

Kannanottopyyntö uudesta näppäimistökaavioehdotuksesta

4.11.2005

Tätä ehdotusta koskevat kannanotot pyydetään toimittamaan viimeistään 28.11.2005 osoitteeseen kotoistus@kotus.fi. Kannanotot käsitellään julkisina ja niihin vastataan kotoistuksen verkkosivuilla osoitteessa <http://kotoistus.fi>.

Ehdotus on kirjoitettu alunperin englanniksi. Se lähetettiin 31.10.2005 liitteenä (jota on päivitetty 2.11.2005) kriittiseen vastauksemme ruotsalaiseen näppäimistökaaviota koskevaan standardiehdotukseen (SIS-remiss 4027). Ehdotus on myös toimitettu kansainvälisen tietotekniikkateollisuuden edustajien (Apple, IBM, Microsoft) sekä muiden Pohjoismaiden standardointijärjestöjen ja saamelaisasioita käsittelevien ministeriöiden kommentoitavaksi.

Ehdotuksen on laatinut kotoistushankkeen näppäimistötyöryhmä sen ohjausryhmän alaisuudessa.

*

Ehdotettu uusi näppäimistökaavio lisää merkittävästi nykyisin yleisessä käytössä olevien SWE/FIN-näppäimistöjen toimintoja.

Vaikka toimintoja on lisätty, ei tavanomaisen (ts. pöytäkoneen) näppäimistön perustoimintojen käyttöön kohdistu häiritseviä muutoksia. Peruskirjaimet ovat entisillä paikoillaan, ja näppäinten lukumäärä ja niiden sijoittelu on sama kuin nykyisissä normaalikokoisissa qwerty-näppäimistöissä.

Ehdotetulla näppäimistökaaviolla on tarkoitus tuottaa UCS/Unicode-koodattuja merkkejä. Mikäli on tarve käyttää rajoitettavampaa koodausta, on myös näppäimistön toiminnallisuutta rajoitettava.

Kaavio on rakenteeltaan avoin. Se perustuu lisäkirjainten ja -tarkkeiden sijoitteluun nykyisissä näppäimistöissä esiintyvien merkkien yhteyteen helposti mielletävällä tavalla. Käyttäjän on näin mahdollisimman helppo oppia sen käyttö omiin tarpeisiinsa.

Käyttäjät

Ehdotetun näppäimistökaavion mahdolliseksi käyttäjiksi on kaavailtu valtakielisiä puhuvia väestöyhtäältä Suomessa ja Ruotsissa ja toisaalta Norjassa ja mahdollisesti myös Tanskassa. Norjaa ja Tanskaa varten siihen tarvitaan muutoksia vain kahteen näppäimeen suomalais-ruotsalaiseen näppäimistöön verrattuna (C10 ja C11, kuten myöhemmin esitetään).

Perustavoitteet

Näppäimistökaaviolla on tarkoitus toteuttaa kolme tärkeää perustavoitetta.

Ensinnäkin sillä on helppo syöttää valtakielten, ts. suomen-, ruotsin-, norjan- ja tanskankielistä tekstiä käyttäen jo entuudestaan tuttua näppäimistökaaviota.

Toiseksi sillä voi vaivattomasti syöttää virallisten alueellisten ja vähemmistökielten tekstiä käyttäen intuitiivisia, helposti omaksuttavia keinoja. Tämän tasoinen tuki kattaa myös ne kielet, joissa esiintyy valtakielistä poikkeavia merkkejä: saamenkielet (pohjois-, etelä-, luulajan-, inarin- ja

koltansaame) Suomessa, Ruotsissa ja Norjassa, romanikielen Suomessa sekä fäärin ja kalaallistutin eli grönlannin Tanskassa.

Kolmantena tavoitteena on, että sillä voi suhteellisen helposti kirjoittaa erityisesti nimiä - kuten ihmisten, paikkojen tai tuotteiden nimiä, joissa esiintyy muissa kielissä käytettyjä latinalaisia merkkejä. Tuettuja kieliä ovat erityisesti EU:n viralliset kielet (paitsi kreikka) ja jäljellä oleva pohjoismainen kieli, islanti.

