# Huomioita Nykysuomen sanakirjan verkkoversiosta

Tähän muistioon olen koonnut erilaisia, osittain hyvinkin yksityiskohtaisia havaintoja Nykysuomen sanakirjan verkkoversiosta. Verkkosivullani <https://jkorpela.fi/nsk.html> esitän käytännöllisiä ohjeita sen käytöstä.

## Yleishuomioita

Nykysuomen sanakirjan (NS) julkaiseminen vapaasti luettavana verkkoversiona on erittäin tervetullutta ja aito kulttuuriteko. Se on myös toteutettu asiantuntevasti ja taitavasti. Kuvaskannauksen laatu on korkea, ja tulosta on paljon helpompi lukea kuin painettua kirjaa. Lisäksi sisältöön voidaan tehdä tekstihakuja, ja sitä voi kopioida.

Ilmeisesti julkaisu on toteutettu jonkinlaisella overlay-tekniikalla, jossa on ikään kuin kuvan päällä vastaava tekstisisältö. Tekstin skannaus vaikuttaa jokseenkin virheettömältä, mikä on varmaankin vaatinut työtä. Onhan monissa vanhojen aineistojen skannauksissa tulos varsin puutteellinen.

Koska skannausvirheitä kuitenkin todennäköisesti on, olisi toivottavaa, että Kotus tarjoaisi yksinkertaisen tavan esittää niistä huomiota ja korjausehdotuksia. Tämä voisi merkitä esimerkiksi lomaketta, jossa vain yksiköidään hakusana, kohta sen sisällä ja havaittu ongelma, mahdollisesti myös ratkaisu kuten virheen korjaus.

Tässä muistiossa käsitellään lähinnä erikoismerkkeihin ja välimerkkeihin liittyviä kysymyksiä. Niissäkin skannauksessa on pääosin onnistuttu hyvin. Lisäksi virheet tällaisissa asioissa eivät vaikuta tekstin hakemiseen ja sanojen kopiointiin. Niillä on kuitenkin merkitystä silloin, kun kirjasta halutaan ottaa lainaus leikkaa ja liimaa -menetelmällä.

## Koko-ongelma

Käytetystä tekniikasta johtunee, että tiedostot ovat hyvin suuria, yli puolen gigatavun kokoisia. Tästä tietysti aiheutuu ongelmia etenkin hitailla yhteyksillä. Sanakirjan saa kyllä käyttöön myös puhelimeen, mutta silloin kannattaa ladata tiedostot siihen lähiverkkoyhteyttä käyttäen. Tästä olisi ehkä hyvä huomauttaa verkkoversion esittelyssä.

Koko-ongelmaa havainnollistaa, että 1. osa on noin 577 megatavun kokoinen, kun taas pelkkänä tekstinä tallennettu sisältö mahtuu noin 4 megatavuun. Eräs mahdollisuus olisi, että Kotuksen sivuille laitettaisiin myös tekstiversiot (.txt-tiedostot). Niiden käyttökelpoisuus sellaisinaan olisi kuitenkin suppea, ja niiden täydentäminen muotoilluksi tekstiksi (rich text, esimerkiksi RTF- tai Word-muotoon) olisi hyvin suuri työ. Joissakin yhteyksissä, kuten runeberg.org-sivustossa (esimerkiksi *Tietosanakirja*, http://runeberg.org/tieto), on pyritty tekemään tällainen työ joukkoistettuna vapaaehtoisvoimin siten. Tässähän perusta (skannattu teksti ja kuvien laatu) on paljon parempi, eikä kyseessä olisi juurikaan ”proofreading”, enemmänkin ”richtexting”. Runeberg.org-aineistojen ”proofreading” on jäänyt melko vähäiseksi, mutta ehkäpä NS kiinnostaisi sisällönsä takia enemmän.

## Muotoilut

Käytetyn tekniikan takia sisältö on kopioitavissa lähinnä vain pelkkänä tekstinä ilman muotoiluja. Tämä on ongelmallista, koska muotoiluilla on suuri merkitys NS:ssa, kuten verkkoversion julkistamisesta kertovassa uutisessakin todetaan. Asia ei liene kohtuullisella työllä korjattavissa yleisesti, mutta ehkä joiltakin osin. Lisäksi siitä olisi hyvä kertoa verkkover­sion käyttöohjeessa.

Verkkoversion tekstissä on ilmeisesti käytetty lähinnä Free Serif ‑fonttia, osittain Times New Romania. Fonttikoko näyttää olevan enimmäkseen 5 pistettä, mikä voi aiheuttaa yllätyksen, kun tekstiä kopioi esimerkiksi Wordiin

Erityisesti seuraavat muotoilut häviävät tekstiä kopioitaessa:

* lihavoinnit
* kursivoinnit
* yläindeksityyli, jolla on esitetty taivutustyyppien numerot; ne voitaisiin esittää myös yläindeksimerkeillä ¹²³⁴⁵⁶⁷⁸⁹⁰ pelkässä tekstissäkin
* kapiteeli, joka on tekstin skannauksessa toteutunut pienaakkosina, eli esimerkiksi ”kianto” on asussa ”kianto”, mikä voi hämmentää; suuraakkoset olisivat sikäli käytännöllisempi valinta, että niistä saa ns. valekapiteelia fonttikokoa pienentämällä (kuten tässä on tehty)
* harvennus, jota esiintyy ohjeissa; on tärkeää, että harvennettu sana on tekstissä sanana eikä erillisinä kirjaimina, mutta samalla häviää se korostus, joka harventamiseen sisältyy.

