

Voikkoindex on LuaLaTeX-paketti, joka tekee hakemistoja ja osaa taivuttaa suomen kielen sanoja.

Voikkoindex on vapaa ohjelma, jota voidaan levittää ja muuttaa GNU General Public License'n versiolla 3 tai uudemmalla.

Voikkoindex is a LuaLaTeX package for texts written in Finnish.

Voikkoindex is free software: you can redistribute it and/or modify it under the terms of the GNU General Public License as published by the Free Software Foundation, either version 3 of the License, or (at your option) any later version.

Käyttö LaTeX-tiedostoissa

```
\usepackage[langcode="fi",  
             path="/usr/local/lib/voikko",  
             nimet=Henkilöhakemisto,  
             paikat=Paikannimihakemisto,  
             sanat=Asiahakemisto]{voikkoindex}
```

`langcode` pitää olla joko `"fi"` tai `"fi-x-sukija"` Huomaa, että parametrissa pitää olla lainausmerkit alussa ja lopussa.

`path` on hakemisto, jossa Voikon sanastotiedostot ovat. Lainausmerkkien pitää olla myös tämän parametrin alussa ja lopussa.

`nimet`, `paikat` ja `sanat` ovat hakemistojen nimet, jotka `splitidx`'n komento `\printindex` tulostaa.

Nuo ovat oletusarvoja eli jos ne kelpaavat, voidaan kirjoittaa näin

```
\usepackage{voikkoindex}
```

Siis `voikkoindex` määrittelee kolme indeksiä.

LaTeXissa on paketti `splitidx`, jolla voi tehdä useita indeksejä, esimerkiksi

```
Maija Virtaselle\index[nimet]{Virtanen, Maija} tuli
Helsingistä\index[paikat]{Helsinki}
kirjoja\index[sanat]{kirja}
Johan Wolfgang von
Goethesta.\index[nimet]{Goethe, Johan Wolfgang@von Goethe, Johan Wolfgang}
```

Typerä lause, mutta se on vain esimerkki.

LuaLaTeX-paketti voikkoindex muuttaa sanat perusmuotoon ennen niiden laittamista hakemistoon, joten lause voidaan kirjoittaa

```
\VXN{Maija}{Virtaselle} tuli
\VXP{Helsingistä}
\VXS{kirjoja}
\VXVON{Johan Wolfgang}{von}{Goethesta}[.]
```

Komennot

Voikkoindex määrittelee 17 komentoa

```
\newcommand{\vxn}[1]{\index[nimet]{#1}}
\newcommand{\vxp}[1]{\index[paikat]{#1}}
\newcommand{\vxs}[1]{\index[sanat]{#1}}
\newcommand{\vxvon}[3]{\vxn{#3, #1@#2 #3, #1}}
```

```
\NewDocumentCommand{\VXN}{m m o}
\NewDocumentCommand{\VXP}{m o}
\NewDocumentCommand{\VXS}{m o}
\NewDocumentCommand{\VXVON}{m m m o}
```

```
\NewDocumentCommand{\VXA}{m o}
\NewDocumentCommand{\VXI}{m o}
\NewDocumentCommand{\VXY}{m o}
\NewDocumentCommand{\VXL}{m o}
\NewDocumentCommand{\VXF}{m m o}
\NewDocumentCommand{\VXX}{m m m o}
```

```
\NewDocumentCommand{\SetExtraSurname}{m}  
\NewDocumentCommand{\SetExtraPlaceName}{m}  
\NewDocumentCommand{\SetExtraWord}{m}  
  
\NewDocumentCommand{\SetSurnameIndex}{m m}  
\NewDocumentCommand{\SetPlaceNameIndex}{m m}  
\NewDocumentCommand{\SetWordIndex}{m m}
```

