,994	9263
34	CV	be	b	e	ベ	1,591,698	3769
35	CV	bo	b	o	ボ	1,104,998	5317
36	CV	da	d	a	ダ	9,953,427	17622
37	CV	de	d	e	デ	7,744,822	5965
38	CV	do	d	o	ド	5,306,817	12158
39	CV	ga	g	a	ガ	11,996,376	17737
40	CV	gi	g	i	ギ	2,009,223	6249
41	CV	gu	g	u	グ	1,210,586	8951
42	CV	ge	g	e	ゲ	2,328,945	6422
43	CV	go	g	o	ゴ	2,549,499	7088
/z/ (voice, fricative)							
44	CV	za	z	a	ザ	1,384,993	5736
45	CV	zi	z	i	ジ	6,550,231	15529
46	CV	zu	z	u	ズ	1,644,763	8670
47	CV	ze	z	e	ゼ	1,646,454	3024
48	CV	zo	z	o	ゾ	936,779	3125
/m/ and /n/ (voice, nasal)							
49	CV	ma	m	a	マ	6,289,027	28364
50	CV	mi	m	i	ミ	3,972,466	21459
51	CV	mu	m	u	ム	2,577,386	8694
52	CV	me	m	e	メ	4,118,162	9174

53	CV	mo	m	o	モ	6,791,877	12180
54	CV	na	n	a	ナ	10,498,607	18207
55	CV	ni	n	i	ニ	14,757,618	10038
56	CV	nu	n	u	ヌ	411,356	2034
57	CV	ne	n	e	ネ	2,771,906	6260
58	CV	no	n	o	ノ	19,326,674	14413
<hr/>							
/r/ (voice, liquid)							
59	CV	ra	r	a	ラ	6,947,107	20442
60	CV	ri	r	i	リ	6,116,529	25431
61	CV	ru	r	u	ル	28,899,891	23656
62	CV	re	r	e	レ	5,777,889	9566
63	CV	ro	r	o	ロ	2,927,552	13925
<hr/>							
/w/ and /j/ (voice; semivowel)							
64	SV	ja	j	a	ヤ	3,432,278	13622
65	SV	ju	j	u	ユ	1,620,066	10491
66	SV	jo	j	o	ヨ	4,484,240	5464
67	SV	wa	w	a	ワ	3,777,201	13922
<hr/>							
/pj/, /tj/ and /kj/ (voiceless, stop)							
68	CSV	pja	pj	a	ピャ	95	9
69	CSV	pju	pj	u	ピュ	29,907	147
70	CSV	pjo	pj	o	ピョ	125,380	82
71	CSV	tja	tj	a	チャ	199,251	1216
72	CSV	tju	tj	u	チュ	995,900	1898
73	CSV	tjo	tj	o	チョ	2,206,487	9007
74	CSV	kja	kj	a	キャ	166,750	551
75	CSV	kju	kj	u	キュ	929,206	1814
76	CSV	kjo	kj	o	キョ	2,721,853	4139
<hr/>							
/sj/ and /hj/ (voiceless, fricative)							
77	CSV	sja	sj	a	シャ	2,138,138	3208
78	CSV	sju	sj	u	シュ	2,858,438	5175
79	CSV	sjo	sj	o	ショ	3,740,881	8688
80	CSV	hja	hj	a	ヒャ	53,284	170
81	CSV	hju	hj	u	ヒュ	3,304	69
82	CSV	hjo	hj	o	ヒョ	577,939	700
<hr/>							
/bj/ and /gj/ (voice, stop)							
83	CSV	bja	bj	a	ビャ	511	45
84	CSV	bju	bj	u	ビュ	16,016	83
85	CSV	bjo	bj	o	ビョ	162,751	504
86	CSV	gja	gj	a	ギャ	59,424	207
87	CSV	gju	gj	u	ギュ	15,343	117
88	CSV	gjo	gj	o	ギョ	937,044	2084
<hr/>							
/zj/ (voice, fricative)							
89	CSV	zja	zj	a	ジャ	188,263	1166
90	CSV	zju	zj	u	ジュ	1,482,387	3623
91	CSV	zjo	zj	o	ジョ	2,041,928	4929
<hr/>							
/mj/ and /nj/ (voice, nasal)							
92	CSV	mja	mj	a	ミャ	14,646	113
93	CSV	mju	mj	u	ミュ	15,494	201
94	CSV	mjo	mj	o	ミョ	33,782	400
95	CSV	nja	nj	a	ニャ	3,087	69
96	CSV	nju	nj	u	ニュ	446,886	738
97	CSV	njo	nj	o	ニョ	10,406	175
<hr/>							
/rj/ (voice liquid)							
98	CSV	rja	rj	a	リャ	50,769	183
99	CSV	rju	rj	u	リュ	214,108	1122
100	CSV	rjo	rj	o	リョ	1,541,224	2442
<hr/>							
/N/ (voice, nasal)							
101	N	N	-	-	ン	35,719,268	81,463
<hr/>							
/R/ (long vowel including two same vowels simply continued)							
102	R	R	-	-	ー	3,607,169	23,598
<hr/>							
/Q/ (geminate)							
103	Q	Q	-	-	ッ	7,858,931	14,435

Note: The token and type frequencies of a long vowel /R/ includes continuous two vowels as in /baai/ meaning 'case'.

Appendix B Additional 33 morae (or kana) used for describing alphabetic loanwords

	Phonemic Comb.	Mora	Cons/semi	Vowel	Katakana	Token Frequency	Type Frequency
102	CSV	sje	sj	e	シエ	14924	149
103	CSV	cje	cj	e	チェ	37633	269
104	CV	ca	c	a	ツア	2465	43
105	CV	ce	c	e	ツエ	4802	38
106	CV	co	c	o	ツオ	675	14
107	CV	ti	t	i	ティ	66313	1205
108	CV	fa	f	a	ファ	60484	681
109	CV	fi	f	i	フィ	50615	474
110	CV	fe	f	e	フェ	16129	346
111	CV	fo	f	o	フォ	31399	408
112	CSV	zje	zj	e	ジエ	30738	255
113	CV	di	d	i	ディ	65994	947
114	CSV	dju	dj	u	デュ	8344	65
115	SV	je	j	e	イエ	20	7
116	SV	wi	w	i	ウィ	12605	170
117	SV	we	w	e	ウェ	25909	226
118	SV	wo	w	o	ウォ	7747	109
119	CSV	kwa	kw	a	クア	21	2
120	CSV	kwi	kw	i	クイ	20	6
121	CSV	kwe	kw	e	クエ	2	1
122	CSV	kwo	kw	o	クオ	85	10
123	CV	ci	c	i	ツイ	13067	15
124	CV	tu	t	u	トゥ	1854	78
125	CSV	gwa	gw	a	グア	25	5
126	CV	du	d	u	ドゥ	1542	55
127	CSV	vwa	vw	a	ヴァ	641	80
128	CSV	vwi	vw	i	ヴィ	733	81
129	CV	vu	v	u	ヴ	283	29
130	CSV	vwe	vw	e	ヴェ	311	38
131	CSV	vwo	vw	o	ヴォ	66	15
132	CSV	tju	tj	u	テュ	72	8
133	CSV	fju	fj	u	フュ	206	7
134	CSV	vju	vj	u	ヴュ	34	4

Note1: The additional 33 kana are listed on the basis of a cabinet announcement of the Government of Japan (1991)

Note2: The frequencies of the 33 kana were counted using words described in katakana.

Note3: The phonemic expressions of these kana were adapted from Matsuzaki (1993) except /v/ for ヴァ, ヴィ, ヴ, ヴェ and ヴォ.