## Hakemisen ongelma

Hakusanan etsiminen verkkoversiosta on hankalaa, koska hakusanat eivät tekstisisällössä eroa muusta tekstistä, vaikka ne painetussa kirjassa ja kuvaskannauksessa näkyvät lihavalla. Artikkelia, joka kertoo iili-sanasta, ei käytännössä voi etsiä hakutoiminnolla, koska se löytää niin suuren määrän merkkijonon ”iili” muita esiintymiä.

Lisäksi hakusanaa ei löydy lainkaan, jos etsitään itse sanalla, mutta tekstisisällössä sen sisällä on pystyviiva. Esimerkiksi etsittäessä sanalla ”aktivoida” ei löydy kyseistä hakusanaa, koska se on muodossa ”aktivo|ida”. Tähän ei taida olla hyvää ratkaisua, koska pystyviivojen poistaminen tekstisisällöstä hävittäisi olennaista informaatiota.

## Sanojen katkeaminen

Hakuja vaikeuttaa myö se, että sanan jatkuminen riviltä toiselle on joskus katkaissut sanan tekstissä. Tämä ilmenee siten, että kun rivin lopussa on NS:ssa sanan alku ja yhdysmerkki, verkkover­sion tekstissä on sanan alku ja rivinvaihto tai kaksi. On myös tapauksia, joissa rivin lopussa on negaation merkki ”¬” NOT SIGN U+00AC ja välilyönti. Joka tapauksessa sana pilkkoutuu. Jo ensimmäisessä hakusana-artikkelissa **a** on kaksi tällaista. Kun NS:ssa kahdelle riville jakautuvasta sanasta ”huuliaukko” tuleekin ”huuli¬ ” ja rivinvaihdon jälkeen ”ko”, ei haku sanalla ”huuliaukko” tietenkään löydä sitä. Verkkoversiossa on kuitenkin myös paljon tapauksia, joissa tällaista ongelmaa ei ole, vaan NS:ssa kahdelle riville jakautunut sana on tullut tekstiin yhtenä. Esimerkiksi artikkelissa **aaltoilla** on sana ”tyrskytä” oikein, vaikka se NS:ssa on jakautunut eri riveille. (Kyseessä ei liene käsin korjattu artikkeli, koska itse hakusanasta puuttuu pystyviiva ja 1. kohdan lopussa on outo virhe ”Meren a:uä” pro ”Meren a.”.) Artikkelin lopussa sana ”Kaapin” on jakautunut kahdelle riville, ja tekstiin on tullut ”Kaa”, kolme rivinvaihtoa ja ”pin”.

## Ajatusviiva ja muut viivat

NS käyttää välimerkkinä ajatusviivaa, joka on varsin pitkä ja ilmeisimmin tulkittavissa pitkäksi ajatusviivaksi, ”—” EM DASH U+2014 pikemminkin kuin lyhyeksi ajatusviivaksi ”–” EN DASH U+2013. Tämä vastaa painetun suomen perinnettä, joka alkoi siirtyä lyhyemmän ajatusviivan kannalle vasta pari vuosikymmentä NS:n julkaisemisen jälkeen. On tulkinnanvaraista, miten tässä pitäisi menetellä, mutta pitkän ajatusviivan korvaaminen yhdysmerkillä ”-” U+002D HYPHEN-MINUS vaikuttaa väärältä ratkaisulta. Verkkoversiossa on ilmeisesti alkupuolella, muun muassa alkusanoissa, käytetty pitkää ajatusviivaa, mutta myöhemmin yhdysmerkkiä.

Sen sijaan on luonnollista, että NS:n käyttämä ”- -” poiston merkkinä on säilytetty, vaikka se poikkeaa nykyisistä normeista, joiden mukaan on käytettävä ajatusviivoja ”– –”. Sama koskee yhdysmerkin käyttöä vaihteluvälin tms. ilmaisemiseen kuten alkusanojen alkupuolella ilmauksessa ”s. 354-”, joka nykysääntöjen mukaan kirjoitettaisiin ”s. 354–”. (NS on kyllä epäjohdonmukainen, sillä sivun VII lopussa esiintyy ”XII—XVIII”, jossa siis on pitkä ajatusviiva).

## Lainausmerkit ja heittomerkki

Tavallinen lainausmerkki on ainakin useassa kohdassa tullut tekstiin kahtena peräkkäisenä Ascii-heittomerkkinä (ns. koneheittomerkkinä eli pystysuorana heittomerkkinä) ' APOSTROPHE U+0027. Se tulisi tietysti esittää suomen kielen mukaisena lainausmerkkinä ” RIGHT DOUBLE QUOTATION MARK U+201D. (Kulmalainausmerkki », joka esiintyy ainakin alkusanoissa, on ilmeisesti tullut tekstiin oikein.)