Pienet komennot

Pienellä kirjoitetut komennot eivät muuta sanoja perusmuotoon, ja ne on määritelty vain kirjoittamistyön vähentämiseksi. Esimerkkilause voidaan kirjoittaa myös näin

```
Maija Virtaselle\vxn{Virtanen, Maija} tuli  
Helsingistä\vxp{Helsinki}  
kirjoja\vxS{kirja}  
Johan Wolfgang von  
Goethesta.\vxvon{Johan Wolfgang}{von}{Goethe}[.]
```

von ja muut sellaiset

Joissakin sukunimissä on osia, joita ei oteta huomioon aakkosjärjestyksessä. Niitä varten on komennot

```
\newcommand{\vxvon}[3]{\vxn{#3, #1@#2 #3, #1}}  
\NewDocumentCommand{\VXVON}{m m m o}
```

Ensimmäisessä parametrissa ovat etunimet, toisessa nimen osat joita ei oteta huomioon aakkosjärjestyksessä, ja kolmannessa sukunimi. Esimerkiksi

```
\VXVON{Magnus Gabriel}{de la}{Gardieta} muistellaan.  
\VXVON{Nils}{af}{Ursinin} nimi voidaan kirjoittaa näin.
```

Vain sukunimi muutetaan perusmuotoon ennen hakemistoon laittamista.

Välimerkit

Jos indeksoitava sana loppuu pisteeseen tai muuhun välimerkkiin, hakemistokomento on laitettava välimerkin jälkeen tai sanan ja välimerkin väliin tulee ylimääräistä tyhjää. Sen näkee helpoimmin y-kirjaimesta

```
\VXS{syksy}.  
\VXS{syksy}[.]
```

Tätä varten voikkoindeksin komennoissa on ehdollinen parametri, johon välimerkin voi laittaa.

Sukunimet ja paikannimet

```
\VXN{<etunimet>}{<sukunimi>}[<välimerkki>]  
\VXP{<paikannimi>}[<välimerkki>]
```

Etunimiä ei muuteta perusmuotoon, mutta sukunimi muutetaan.

Paikannimessä voi olla useampi kuin yksi sana. Tällöin vain viimeinen sana muutetaan perusmuotoon ja muut pysyvät ennallaan, esimerkiksi

```
\VXP{Etelä-Suomen läänissä} =>  
Etelä-Suomen läänissä\vxp{Etelä-Suomen lääni}
```

Asiasanat

```
\VXA{<asiasana(t)>}[<välimerkki>]  
\VXI{<asiasana(t)>}[<välimerkki>]  
\VXS{<asiasana(t)>}[<välimerkki>]  
\VXY{<asiasana(t)>}[<välimerkki>]
```

\VXA toimii kuten \VXS, mutta muuttaa kaikki kirjaimet pieniksi kirjaimiksi. A niin kuin alas.

\VXI toimii kuten \VXS, mutta muuttaa sanan ensimmäisen kirjaimen isoksi kirjaimeksi ja muut pieniksi. I niin kuin isolla.

\VXS muuttaa sanan perusmuotoon, mutta ei muuta kirjainten kokoa.

\VXY toimii kuten \VXS, mutta muuttaa kaikki kirjaimet isoiksi kirjaimiksi. Y niin kuin ylös

Jos asiasanassa on useampi kuin yksi sana, vain viimeinen muutetaan perusmuotoon, esimerkiksi

```
\VXS{auringon maassa}[,] =>
  auringon maassa[,]\vxs{auringon maa}
```

Mikäli tämä ei ole tarkoitus, käytä komentoa \VXF.

Listat

```
\VXF{<sanat>}{<formaatti>}[<välimerkki>]
```

Esimerkiksi

```
\VXF{Helsingin uudelle nuorisoseuralle}{p=s,s} =>
  Helsingin\vxp{Helsinki} uusi
  nuorisoseura\vxs{Helsingin uusi nuorisoseura}
```

```
\VXF{Helsingin seudun nuorisoseurassa}{p=,=,s}
  Helsingin\vxp{Helsinki} seudun
  nuorisoseurassa\vxs{Helsingin seudun nuorisoseura}
```

Ensimmäinen parametri on lista sanoja ja toinen parametri sanojen formaatti pilkuilla toisistaan erotettuina. Sallitut formaatit ovat

```
p=  =
ps  s
pa  a
pi  i
py  y
```

