

Kvantitativ undersøgelse af sammenhængen mellem stavemåde og vokal i dansk retskrivning

Indledning

Første mål er at finde ud af om alfabetprincippet er sandt. Andet mål er at finde ud af hvilke vokaler der er normale for hvilke bogstaver. Normal er her defineret som en sammenhæng fra bogstav til lyd i 10% eller mere af tilfældende. Metoden er at der kikkedes på tilfældigt udvalgte ord med hver vokalbogstav i dansk og deres udtale noteres vha. et ordbog.

Terminologi

PNO = Politikens Nudansk Ordbog, 2005, [link](#).

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

Metode

Der tages udgang i alfabetprincippet. Alfabetprincippet er at den mest normale vokal for et bogstav er den som det udtales med i alfabetet, altså hhv. A [a], E [e], I [i], O [o], U [u], Y [y], Æ [æ], Ø [ø], Å [å]. Denne undersøgelse undersøger om dette nu også er tilfældet, da undersøgelsen forsøger at klarlægge udtalerne af alle vokalbogstaver.

Metoden er at vælge 100 tilfældige ord stavet med det valgte bogstav vha. DAT. Derefter slås de op i PNO og undersøges hvad udtalen er. Dette gælder også selvom jeg er uenig i udtalen angivet i PNO (hvilket jeg var i et par tilfælde). Den vokal som det valgte bogstav udtales med noteres i en tabel. Hvis der er varianter noteres begge og men tælles kun for 0.5 hver. I mange tilfælde findes det specifikke ord ikke i ordbogen, i stedet forsøges at bruge et lign. ord eller at bruge de ord/morfemer som ordet består af til at finde udtalen. Det noteres hvordan man fandt udtalen. Hvis ingen af delene lykkedes, så droppes ordet og et nyt tilfældigt ord findes vha. DAT.

Kritik

Nogle bogstaver forekommer meget ofte i bøjninger som ikke findes i ordbogen. Det drejer sig især om pluralis-bøjninger og bestemt form bøjninger. På dansk er det især bogstaverne E, R, N og T der

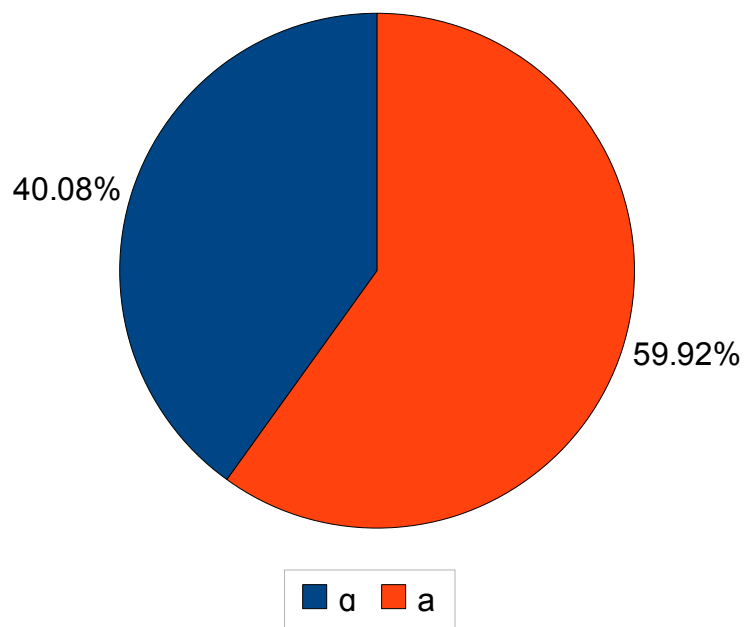
forekommer i sådanne bøjninger. R, N og T er ikke relevante for denne undersøgelse, da de ikke er forbundet med nogen vokal. E er relevant og man må blot notere sig, at undersøgelsen ikke tager højde for dette.

Nogle ord forekommer meget oftere end andre i skrevne tekster. At udvælge ord tilfældigt fra en ordbog tager ikke højde for dette. For at tage højde for dette, så skulle man bruge en ordbog som også har information om frekvensen af ordene. Sådan en kender jeg ikke til.