Yksinkertainen lainausmerkki, jota käytetään myös heittomerkkinä, on tullut tekstiin osittain oikein merkkinä ’ RIGHT SINGLE QUOTATION MARK U+2019, mutta usein virheellisesti pystynä heittomerkkinä ' APOSTROPHE U+0027. Ensimmäinen tällainen esiintymä on sivulla VIII, jossa loppuaspiraation merkiksi selitettävä symboli on selvästi kaareva.

## Yläpiste

Sivulla VII kuvataan, että ”pääpaino merkitään painollisen vokaalin yläreunassa olevalla pisteellä”. Se on tekstissä rivinkeskinen piste ”·” MIDDLE DOT U+00B7, joka on monikäyttöinen merkki, mutta on tarkoitettu sijaitsemaan ja sijaitsee useimmissa fonteissa x-korkeuden puolessavälissä: a·. NS:ssa käytetty symboli sijaitsee, kuten sen teksti kuvaa, kuitenkin ”vokaalin yläreunassa”, joka tarkoittaa suunnilleen x-korkeuden tasoa. Siksi sopivampi merkki olisi kreikkalainen puolipiste ano teleia, ”·” ANO TELEIA U+0387: a·.

Ongelmaksi voidaan toki kokea, että ano telia kuuluu kreikan kirjoitusjärjestelmään ja on Unicode-standardissa lapsipuolen asemassa: se on määritelty kanonisesti ekvivalentiksi rivinkeskisen pisteen kanssa, ja standardi lisää: ”00B7 is the preferred character”. Kuitenkin ano telia sopii typografisesti paljon paremmin kreikan puolipisteeksi ja myös sellaiseksi ”vokaalin yläreunassa” olevaksi pisteeksi, jota NS:ssa käytetään.

## Yhteen ja erikseen kirjoittamisen merkit

Yhteen ja erikseen kirjoittamisen merkit kuvataan NS:n sivulla X, mutta tekstistä ne puuttuvat kokonaan, eli tekstiskannaus on tässä epäonnistunut. Sana-artikkeleissa ne ovat ilmeisesti ainakin yleensä merkkeinä ”∪” UNION U+222A ja ”∩” INTERSECTION U+2229. Tämä on epätyydyttävää sekä merkityksen (nämä ovat joukko-opin operaattoreita) että muodon kannalta (nämä ovat U:n ja käännetyn U:n muotoisia, eivät loivahkoja kaaria, kuten NS:ssa).

Unicodessa on useita muodoltaan jotenkin sopivia merkkejä, mutta parhaiten tähän kai sopisivat ”‿” UNDERTIE U+203F ja ”⁔” INVERTED UNDERTIE U+2054. Tosin ne sijaitsevat alempana (peruslinjan) alla.

## Painohistorian vaikutus

Edellä mainittu sivu sanoo: ”Sanakirjan alkuperäistä sisältöä ei ole muutettu yhdessäkään monista eri painoksista.” Tämä usein esitetty huomautus pitää paikkansa, mutta voi silti olla harhaanjohtava etenkin tässä yhteydessä.

Alkuperäisessä, vuosina 1951–1961 julkaistussa painoksessa, josta skannaus on ilmeisesti tehty, oli kunkin osan lopussa Oikaisuja-osasto ja viimeisen osan lopussa lisäksi Täydennyksiä-osasto. Alkuperäisessä kuusiosaisessa nideasussa kirja ilmeisesti julkaistiin vielä kahdesti, Fennican tietojen mukaan vuosina 1992 ja 2002. Jälkimmäisen julkaisun bibliografisissa tiedoissa on ”15. p.” Vuonna 1967 julkaistiin kolmena nidoksena ”lyhentämätön kansanpainos” ja siitä lisäpainos vuonna 1970. Ilmeisesti useita muitakin painoksia on ollut, vaikka niistä on Fennicassa tietoa hyvin sekavasti. Esimerkiksi tietue  
<https://kansalliskirjasto.finna.fi/Record/fikka.3515582>  
puhuu osasta 4, ”O-R”, mutta kuitenkin sanoo: ”Perustuu Nykysuomen sanakirja 1-3:n 1967 ilmestyneeseen 2. painokseen”.

Ilmeisesti kaikki painokset on tehty samoilla painolaatoilla, joihin kuitenkin tehtiin muutoksia vuoden 1967 kansanpainoksessa siinä mielessä, että alkuperäisen julkaisun osien lopussa esitetyt oikaisut tehtiin itse tekstiin (ja osien lopusta tietysti poistettiin Oikaisuja-osastot) ja lisäksi Täydennyksiä-osasto otettiin huomioon siinä mielessä, että sanakirjan sisältöön lisättiin reunukseen merkinnät, joilla osoitetaan, mihin kohtiin on täydennyksiä. Ilmeisesti kuusiniteiset painoksetkin tehtiin näin korjatuilla painolaatoilla.