Jos formaatin ensimmäinen kirjain on p sana oletetaan paikannimeksi ja se laitetaan paikannimihakemistoon perusmuodossa. Muut formaatit ovat

= sanalle ei tehdä mitään
s sana muutetaan perusmuotoon
a perusmuoto muutetaan pieniksi kirjaimiksi (alas)
i perusmuodon eka kirjain muutetaan isoksi, muut pieniksi (isolla)
y perusmuoto muutetaan isoiksi kirjaimiksi (ylös)

Formaatin lopussa voi olla jompikumpi kirjaimista NL.

N perusmuodoksi hyväksytään vain nimisana tai nimilaatusana
L perusmuodoksi hyväksytään vain laatusana tai nimilaatusana

Tämä on tarpeen, koska eräät taivutusmuodot voivat olla sekä nimisanan että inen-loppuisen laatusanan taivutusmuotoja, esimerkiksi

talvista => talvi, talvinen

\VXL on sama asia kuin \VXF{<lista>}{p,s}[<välimerkki>]

mutta <lista>ssa voi olla vain kaksi sanaa. Esimerkiksi

\VXL{Helsingin nuorisoseuralle}[.] =>
Helsingin\vxp{Helsinki} nuorisoseuralle.\vxs{Helsingin nuorisoseura}

\VXX{<sanat>}{<formaatti>}{<lisäys>}[<välimerkki>]

Tämä toimii samoin kuin VXF, mutta <lisäys> lisätään indeksiin muun tekstin jälkeen. Esimerkiksi

\VXX{Helsingin uudelle nuorisoseuralle}{p,s,s}{1900} =>
Helsingin\vxp{Helsinki} uusi
nuorisoseura\vxs{Helsingin uusi nuorisoseura 1900}

Lisäys voi tietenkin olla muuta kuin vuosiluku.

Yksikkö ja monikko

Voikko muuttaa kaikki nominit yksikön nominatiiviin. Esimerkiksi Yhdysvalloista tulee Yhdysvalta. Tätä varten voikkoindexissä on kolme komentoa. Kaikilla on kaksi parametria, joista ensimmäinen on Voikon palauttama arvo tai voikkoindexiin lisätty sana (katso seuraava kohta), ja toinen on sana (jonka ei tarvitse olla monikko), joka halutaan laittaa hakemistoon. Esimerkiksi

```
\SetSurnameIndex{Wirtanen}{Virtanen}  
\SetPlaceNameIndex{Yhdysvalta}{Yhdysvallat}  
\SetWordIndex{nuorisoseura}{nuorisoseurat}
```

Voikosta puuttuvat sanat

Voikosta puuttuvat sanat voidaan lisätä voikkoindexiin tällä tavalla

```
\SetExtraSurname{Hyvämä|ki}  
\SetExtraPlaceName{Takahikiä}  
\SetExtraWord{teikar|i}
```

Koska sanat etsitään luettelosta siinä järjestyksessä, kun ne on siihen laitettu, jos lisättävä sana on sama kuin pitemmän sanan alkuosa, pitempi sana pitää olla ensin. Esimerkiksi

```
\SetExtraWord{qwertyuiop}  
\SetExtraWord{qwerty}
```

Muuten pitemmän sanan asemesta hakemistoon lisätään aina lyhyempi sana.

Koska nämä sanat eivät siis ole Voikossa, niitä ei voi muuttaa perusmuotoon, mutta jos hakemistoon lisättävä merkkijono alkaa samalla tavalla kuin jokin tällä tavalla lisätyistä sanoista, sana lisätään hakemistoon. Jos lisätyssä sanassa on pystyviiva, sanaa tutkitaan vain siihen asti. Esimeriksi kaikki **teikar**-alkuiset merkkijonot lisätään asiasanahakemistoon muodossa **teikari** ja kaikki **Hyvämä**-alkuiset

sukunimet nimihakemistoon muodossa Hyvämäki. (Mutta katso myös edellinen kohta.)

