

Kvantitativ undersøgelse af sammenhængen mellem vokal og stavemåde i dansk retskrivning

Introduktion

Jeg har [tidligere undersøgt sammenhængen mellem stavemåde og vokal](#). I denne undersøgelse går jeg den anden vej ved hjælp af de data jeg kom frem til i den anden undersøgelse. Jeg vil forsøge at finde ud af hvilke vokalbogstaver de forskellige vokaler er forbundet med. Før fandt jeg ud af hvilke vokaler de forskellige vokalbogstaver er forbundet med. Første mål er at finde ud af hvor mange ord der er med de forskellige vokalbogstaver. Andet mål er at finde ud af sammenhængen mellem alle vokalerne og alle vokalbogstaverne i alle kombinationer.

Forkortelser

DAT = Dictionary Analysis Tool, [link](#).

Metode

Jeg ved allerede hvilke vokaler de forskellige vokalbogstaver er forbundet med. Vha. DAT kan jeg se hvor mange ord der er med de forskellige vokalbogstaver. Derfra kan jeg slutte til hvor mange ord med de forskellige vokaler der er. Dette kan jeg gøre med alle vokalbogstaver og alle vokalerne og se hvilke vokalbogstaver der oftest er forbundet med hvilke vokaler. Selvom fx at Å er stærkt forbundet til [å] (63%), så er også O forbundet til [å] (13%). Hvis der er mange flere ord med O end med Å, så er sammenhængen mellem [å] og O måske stærkere end den mellem [å] og Å. (I dette tilfælde skal der være ca. >4.8 gange så mange ord med O.)

Resultater

Vha. DAT kan man nemt finde ud af hvor mange ord der er med de forskellige vokalbogstaver.

Bogstav	Antal ord
A	27642
E	45191
I	30213

O	21804
U	12943
Y	5351
Æ	7504
Ø	6342
Å	2837
Alle ord ¹	62832

Sammenhængen mellem vokaler og vokalbogstaver

Vokal/Bogstav	A	E	I	O	U	Y	Æ	Ø	Å	SUM
[a]	16562									16562
[ɑ]	11080	1255								12335
[e]		11298	9684							20982
[ɛ]		19771								19771
[ä]		1883					1895			3778
[æ]		5335	1439				5609			12383
[ɔ]		5335		9920	237			995	1050	17537
[i]			20861			105				20966
[o]				9232	119					9351
[å]				2652	2612				1787	7051
[y]					237	4013				4251
[u]					8668					8668
[ø]						682		2736		3418
[ö]						498		1399		1897
[õ]								1212		1212
Fordeling %										SUM
[a]	100.00%									100.00%
[ɑ]	89.82%	10.18%								100.00%
[e]		53.85%	46.15%							100.00%
[ɛ]		100.00%								100.00%
[ä]		49.85%					50.15%			100.00%
[æ]		43.08%	11.62%				45.30%			100.00%
[ɔ]		30.42%		56.57%	1.35%			5.67%	5.99%	100.00%
[i]			99.50%			0.50%				100.00%
[o]				98.73%	1.27%					100.00%
[å]				37.61%	37.05%				25.35%	100.00%
[y]					5.59%	94.41%				100.00%
[u]					100.00%					100.00%
[ø]						19.95%		80.05%		100.00%
[ö]						26.27%		73.73%		100.00%
[õ]								100.00%		100.00%

I øverste del ses antal ord med en given vokal og et givet vokalbogstav. Nedenunder ses sammenhængen mellem vokaler og vokalbogstaver. Det læses således: vælg en vokal ude til venstre, fx [ɑ]. Derefter følg rækken til du rammer et tal, dvs. 89.82%. Derefter kik op for at finde ud af hvilket bogstav der er tale om, dvs. A. Tallene betyder at 89.82% af ord med [ɑ] staves med A. Forsæt i rækken til du finder et andet tal, dvs. 10.18%. Altså, 10.18% af ord med [ɑ] staves med E, etc. Af overraskende resultater kan det jo bl.a. nævnes at O er det bogstav med størst sammenhæng med [å], og derefter kommer U og Å, med hhv. 37.61%, 37.05% og 25.35%!

¹ Bemærk at dette ikke er lig med summen af alle de ovenstående. Hvis man tog summen, så ville man tælle en masse ord flere gange.