Tässä päättelyssä on tietysti epävarmuuksia. Käytössäni oleva painos on kolminiteinen ja vuodelta 1976 ja merkitty viidenneksi painokseksi. Siitä näkee selvästi esimerkiksi **viikonloppu**-kappaleen ulkoasusta, että sen sisällä on tehty artikkeliin **-lippu** iso, Korjauksia-osastossa kuvattu muutos, joka on lyhentänyt tekstiä huomattavasti: kappaleen jälkeen on lähes rivin verran tyhjää tilaa, ja sen neljä viimeistä riviä ovat ladonnaltaan hiukan poikkeavia.

Oletettavasti painosten numerointi on tehty juoksevasti ottaen mukaan sekä kuusi- että kolminiteiset painokset. Olisi tavallaan kiinnostavaa nähdä koko julkaisuhistoria.

Käytännössä verkkoversion käyttäjälle olennaista on, että se perustuu alkuperäiseen julkaisuun ja siten ilmeisesti poikkeaa kaikista myöhemmistä: osien lopussa on suuri määrä korjauksia, jotka myöhemmin tehtiin tekstiin, eikä Täydennyksiä-osastoon ole viittauksia sisällössä. Tämä olisi varmaankin hyvä selostaa verkkoversion kuvauksessa.

Hankalampi kysymys on, pitäisikö korjaukset tehdä tekstiin. Niitä on kyllä paljon, mutta ei suunnatonta määrää, ja suurin osa on pieniä korjauksia. Se, että skannaus tehtiin alkuperäisestä julkaisusta, on ymmärrettävää ja kulttuurihistoriallisesti arvokasta, mutta käytettävyyttä lisäisi, jos korjaukset vietäisiin tekstiin sopivasti merkittyinä. Nyt tilanne on se, että jos lukija haluaa varmistua esimerkiksi siitä, että jokin 1. osassa oleva kohta on oikein (siinä mielessä, että siihen ei ole oikaisua), hänen on käytävä läpi kaikkien osien Oikaisuja-osastot. Lisäksi on tietysti mahdollista, että siihen on tullut täydennys.

## Tekstiskannauksen epäonnistumisia

Tekstin skannaus on epäonnistunut joillakin sivuilla niin, että **kaksipalstaisuus** on jäänyt huomioon ottamatta. Tämä tietysti aiheuttaa tekstisisällön sekoittumisen, ja sen huomaa esimerkiksi, kun yrittää valita jostakin palstasta tekstiä enemmän kuin yhden rivin: mukaan tulee väkisin toisen palstan tekstiä. Näin on käynyt ainakin seuraavilla sivuilla:

* 1. osan sivut 8, 37, 151, 228, 291, 306, 586, 616 ja 661
* 2. osan sivu 607
* 4. osan sivu 419 ja sivun 600 alku (P-kirjaimen loppu)
* 5. osan sivu 144.

(Sivunumerot ovat tässä painetun kirjan sivunumeroita, jotka eivät vastaa sivun järjestysnumeroa PDF-tidostossa.)

## Taivutuskaavojen ongelmia

Sivulla XII taivutuskaavan 12 kaksi viimeistä kohtaa puuttuvat. Kokonaisuudessaan kirjan alkuosassa olevat taivutuskaavat ovat niin ongelmallisia mm. taulukkorakenteensa takia, että olen laatinut niistä korjatut esitykset:

* <https://jkorpela.fi/nsk-nominit.html>
* <https://jkorpela.fi/nsk-verbit.html>

## Ongelma tekstinä tallentamisessa

Kun 1. osan tallentaa Adobe Acrobat Readerilla tekstimuodossa, jää sivu XIV (osa nominien taivutustaulukkoa) pois niin, että siitä näkyy vain ”to|si”. Voisiko tämä johtua siitä, että skannatulla sivulla tämä sana on vahvasti alleviivattuna? Kun tekstin maalaa kopioimista varten, mukaan tulee lähes koko sisältö, mutta pois jää osa teksteistä ilman mitään ilmeistä logiikkaa (esimerkiksi yksikön genetiivit tyypeistä 38 ja 39) ja ”to|si” tulee mukaan sekä omalla paikallaan että tekstin loppuun.

Samanlainen ilmiö on sivulla XVII, jolla on kaksi alleviivattua kohtaa.

Samalla tavoin käy 1. osan sivulle 54. Maalattaessa koko sisältö tulee mukaan, mutta loppuun tulevat lisäksi rivit ”subj:a” ja ”—”.

Ongelma koskee myös 4. osan sivuja 344 ja 345 (pippuriköynnös-sanan loppuosasta pirteä-sanan alkuosaan). Tekstiä voidaan kopioida niiltä maalaamalla, mutta tallennettuun tekstitiedostoon niistä tulee vain muutama merkki.