Tämäkin toimii

```
\SetExtraSurname{Wirta|nen}  
\SetSurnameIndex{Wirtanen}{Virtanen}
```

Myös sellaiset erisnimet pitää lisätä, joitten perusmuoto on sama kuin jonkun muun sanan taivutusmuoto, tai jos niitten taivutusmuodolla on useampi kuin yksi perusmuoto. Esimerkiksi Koski (koskea), Kemien (Kemi, Kemie).

Voikkoindex hyväksyy erisnimien perusmuodoiksi kaikki ne merkkijonot, joille se löytää perusmuodon, vaikka perusmuoto ei olisi sanastossa erisnimenä.

voikkoindex-mode.el

Emacsin paketti `voikkoindex-mode.el` määrittelee komennoille näppäinyhdistelmät

ESC M-n VXN
ESC M-p VXP
ESC M-s VXS
ESC M-a VXA
ESC M-i VXI
ESC M-y VXY
ESC M-l VXL
ESC M-f VXF
ESC M-v VXVON
ESC M-e (*)

(*) ESC M-e toimii kuten ESC M-n, mutta lisää komennon `\emph{}` sukunimen ympärille, ja jos sukunimen jäljessä on piste tai pilkku tai kaksoispiste, laittaa sen `\emph{}`:n sisään. Esimerkiksi

Matti Virtasesta, tulee `\VXN{Matti}{\emph{Virtasesta,}}`

Paketti asennetaan Linuxissa kopioimalla tiedosto `voikkoindex-mode.el` hakemistoon, josta emacs löytää sen ja lisäämällä tiedostoon `.emacs`

```
(autoload 'voikkoindex-mode "voikkoindex-mode"  
  "Toggle voikkoindex-mode." t)
```

Moodi otetaan käyttöön ja poistetaan käytöstä komennolla

M-x voikkoindex-mode

Näppäinkomennot toimivat siirtämällä kohdistin sanaan, joka halutaan lisätä hakemistoon ja antamalla jokin näppäinkomento. Esimerkiksi M-s muuttaa sanan **esimeriksi** muotoon `\VXS{esimeriksi}` ja niin edelleen.

VXN ja VXL lisäävät indeksiin kaksi sanaa ja VXVON kolme sanaa. Tällöin kohdistin täytyy laittaa ensimmäisen sanan päälle.

Mikäli (viimeisen) sanan lopussa on piste tai pilkku tai kaksoispiste, se laitetaan hakasulkujen sisään, esimerkiksi `[.]`

Komennolla VXF voidaan hakemistoon laittaa kuinka monta sanaa tahansa. Niinpä M-f vain lisää ensimmäisen sanan (jonka päällä kohdistin on) alkuun `\VXF{` eikä tee mitään muuta.

Mitä tarvitaan ja mistä ne saa?

- voikkoindex:
<https://github.com/ahomansikka/voikkoindex> Koska luet tätä tekstiä, olet jo imuroinut tämän. (-:
- corevoikko:
<https://github.com/voikko/corevoikko> Uusin versio. Sen pitää olla käännetty niin, että perusmuodot tulevat mukaan.
- lualibvoikko:
<https://github.com/ahomansikka/lualibvoikko>
- lualatex:
Tulee ainakin TeXLiven mukana.
- splitidx:
Tulee ainakin TeXLiven mukana.
- luarocks:
Se voidaan asentaa komennolla `sudo apt install luarocks`
- luautf8:
<https://github.com/starwing/luautf8> Se voidaan asentaa komennolla `sudo luarocks install luautf8`

Jos lualatex ei löydä luautf8:aa, pitää määritellä ympäristömuuttuja (esim. tiedostossa `.bashrc`), joka kertoo missä tiedosto `lua-utf8.so` on:

```
export CLUAINPUTS=/usr/local/lib/lua/5.3//
```

Sen jälkeen voikkoindex'ia voi käyttää komennolla `lualatex`, esimerkiksi

```
lualatex tiedosto.tex  
splitindex <parametrit> tiedosto.idx  
lualatex tiedosto.tex
```