

## Kirjainten ja äänteiden useudet Suomessa

Tämän kirjoittaja on koonnut tilastoa sekä yhdistänyt omaan tilastoonsa muiden tekemää tilastoa kirjainten ja kirjainparien esiintymisestä suomenkielisessä tekstissä, joka on valittu eri kirjailijain sekä sanomalehtiartikkelien teksteistä. Tilasto on tehty kirjoituskoneen työtutkimuksen yhteydessä, joten se ei kaikissa kohdissaan täytä sitä, mitä äänneopilliselta tilastolta olisi vaadittava, mutta kun tämänkin tilaston tulokset saattavat olla kiintoisat, esitän ne tässä.

Tilastoitu määrä on 66 344 kirjainta, ja tulokset on laskettu 10 000 kirjainta kohti, pääasiassa tasoittaen kokonaislukuihin. Kirjainpariin kuuluvien kirjainten järjestystä ei ole otettu huomioon (koska sillä ei ole merkitystä kirjoituskoneen näppäimistön suunnittelussa). Siis esim. *na* 181 merkitsee seuraavassa taulukossa 1, että *pareja na* ja *an* on yhteensä 181 esiintymää 10 000 kohti.

### 1. Kirjainparien useus suomessa (10 000 kirjainta kohti).

ta	263	ke	80	kä	35	lk	18	rs	8	rp	2,3
si	251	su	79	vä	34	rj	17	np	8	hn	2,3
aa	190	ra	79	kk	34	hd	17	ea	7	pt	2,0
ti	188	ko	78	mu	33	lm	17	yi	7	mk	2,0
te	183	ku	75	uu	33	do	17	sö	7	ln	2,0
na	181	mm	70	pi	31	nk	16	dy	6	ga	1,7
ne	176	nä	69	ha	31	po	15	ij	6	mö	1,7
ni	163	re	64	pu	29	äy	15	hk	6	ge	1,5
se	153	he	63	ss	29	oo	15	ju	6	pl	1,4
tu	132	ii	62	ht	28	hu	15	kö	6	ps	1,4
sa	130	lu	61	hä	27	jä	15	py	6	vö	1,4
tt	129	ou	61	ru	26	pp	14	äe	5	ya	1,4
la	128	pa	59	sy	26	ny	14	sm	5	rö	1,2
tä	123	jo	58	rä	25	pe	14	öi	5	nm	1,2
le	122	ty	57	ky	25	lt	13	ls	5	go	1,2
va	118	nu	56	my	24	ao	13	lv	5	vk	1,1
st	116	mä	55	ui	23	tk	12	hm	5	äa	0,8
me	114	lo	55	du	23	yö	12	hr	4	mv	0,6
ki	114	nt	52	äi	22	da	11	ey	4	pö	0,6
ai	113	ri	49	ve	22	ly	11	lj	4	hö	0,5
ka	112	so	49	ro	22	ry	10	oe	4	öö	0,5
oi	105	ää	46	vo	22	rt	10	lö	3	rn	0,3
ma	98	vi	45	yy	21	dä	9	gi	3	nd	0,3
ks	96	au	45	je	21	pä	9	jy	3	kg	0,3
to	95	lä	45	vu	21	vy	9	yo	3	dö	0,15
li	95	nn	44	di	21	ng	9	hj	3	rg	0,15
mi	93	ee	44	ho	20	nö	9	hl	3	sj	0,15
no	87	mo	43	eu	20	mp	8	sv	3	td	0,15
ei	84	sä	42	rk	20	rr	8	rm	3	äo	0,15
ll	84	de	41	ns	18	öt	8	jn	2,4		
ja	80	hi	36	hy	18	rv	8				

Kirjainparien avulla saadaan helposti erikseen lyhyiden ja pitkien ääntiöiden, yksinkertaisten ja kerrattujen kerakkeiden esiintymät. Tämän nojalla on laadittu kolmas taulukko.

Saneenalkuiset ja -loppuiset kirjaimet on tilastoitu pareina »kirjain + tyhjä väli» (taulukko 2). Niistä saadaan myös sanevälkien ja siis saneiden lukumäärä  $1\ 195$ . Saneen keskipituus oli  $10\ 000/1\ 195 = 8,35$  kirjainta. Jos pitkät ja kahdennetut äänteet luetaan yhdeksi, niin  $10\ 000$  kirjaimella on merkitty  $7\ 836$  äännettä ja saneen pituus oli  $6,57$  äännettä.

