

# Unicode: En ganske kort introduktion

eller

## Hvordan computeren opfatter andre alfabeter end det danske

Søren Holst, juni '07

Mange teologer har oplevet at skrive et par hebraiske eller græske brokker i en tekst, tage filen med et andet sted hen for at printe den eller skrive videre på den – og så opdage, at de pæne bibelske bogstaver er blevet til noget (endnu mere) ulæseligt krimskrams, når man åbner den på en anden computer. Det sker kun, når man benytter en forældet teknologi, hvor man ”narrer” computeren til at vise fx et hebraisk *alef* på skærmen, selv om man har tastet den elektroniske bogstavkode for et dansk ”a”, eller hvad det pågældende system nu bruger som tegn for *alef*.

Fremtiden mht. sære skrifttegn på computeren hedder *Unicode*. Det er navnet på en international standard, hvor der til hvert bogstav i næsten alle tænkelige alfabeter svarer en særskilt kode inde i computeren. Det var der simpelthen ikke plads til i tidligere tiders såkaldte ASCII-kodning, som kun kunne skelne mellem i alt 256 tegn – det rækker ikke til så mange alfabeter. Men fra nu af kan computeren altså skelne mellem ”a” og *alef* og behøver aldrig blande dem sammen.

Det kan naturligvis stadig tænkes, at en given Unicode-skrifttype slet ikke indeholder hebraisk: Hvis en softwareudvikler i Korea designer en font til brug på en thailandsk web-side, har han ikke behov for at inkludere det hebraiske alfabet. Men pointen med Unicode er, at standarden skelner alle alfabeter og tegnsæt fra hinanden, så hebraisk vil blive gengivet korrekt som hebraisk i alle fonte, der indeholder hebraisk, og aldrig blive blandet sammen med thai, koreansk eller dansk. Hvad du taster ind er altså altid entydigt, og hvis du ikke ligefrem skriver runer eller sumerisk, vil der næsten altid være en font på en up-to-date computer, der kan gengive det.

Fra og med Windows XP er muligheden for at skrive i Unicode nemlig indbygget i de almindeligste programmer. Så vidt jeg véd, skulle det også køre perfekt i et gratis tekstbehandlingsprogram som Open Office Writer, men jeg har ikke prøvet. Både Mac og Linux skulle efter sigende også kunne være med.

Entydigheden har dog en pris, men den er beskedent: Da der ikke er et særskilt hebraisk og græsk alfabet på et dansk tastatur (det ville blive et pokkers stort ét), bliver man nødt til at benytte et ”logisk tastatur”, dvs. et lille program, som fortæller PC'en, at det, der lige nu tages ind på det danske ”fysiske tastatur”, skal opfattes som et andet alfabet. Til klassisk græsk findes det allerede i Windows XP og skal blot hentes frem, til bibelhebraisk skal det installeres. Vejledning i begge dele findes på web-siden [www.itu.dk/people/sho](http://www.itu.dk/people/sho) (flytter nok på et tidspunkt til [www.sorenholst.dk](http://www.sorenholst.dk)).

Meget mere om elektronikken bag bogstaverne kan læses på Rodney Decker's side <http://www.ntresources.com/unicode.htm>, hvor man også kan hente den storartede artikel ”What a Biblical Scholar Should Know about Unicode”. Decker er desværre NT-professor, så han gør mere ud af græsk end hebraisk. Men artiklen er yderst læseværdig.

God fornøjelse.