Suunnittelun lähtökohtana on ollut niin sanottu eurooppalainen merkkivalikoima MES-2 (Multilingual European Subset 2 of the UCS/Unicode). Tavoitteena on mahdollistaa siihen sisältyvien, nykyaikaisissa teksteissä käytettyjen latinalaisten kirjainten tuottaminen. Näiden lisäksi on otettu huomioon vietnamilaisien nimien oikeinkirjoitustarpeet.

Keinot

Ne tarvittavat lisäkirjaimet, joita ei voi muodostaa eri merkkien yhdistelminä, on sijoitettu lisämerkkeinä intuitiivisesti tunnistettaviin näppäimiin. Tätä silmälläpitäen tuotetaan lähes kaikki ne merkit, jotka ovat UCS/Unicode-mielessä yhdistelmiä, systemaattisesti käyttäen tarkenäppäimiä peruskirjainten yhteydessä. Tarkkeellisista kirjaimista ovat å, ä ja ö kuitenkin luonnollisesti säilyttäneet nykyiset paikkansa. (Kirjaimet ä ja ö kuitenkin vaihtavat Norjassa paikkaa näppäimessä niiden vastineiden æ:n ja ø:n kanssa ja Tanskassa myös keskenään.)

On tärkeää, että myös tuettujen kielten oikeinkirjoitussääntöjen mukaiset välimerkit ovat käytettävissä. Nykyisin useat sovellukset muuntavat näppäimistön perus-ASCII-lainausmerkit ja -yhdyserkin (tavuviiva-miinusmerkin) enemmän tai vähemmän oikein kotoistetuiksi välimerkeiksi. Nämä menetelmät tulevat tietysti edelleen olevan käytettävissä, mutta näppäimistöön nyt lisättävillä uusilla välimerkeillä voidaan tuottaa halutut merkit käytetystä sovelluksesta ja sen kieliasetuksista riippumatta. Näin täydennettynä näppäimistö tukee huomattavasti paremmin välimerkkien osalta pohjoismaisten kielten oikeinkirjoitusta. Näppäimistöön lisätään myös yhdistävä välilyönti (ja toteutuksen salliessa myös kapea yhdistävä välilyönti), jota käyttäen voidaan helposti ohjata tekstiä tulostettaessa syntyvät rivinvaihdot pois epätoivotuista kohdista.

Edellä olevan perusteella ehdotusta voitaisiin perustellusti kutsua pohjoismaiseksi kansainväliseksi ”Nordic International” -näppäimistökaavioksi. Sen merkkivalikoima ja muut ominaisuudet ylittävät nykyisen ruotsalaisen ehdotuksen (SIS remiss 4027) mukaisen ”Swedish International” -kaavion, mutta sen suunnittelun lähtökohta on täysin erilainen.

Yhdistelmät

Nykyisessä näppäimistökaaviossa on kaksi erityistä tarkenäppäintä. Ehdotuksessa näihin näppäimiin on lisätty tarkkeita eli diakriittisiä merkkejä. Lisäksi yksi on siirretty samassa näppäimessä uuteen paikkaan: suomessa ja useissa alueellisissa ja vähemmistökielissä käytetyn, näppäimistöön lisättävän hattu-tarkkeen ˇ paikaksi ehdotetaan nykyistä treeman ¨ paikkaa. Muutoksen varjopuolena tulee melko yleisessä käytössä olevan ü-kirjaimen tuottaminen jonkin verran nykyistä työläemmäksi.

Alkuperäisten tarkenäppäinten lisäksi on useaan muuhun näppäimeen lisätty tarkeominaisuus, kun niitä käytetään yhdessä Alt Gr -näppäimen (ja mahdollisesti myös Shift-näppäimen) kanssa. Esimerkiksi pilkku-näppäimestä tulee Alt Gr -näppäintä käytettäessä sediljin tuottava tarkenäppäin. Tässä dokumentissa termiä ”tarkenäppäin” käytetään kuvaamaan sekä varsinaisten tarkenäppäinten toimintaa että muihin näppäimiin lisättyjä tarkkeenomaisia toimintoja, joille on yhteistä se, ettei

”tarke” sinänsä, so. yksinään, tuota mitään näkyvää merkkiä, vaan tämä syntyy vasta sitä seuraavan näppäimen painalluksen jälkeen.