Äänteet voidaan jakaa viiteen ryhmään, joiden useus eroaa jyrkästi toisistaan (taulukko 4).

## 2. Saneenalkuiset ja -loppuiset kirjaimet suomessa (10 000 kirjainta kohti)

n	504	i	195	o	126	p	63	h	29
a	376	e	180	m	106	l	55	r	19
ä	219	s	147	v	89	u	47	ö	1
t	209	k	145	j	87	y	30		

## 3. Yksittäisten kirjainten (äänteiden) useus suomessa (10 000 kirjainta kohti)

merkki	kirjainta	äänteitä		merkki	kirjainta	äänteitä	
		yhteensä	pitkiä			yhteensä	pitkiä
a—aa	1 120	930	190	y—yy	162	141	21
i—ii	952	890	62	h	156	156	0
t—tt	874	745	129	j	151	151	0
n—nn	735	666	44	p—pp	138	124	14
e—ee	728	684	44	d	73	73	0
s—ss	603	574	29	ö—öö	30	29	1
o—oo	457	442	15	n—ng (η, ηη) (25)		25	9
k—kk	457	423	34	f	14	14	0
ä—ää	424	378	46	g	9	—	—
u—uu	413	380	33	b	5	5	0
l—ll	408	324	84	c	1	—	—
m—mm	377	307	70	,	149		
r—rr	194	186	8	.	123		
v	189	189	0	-	60	(saneenkeskeisiä 44)	

## 4. Kirjainten (äänteiden) useusryhmät

Kirjaimet (äänteet)	useuden rajat	yhteisuseus	keskiuseus (ryhmän)
a i t n e s	1 120—603	5 012	835
o k ä u l m	457—377	2 536	423
r v y h j p	194—138	990	165
d ö ng (η)	73—25	128	43
f g b c	14—1	29	7

Saamistamme taulukoista voimme vielä laskea eri äänneryhmien suhteelliset useudet tutkitun aineiston perusteella. Niinpä saamme *vokaaleista*:

etuvokaaleja .....	2 122 = 54,7 %	vokaaleista
takavokaaleja.....	1 752 = 45,3 »	»
pitkiä vokaaleja.....	412 = 10,6 »	»
diftongeja .....	(ei yli) 348 = 9,0 »	»
pitkiä ja diftongeja yht. ....	760 = 19,6 »	»
vokaalien summa .....	3 874	

Diftongien lukumäärä on kuitenkin liian suuri, koska niihin sisältyy myös sellaisia vokaalipareja, jotka kuuluvat eri tavuihin, ja näiden lukumäärää ei ilman muuta voi erottaa tästä tilastosta (esim. saneen 'vesi-eläin' keskellä oleva *ie* esiintyisi ryhmässä *ei*).

*Konsonanttien* jakautuminen:

	klusilli	frikatiivi	lateraali	tremul.	reson.	summa
dentaali .....	818	574	324	186	666	2 568
labiaali .....	129	213			307	649
palataali .....	423	151			25	599
laryngaali .....		156				156
yhteensä	1 370	1 094	324	186	998	3 972
konsonanttien yhteissumma .....						3 972

Dentaali: labiaali: palataali: laryngaali = 4: 1: 1: 0,25 = 16: 4: 4: 1

Vokaalit: konsonantit = 3 874: 3 972 = 100: 103

Mitä taka- ja etuvokaalien esiintymisen useuteen tulee, ei edellä esitetty suhde (45,3: 54,7) anna oikeaa kuvaa taka- ja etuvokaalisten *saneiden* suhteesta toisiinsa, koska *i* ja *e* esiintyvät kummassakin ryhmässä. Itse asiassa on takavokaalisia saneita paljon enemmän kuin etuvokaalisia, kuten jo osoittaa sekin seikka, että vokaaleja *a*, *o*, *u* on 1 752, *ä*, *ö*, *y* 548 ja *e*, *i* 1 574 eli suhteessa 3,2: 1: 2,9. Tarkemmat saneiden suhteet tulen esittämään myöhemmin saneiden useuksia koskevien muiden tulos-ten yhteydessä.

VILHO SETÄLÄ