## Runojalat

Runojalkojen merkinnöissä on vaihtelua, joka ei johdu painotekstin asun vaihtelusta vaan siitä, että teksti on skannautunut eri tavoin:

* anapesti: ∪ ∪ -
* daktyyli: — ◡◡
* jambi: ∪ -
* spondee: --
* trokee: -∪

Tässä pitäisi varmaan olla johdonmukainen välien käytössä (ei välejä) ja tavujen merkinnöissä. Pitkä/painollinen tavu on aika luontevasti pitkä ajatusviiva ”—”. kuten nyt daktyylin merkinnässä). Lyhyt/painoton tavu on ongelmallisempi. Englanninkielinen Wikipedia käyttää lyhyymerkkiä U+02D8 BREVE ˘. joka on tavallaan semanttisesti sopiva, mutta ulkoasultaan sopimaton. Jos sen pariksi ottaisi pituusmerkin U+00AF MACRON, saataisiin johdonmukainen asu kuten ˘ ˘¯, mutta se olisi vaikealukuinen ja kaukana painoasusta. Daktyylin merkinnässä nyt esiintyvä U+25E1 LOWER HALF CIRCLE on ehkä sopivin. Ehkä daktyylin merkintä onkin korjattu käsin, koska kaikissa muissa on yhdysmerkki ja unionin merkki; siinä kuitenkin on välilyönti ensimmäisen tavun merkin jälkeen, mutta ei toisen.

## Tekstiskannauksen virheitä

### Virhetyyppejä

Tekstiskannauksessa eli tekstin tunnistamisessa kuvista on sattunut muun muassa seuraavanlaisia virheitä:

* Merkin osa on jäänyt pois, koska se on liian ohutviivainen. Esimerkiksi merkki ū on tunnistunut pelkäksi u:ksi, koska viiva on ohut. Joissakin tapauksissa merkin osaa ei edes näy skannatussa kuvassa, vaikka painotekstissä se kyllä on.
* Skannaus on tietysti tehty painetun kirjan tietystä kappaleesta. Siinä on jonkin verran käsin tehtyjä merkintöjä. jotka ovat vaikuttaneet häiritsevästi tekstin tunnistamisen.
* Erikoismerkit ovat usein tunnistuneet väärin. Esimerkiksi 1. osan sivun 232 toisen palstan alussa kolmannen rivin alussa on erikoinen symboli, jolla ei ole vastinetta tekstissä. Se on ilmausten *ennen kuin* ja *ennenkuin* välissä ja selvästikin tarkoittaa niiden vaihtoehtoisuutta. Se on vinossa oleva aaltoviiva, jonka voisi tulkita merkiksi ”~” TILDE U+00E7, mutta koska tämän merkin ulkoasu vaihtelee eikä sillä ole selvää merkitystä, sopivampi voisi olla ”∼” TILDE OPERATOR U+223C.
* Tavallisissakin merkeissä on sattunut virheitä. Esimerkiksi e-kirjain on voinut tunnistua c:ksi, ja iso Z on ilmeisesti useimmiten tunnistunut L:ksi.
* Murtolukumerkit kuten ½ ovat usein tunnistuneet väärin tai parhaimmillaan merkkiyhdistelmiksi kuten 1/2.