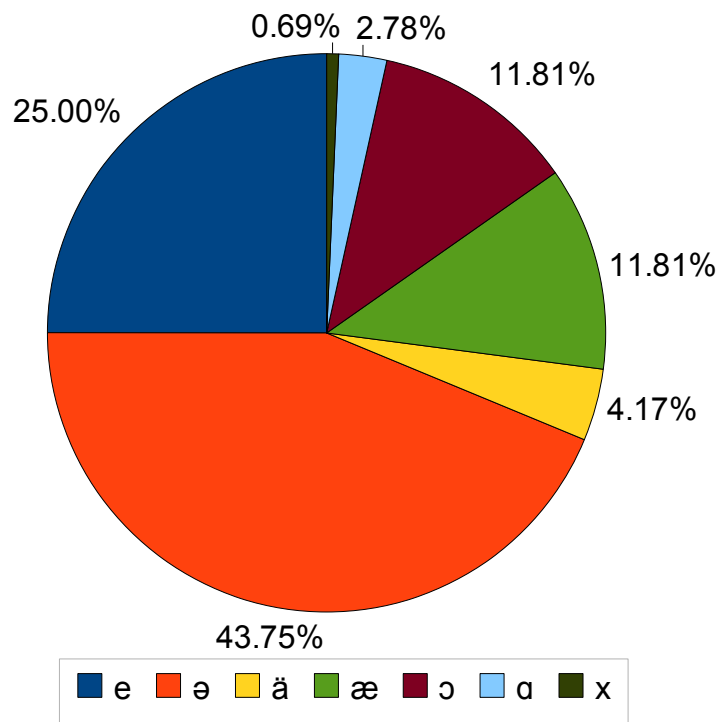
Resultater

Cirklen og procenttallene viser fordelingen. Boksen angiver hvad cirklens dele illustrerer, dvs. hvilke vokaler der er tale om. "x" indikerer at bogstavet er stumt. De mere specifikke resultater findes særskilt.