Luonnollinen toteutustapa on, että tarkenäppäin tuottaa loogisesti UCS/Unicoden mukaisen yhdistyvän tarkkeen, joka seuraa tarkenäppäimen jälkeen syötettyä merkkiä. Niinpä, ylemmän tarkenäppäimen (kaiverrukset: ´ ja `) ja kirjainnäppäimen painallus tuottaa kyseisen kirjaimen, jota seuraa yhdistyvä akuutti-korkomerkki-tarke (0301). Näppäimistöajurin kuitenkin tulisi synnyttää tästä valmis yhdistemerkki, kuten á (00E1) tai é (00E9), aina kun se on mahdollista, ts. silloin kun tällainen merkki sisältyy Unicode-standardin versioon 3.0. Tästä ei seuraa mitään muutosta nykyiseen ISO Latin1 -merkkien käsittelytapaan, vaan se laajennetaan määritellyksi yleiskäytännöksi. (Nimenomaisen Unicode 3.0 -viittauksen tarkoituksena on tehdä määrittelystä samalla tavalla vakaa kuin Unicoden normalisointiosäännöt ovat).

Määritykset sallivat myös usean eri tarkkeen liittämisen samaan peruskirjaimeen. Niinpä esimerkiksi ¨ (treema) + ˇ (hattu) + u-kirjain tuottavat merkin ů (01DA valmiina yhdistelmämerkinä). Tätä ominaisuutta tarvitaan esimerkiksi balttilaisten ja vietnamlaisien nimien oikeinkirjoitukseen.

Koska saman tarkkeen yhdistäminen merkkiin kahdesti ei ole mielekästä, voidaan yhdistelmälle, jossa tarkenäppäintä painetaan kahdesti, sopia erikoismerkitys:

Akuutti-korkomerkin ´ (0301) tarkenäppäimen painaminen kahdesti peräkkäin aiheuttaa yhdistyvän kaksoisakuutti-korkomerkin ¨ (030B) lisäämisen seuraavan merkin jälkeen. Tätä näin synnytettyä tarketta käsitellään sen jälkeen muiden tarkkeiden tapaan. Esimerkiksi kun yhdistyvää akuutti-korkomerkkiä painetaan kahdesti ja sen jälkeen kirjainta u, tuloksena on ů (0171), jota käytetään unkarin kielessä.

Vastaavasti gravis-korkomerkin ` (0300) tarkenäppäimen painallus kahdesti voisi tuottaa yhdistyvän kaksoisgravis-korkomerkki-tarkkeen (030F). UCS/Unicodessa on määritelty useita yhdistyviä kaksoistarkkeita, jotka voitaisiin tuottaa samalla periaatteella, mutta niiden kaikkien tukemiselle ei meillä ole mitään tarvetta.

Erityistoimintona voitaisiin myös määritellä, että yhdistyvän alapuolisen pisteen (0323) tarkenäppäimen painallus kahdesti tuottaisi rivinkeskisen pisteen seuraavasti: Kun näitä seuraisi välilyönti, syntyisi erillinen rivinkeskinen piste · (00B7). Kun näitä seuraisi kirjain l tai L, syntyisi ł (0140) tai Ł (013F), jotka eivät sinänsä ole yhdistelmämerkkejä.

Yhdistelmämerkeille toteutettua periaatetta voi soveltaa laajemminkin, On joukko merkkejä, jotka eivät ole UCS/Unicode-mielessä yhdistelmämerkkejä, mutta joiden voidaan katsoa koostuvan latinalaisesta perusmerkistä, jonka yli on vedetty poikkiviiva, kuten đ. Näissä merkeissä poikkiviiva ei ole tarke, mutta ehdotuksen mukaan ne synnytetään käyttäen erityistä poikkiviivanlisäysnäppäintä. Vaikkakaan poikkiviiva ei ole erotettavissa merkistä, näppäimistöajuri voi käsitellä lisäysnäppäintä paljolti samalla tavoin kuin tarkenäppäimiä. Näiden merkkien syöttö on näin helppoa ja luonnollista, kun käyttäjä on oppinut niihin sovellettavan yksinkertaisen periaatteen. Ehdotettu näppäimen kaiverrus puolestaan viittaa tähän periaatteeseen ja muistuttaa siitä.