### Skannauksen väärin tunnistamia merkkejä ja merkkijonoja

|  |  |  |  |
| --- | --- | --- | --- |
| **hakusana** | **on** | **pitää olla** | **huomautuksia** |
| aarto | aārto | aarto |  |
| abitura | [u-] | [-ū-] |  |
| adverbiaali | adjektivin | adjektiivin |  |
|  | cellainen | sellainen |  |
| adverbiaalinen | advevbiaalina | adverbiaalina |  |
| adversatiivikonjunktio | ralousta nlmaiseva | rajoitusta ilmaiseva |  |
| ahdistaa | Yksipcrs. | Yksipers. | e tunnistunut c:ksi |
| aikatieto | Elämäkcrran | Elämäkerran | e tunnistunut c:ksi |
| aksentti | é, è, ) | é, è, ê) | puuttuu sirkumfleksi-e; skannattu kuva on tässä epäselvä, mutta sirkumfleksi mainitaan tekstissä, ja kansan­painoksessa merkki on selvä |
| alennus 1 | 41/2 | 4½ | tai ainakin 4 1/2 |
| alfa | (a) | (α) |  |
| alkumerkki 1 | .--.--.- | .—.—.— |  |
| apollinaris | [-ä-] | [-ā-] |  |
| artikla | sotaa:  t. | sota-a:t |  |
| aste 3 b | 231/ a:tta | 23½ a:tta | tai ainakin 23 1/2 a:tta |
| auki 4 | iltaauringon | ilta-auringon | rivinloppuinen ”-” tulkittava ”kovaksi” |
| avata 8 | šsakkipeli | Šakkipeli | outo skannausvirhe! |
| beeta | 9) | (β) |  |
| bilabiaali | 3 | β | ilmeisesti kursiivi-beeta |
| biljoona | 1.000.000.000.-  000. | 1.000.000.000.000. |  |
| bungalow | [bangalou] | [baŋgəlou] |  |
| cabaret | [kabarē] | [kabarē·] |  |
| delta 1 | (s) | (δ) |  |
| dyne | [dyne t. dyne] | [dyne t. dȳne] | Viiva ei näy kuvassa, mutta on painoasussa. |
| eksponentti | lausekkeessa a), | lausekkeessa aⁿ), |  |
| enempi 2 b | c | † |  |
| ennen 2 c | ennen kuin ennenkuin | ennen kuin ∼ ennenkuin | Ks. aik. huomautusta. |
| eräs 2 | -- kuten eräillä muilla - | — kuten eräillä muilla — |  |
| esperanto | Lamenhofin | Zamenhofin |  |
| fa | [fã] | [fā] |  |
| folklore | [-l̄re] | [-lōre] |  |
| franko | (4) | (¹) | yläindek­si tunnistunut väärin |
| gamma | (/) | (γ) |  |
| histologi | histo|loei | histo|logi |  |
| hätinä | ehtinyt. | ehtinyt, | pilkku pro piste |
|  | pävär. | päivär. |  |
| i 3 | - vI. | = √−1 | ”vinculum” ei esitettävissä tekstinä |
| ikkunalista | 5y5 | 5×5 | tai isompi kertomerkki |
| ilmakehäinen | - 191 | −191 | miinusmerkki, ei väliä |
| ilmatulppa | i0maa | ilmaa |  |
| itseisarvo | +tai - | + tai – | tässä kuten muutamassa muussakin kohdassa tekstisisältöön tullut ”-” pitäisi korjata miinusmerkiksi |
| jaa-ah | ã | ā |  |
| jalka 4 | merkitään ' | merkitään ′ | U+2032 PRIME |
| 1. juuri 6 a | (mer  ia (4l va-evtaa , tvuea 8 v  tenssi | (merkitään: ⁿ√a) tarkoittaa lukua, jonka n:s potenssi | epätasainen riviväli sekoittanut |
| kabinettikokoinen | 17x11 | 17×11 |  |
| kahdeksannes | 1g | ⅛ |  |
| kahdeksannesnuotti | , . | , 𝅘𝅥𝅮... | U+1D160 MUSICAL SYMBOL EIGHTH NOTE |
| kaksois-b | (|) | (♭♭) | tai U+1D12B MUSICAL SYMBOL DOUBLE FLAT |
| kaksoispiste 1 | s:ssä | §:ssä |  |
| kaksoisristi | (() | (×) | t. U+1D12A MUSICAL SYMBOL DOUBLE SHARP 𝄪 |
| 3. kappa | () | (ϰ) | KAPPA SYMBOL pikemmin kuin KAPPA κ |
| kertoma | 1-2-3. | 1⋅2⋅3. | DOT OPERATOR |
| kertomerkki | ( tai ). | (⋅, ×) | DOT OPERATOR, TIMES |
| keskiverto | A:B = B:C | A : B = B : C |  |
| kiintokuutiometri | k-m3 | k-m³ |  |
| kokonuotti | merkki, o. | merkki, 𝅝. | U+1D15D MUSICAL SYMBOL WHOLE NOTE |
| konvertoida | 51/ | 5½ | tai ainakin 5 1/2 |
| kopeekka | 1/1oo | 1/100 | typogr. murtoluku ei esitettävissä tekstinä |
| korvata 1 | Dvitamiinin | D-vitamiinin |  |
| kuin I 3 b | 5»3 = [on yhtä k.] 15. a2b | 5×3 = [on yhtä k.] 15. a²b |  |
| kuutio | 4 4 4 | 4 ⋅ 4 ⋅ 4 | DOT OPERATOR |
| kuutiodesimetri | dm3 | dm³ |  |
| kuutiojalka | dm3 | dm³ |  |
| kuutiometri | m3 | m³ |  |