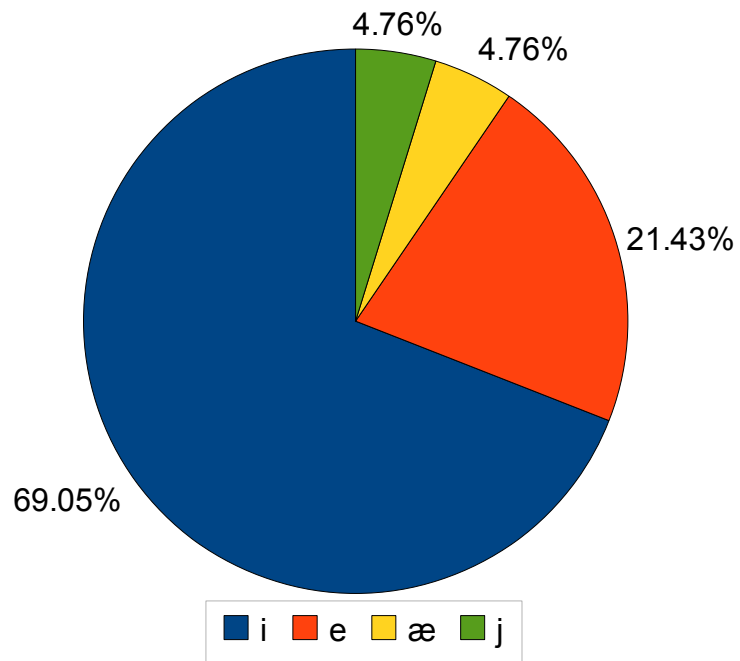
A



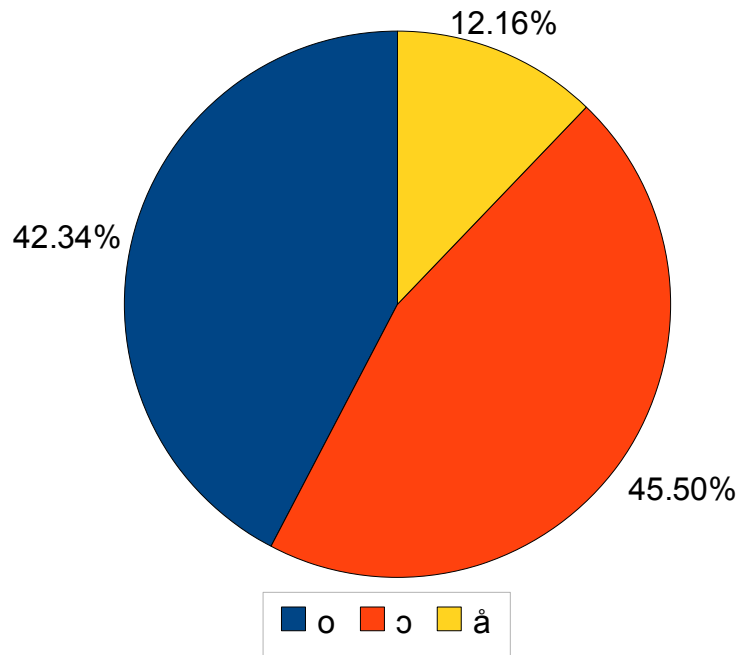
E



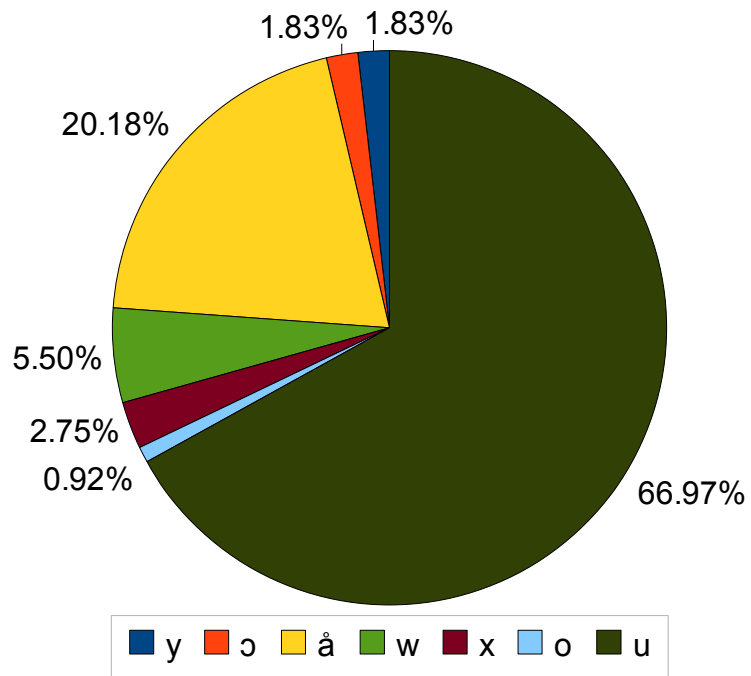
I



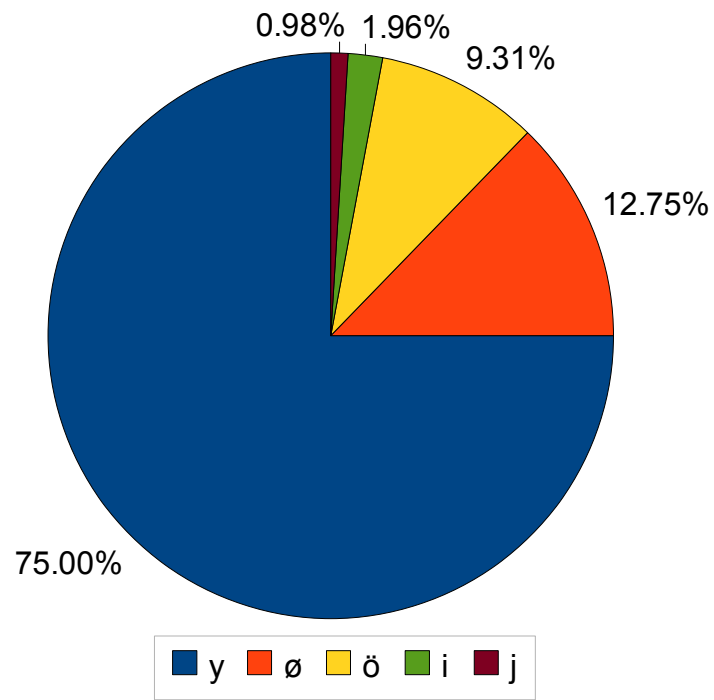
O



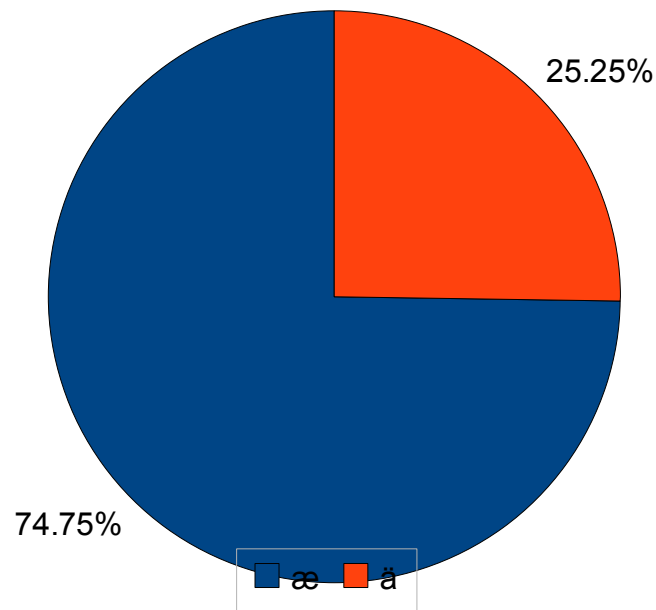
U



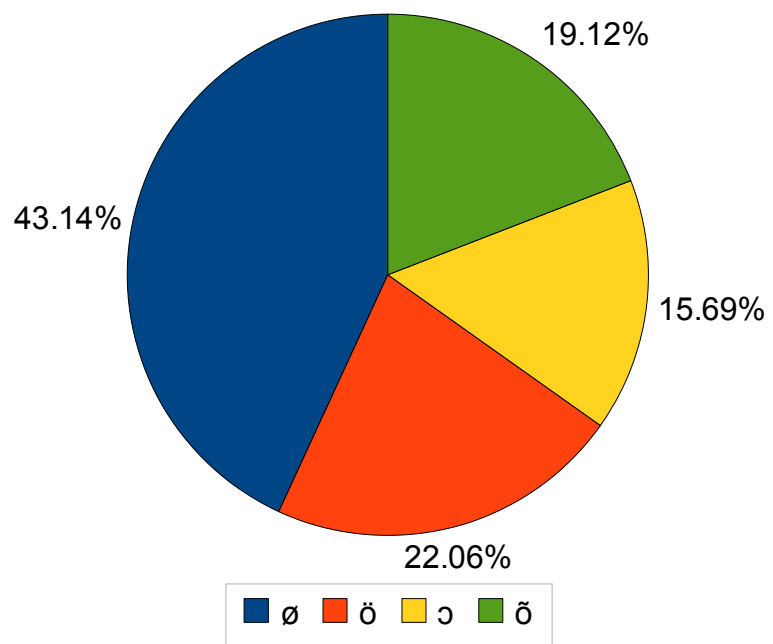
Y



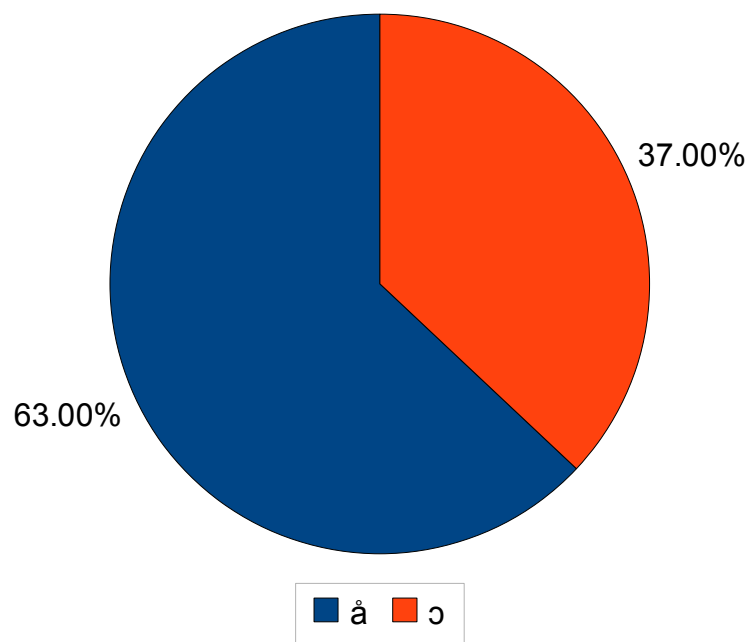
Æ



Ø



Å



Diskussion

Alfabetprincippet

Alfabetprincippet kan ikke siges at være bekræftet. Resultaterne er i overensstemmelse med bogstaver A, I, U, Y, Æ, Ø og Å. Resultaterne er ikke i overensstemmelse med E, og heller ikke med O selvom det er tæt løb. Den statistiske usikkerhed gør at man ikke kan konkludere noget ved bogstavet O. Alfabetprincippet er altså falsk.

Måske er en revideret version af alfabetprincippet dog sandt, hvis man siger at det kun gælder for E'er som ikke står i sidste stavelse i flerstavellesord, samt O efter R. En ny undersøgelse er nødvendig for at undersøge om dette er tilfældet. En anden hypotese lign. alfabetprincippet er at den mest normale *lange* vokal for et bogstav er den som det udtales med i alfabetet. Dette tror jeg umiddelbart er sandt, men jeg har ikke noget bevis for det. Denne undersøgelse tog ikke højde for lange og korte vokaler, så ud fra de data der findes i denne undersøgelse kan man ikke konkludere noget.

Normale vokaler

Skemaet herunder viser de normale vokaler for de forskellige bogstaver, opstillet efter udbredelse.

Bogstav	Normale vokaler
A	[a], [ɑ]
E	[ɘ], [e], [æ], [ɛ]
I	[i], [e]
O	[ɔ], [o], [å]
U	[u], [å]
Y	[y], [ø], ([ö])
Æ	[æ], [ä]
Ø	[ø], [ö], [õ], [ɔ]
Å	[å], [ɔ]