| kuutiomillimetri | mm3 | mm³ |  |
| kuutiosenttimetri | cm3 | cm³ |  |
| kuutiosyli | 2x2»2 m = 8 m3. | 2×2×2 m = 8 m³ |  |
| kvadriljoona | (12) | 10²⁴ |  |
| kvintiljoona | 103p | 10³⁰ |  |
| laskea A II | tchdä | tehdä |  |
| laudanpää | kuin 2', leveys vähintään 5'' ja pituus 1/-51/''. | kuin 2″, leveys vähintään 5″ ja pituus ½—5½″. | ″ on U+2033 DOUBLE PRIME |
| logaritmi | a = b | ax = b | x:n pitäisi olla yläindeksi; U+02E3 MODIFIER LETTER SMALL X ˣ? |
|  | loga | logₐ | U+2090 LATIN SUBSCRIPT SMALL LETTER A |
| logaritmifunktio | y=a | y = ax | x:n pitäisi olla yläindeksi |
|  | loga | logₐ | U+2090 LATIN SUBSCRIPT SMALL LETTER A |
| logaritmikäyrä | loga | logₐ | U+2090 LATIN SUBSCRIPT SMALL LETTER A |
| lähipäivä | lähipäivä s | lähipäiv|ä s. | heikko painojälki |
| maissi 1 | Lea | Zea |  |
| maksaa 1 a | Paljonko n:t | Paljonko m:t |  |
| menoisa | Dvoäkin | Dvořákin |  |
| merkki I 2 a | Koiraan m. on 6 ja naaraan . | Koiraan m. on ♂ ja naaraan ♀. |  |
| mesotsooinen | n esotsooinen | mesotsooinen |  |
| miinus | ab. 0 | ab. ±0 |  |
| mikkeli | Mikon äivä | Mikon päivä |  |
| mikrofaradi | (lyh. ) li.000.000 taradia | (lyh. μF) 1/1.000.000 faradia |  |
| mikrogramma | 1/1.000 000 | 1/1.000.000 | ei ehkä skannausvirhe, vaan painovirhe, joka on korjattu kansanpainoksessa; tästä ei kuitenkaan ole oikaisua, vaikka artikkeliin on muu oikaisu |
| mikroni | on (lyh. ) 1l.000.000 | (lyh. μ) 1/1.000.000 |  |
|  | Yhld. | Yhd. |  |
| millimikroni | vars. valo  millimikronis (lyh. m|ut. )  opissa: | millimikroni s. (lyh. mμ t. μμ) vars. valo-opissa: | skannauksessa rivitkin sekoittuneet |
| minuutti 2 | merkitään . | 30° 15' | merkitään ′. | 30° 15′ | U+2032 PRIME ′ |
| minä III 2 a | keskeisim män | keskeisimmän | sana jakautunut eri sivuille |
| minä III 2 b | rikas tuttaa | rikastuttaa |  |
| molekyylikaava | HLO | H₂O |  |
| 2. myy | () | (μ) |  |
| neliö 3 | 2.2 | 2 ⋅ 2 | U+22C5 DOT OPERATOR |
| neljännesmaili | 1/4 | ¼ | tulkinnanvaraista, kumpi on parempi |
| neljännesnuotti | (merkitään ) | (merkitään ♩) | U+2669 QUARTER NOTE |
| niin A I a | taivio | talvio | kapiteeli-l tunnistunut i:ksi |
| normaalikoko | 15x10,5 | 15×10,5 | tai isompi kertomerkki |
| nupukivi | 10-15 15-25 | 10—15 × 15—25 | tai isompi kertomerkki |
| näkyä 1 b | Kuvasta n·y | Kuvasta n:y |  |
| näkö | Arvostella jtak äön  perusteella | Arvostella jtak näön  perusteella. | Kyseisessä palstassa (3. osa, s. 704, toinen palsta) paljon muitakin virheitä heikon painoasun takia. |
| o | о̄ | ō | on kyrillinen o ja yhdistyvä pituusmerkki |
| o | о̄ | ō | toisessa o-hakusanas­sa sama ongelma kuin yllä |
| ohjauselin 2 | Likenteen | Liikenteen |  |
| ohmi | [̄mi] | [ōmi] |  |
| ohminen | [ömi-] | [ōmi-] |  |
| ohoo | [oh̄·] | [ohō·] | ks. myös [Yläpiste](#_Yläpiste) |
| ohutlauta | 5/s-7/4 paksu ja vähintään 5' | ⅝—⅞″ paksu ja vähintään 5″ |  |
| oktantti | 1/s | ⅛ |  |
| olla | ole- ∪ ol- | ole- ∼ ol- |  |
| olla A I 2 | 7. š:ssä | 7. §:ssä | s. 52 1. palsta |
| olla C I 4 | vī | √1 | s. 54 2. palsta |
|  | a2-b2 | a² − b² |  |
| olla | 3039 | 30—39 | s. 56 2. palsta |
| omega | 5- | ō |  |
|  | (42, o) | (Ω, ω) |  |
| osa 1 d | 31/4 | 3¼ | tai ainakin 3 1/4 |
| ooho | [̄·ho· t. ōho·] | [̄ō·ho· t. ōho·] | ks. myös [Yläpiste](#_Yläpiste) |
| ottaa G I 5 | Pallo ot4i | Pallo otti |  |
| paperipuu | 1--4 | 1—4 |  |
| peitepiirros | jR | JR |  |
| pelti 1 | 41,5 | 4×1,5 | tai isompi kertomerkki |
| peninkulma | 1/Ls | 1/15 |  |
| 2. pii | t, | π, |  |
| porsaskausi | 21/E | 2½ |  |
| potenssi 2 a | an | am | m:n pitäisi olla yläindeksi |
| promille | 0/; | ‰; |  |
| prosentti | 61/2 | 6½ | tai ainakin 6 1/2 |
| prosit | [pr5-] | [prō-] |  |
| pukki 2 | (v) | (✓) | vaikea merkki, mutta ei kirjain |
| puolinuotti | (merkitaan 20 | (merkitään 𝅗𝅥) | U+1D15E MUSICAL SYMBOL HALF NOTE; ulkoasu FreeSerifissä kehno |
| puolioikea | P:a | P:ssaa |  |
| 1. pää 3 d | 9 x 14! | 9 × 14! |  |
| riboflaviini | B-vitamiini | B₂-vitamiini |  |
| richelieukirjonta | [rišelj-] | [rišeljȫ] | ö ja yhdistyvä pituusmerkki U+0304) |
| riippua 2 | 3x | 3x² |  |
| romaaninkirjoittaja | 7ola | Zola |  |
| ruhtinaspiispa | Liegen | Liègen |  |
| schwabach | [šväbah] | [švābah] |  |
| sekstiljoona | 10-6 | 10³⁶ |  |
| sekunti 2 | merkitään '. | 30 5' 24 | merkitään ″. | 30° 5′ 24″ |  |
| septiljoona | 1042 | 10⁴² |  |
| shortsit | šöt- | šōt- |  |
| si | [si] | [sī] |  |
| sibilantti | 2 | z | kursiivi-z ei ole tunnistunut oikein |
| sigma | (E, g) | (Σ, σ) |  |
| sirkumfleksi | nousevaalaskevaa | nousevaa-laskevaa | rivinloppuinen ”-” tulkittava ”kovaksi” |
|  | (4) | (^) |  |
|  | (-) | (~) |  |
| suoritustila | 3/, | ¾ |  |
| supistua | 3/5 | ⅗ | on oikein, mutta murtolukumerkkiäkin voisi käyttää |
| sääski | aides | aïdes |  |
| šakki | 8x8 | 8×8 | x-kirjaimen tilalle kertomerkki, tai ehkä 8✕8 (✕: MULTIPLICATION X U+2715) |
| taa | [tä'] | [tā’] |  |
| taistelu | Latopekin | Zatopekin |  |
| tauko 3 | 1/4--, 1/s--, 1/ie-t. | ¼--, ⅛ --, 1⁄16-t. | 1/16:n esityksessä tässä U+2044 FRACTION SLASH |
| thomas- | [tömas-] | [tōmas-] |  |
| tiamiini | B,-vitamiini | B₁-vitamiini |  |
| tiili | 27 x 13 7,5 cm | 27 × 13 × 7,5 cm |  |
| treema | poèta | poëta |  |
| triljoona | 1018 | 10¹⁸ |  |
| trinomi | a +- b+- c | a + b + c |  |
| tuoppi | 1/ | ½ | tai ainakin 1/2 |
| 1. tuuma | merkitään '' | merkitään ″ | U+2033 DOUBLE PRIME |
| uusaste | (6) | (ᶜ) | vaikea tulkita, lähinnä yläindeksi-c, ehkä U+1D9C MODIFIER LETTER SMALL C |
| valloittaa 2 b | tuikitaan | tulkitaan |  |
| valmistaa III | valmentaa]' | valmentaa} | tahra tulkittu merkiksi |
| viestimies | 4 100 | 4 × 100 |  |
| viestinhiihto | 4 x 10 | 4 × 10 |  |
| viestinuinti | 4 50 | 4 × 50 |  |
| viestinviejä | 4 400 | 4 × 400 |  |
| vuotinen 2 | kaksi, neljä vuotta | ’kaksi, neljä vuotta | rivinalkuinen heittomerkki jäänyt pois |
| yhyy | yhy | yhȳ | pituusmerkki lisättävä yhdistyvänä |
| ypsilon | (Y,v) | (Υ,υ) | kreikkalaiset kirjaimet |
| zambo | [9-] | [θ-] | theeta-kirjain on varmaan muuallakin tuottanut ongelmia, jos sitä esiintyy; artikkelissa thriller se kuitenkin on oikein; se ei liene NS:ssa kovin tavallinen; sitä ei edes mainita ääntämisen merkinnöissä sivulla VIII |
| zarathustralainen | Larathustran | Zarathustran | ilmeisesti versaali-Z on lähes aina tunnistunut L:ksi |
| zeoliitti | Leoliitit | Zeoliitit |  |
| zirkoni | L:n | Z:n |  |
| zirkoniumsilikaatti | L:a | Z:a |  |
| zooklorelli | L:t'eräissä | Z:t ’eräissä | myös sanaväli puuttuu |
| zoologinen | L. | Z. |  |
| zoonoosi | L:eja | Z:eja |  |
| zootominen | z. | Z. |  |
| zulu | z:t | Z:t |  |
| å | ä [ö] 1. kirjain, o-äänteen merkki skandinaavi-ängström7 s. fys. (lyh. A) aallonpituuden mittasissa nimissä. 2. lyh. fys. (A) àngström. yksikkö: 10-8 cm, Ängströmin yksikkö. | å [ō] 1. kirjain, o-äänteen merkki skandinaavisissa nimissä. 2. lyh. fys. (Å) ångström. | Å-kohdan kaksi hakusanaa sekoittuneet toisiinsa, koska palstoitusta ei ole tunnistettu |
| ångström |  | ångström7 s. fys. (lyh. Å) aallonpituuden yksikkö: 10⁻⁸ cm, Ångströmin yksikkö. | ks. ed. |
| ä | [ã] | [ǟ] | pituusmerkki lisättävä yhdistyvänä; pitäisi näkyä niin, että viiva on ä:n pisteiden yläpuolella. |
| ö | [ō] | [ȫ] |  |