Jotta monien poikkiviivallisten, erityisesti saamen kielissä käytettyjen kirjainten kirjoittamisen olisi mahdollisimman helppoa, poikkiviivan lisäystoiminto on sijoitettu näppäimen perustoiminnoksi. Tästä aiheutuu välttämättä muutos nykyiseen näppäimistöön verrattuna. Sekä muutoksen vaikutuksen minimoimiseksi että uuden toiminnon havaitsemisen helpottamiseksi toiminto on sijoitettu E00-näppäimeen (niissä näppäimistöissä, joissa se on käytettävissä, ts. valtaosassa). Tämän näppäimen perustoimintona ollut pykälämerkki § on puolestaan siirretty saman näppäimen Alt Gr -positioon, mistä aiheutuu jonkin verran epämukavuutta kirjoitettaessa lakitekstiä.

Tämän toteutuksen ansiosta voidaan kaikki nämä saamen kielten kirjaimet kirjoittaa kahdella näppäimen painalluksella: poikkiviivanlisäysnäppäin ja itse kirjainnäppäin. Merkit, jotka näin synnytetään, ovat đ - d ja poikkiviiva (0111, pohjois-, inarin- ja koltansaamassa), g - g ja poikkiviiva (01E5, koltansaamassa) sekä t - t ja poikkiviiva (0167, pohjoissaamassa). Poikkiviivaa käytetään myös tuottamaan esimerkiksi ł - l ja vino poikkiviiva (0142, puolan kielessä) tai h - h ja poikkiviiva (0127, maltan kielessä). Koska poikkiviiva ei ole erotettavissa oleva merkin osa, kaikki lisäysnäppäimen ja peruskirjaimen yhdistelmät, joista muodostuu kirjain, voidaan määritellä erikseen. Loogisesti yksinkertaisinta olisi käyttää tätä menettelyä kaikkien niiden kirjainten tuottamiseen, joiden normatiivisessa UCS/Unicode-nimessä esiintyy ”stroke” (= poikkiviiva), mutta meillä ei ole tarvetta tukea kaikkia niitä.)

Jos poikkiviivanlisäysnäppäintä käytettäisiin missään muussa kuin hyväksyttävässä yhdistelmässä, sillä ei olisi vaikutusta. Vaihtoehtoisesti voitaisiin määritellä, että silloin syntyisi pykälämerkki §. Tästä seuraisi, että lähes kaikissa tilanteissa, joissa pykälämerkki § haluttaisiin tuottaa, se syntyisi entiseen tapaan, joskin se tulisi näyttöön vasta sitä seuraavan näppäimen painalluksen jälkeen.

Muutokset

Ehdotuksesta aiheutuu kaksi muutosta nykyisin käytössä olevaan näppäimistökaavioon:

1. Näppäimessä E00 on pykälämerkki § siirretty Alt Gr -positioon, poikkiviivallisten merkkien syötön helpottamiseksi. (Ellei näppäin E00 lainkaan sisälly näppäimistöön, työläämpi menettely poikkiviivan lisäykselle toteutetaan näppäimessä C09).

2. Näppäimessä D12 on yhdistyvä treema-tarke siirretty Alt Gr + Shift -positioon hatullisten kirjainten syötön helpottamiseksi.

Jälkimmäisen muutoksen tarkoituksena on helpottaa erityisesti suomen kielessä käytettyjen š- ja ž- kirjainten sekä useiden muiden saamen kielissä ja romanikielessä käytettyjen kirjainten syöttöä. Näitä ovat č eli hattu-c (010D, pohjois-, inarin- ja koltansaamassa), ģ eli hattu-g (01E7, koltansaamassa), ħ eli hattu-h (021F, Suomen romanikielessä), ģ eli hattu-k (01E9, koltansaamassa) ja ž eli hattu-ezh (01EF, koltansaamassa). Kirjaimista, joiden osana on treema, ä ja ö esiintyvät luonnollisesti näppäimistössä, ja ü on selvästi yleisin muista jäljellä olevista kirjaimista pohjoismaisilla kielillä kirjoitetuissa teksteissä, kun taas ë ja ï ovat harvinaisia ja ý on erittäin harvinainen.

Ehdotus myös muuttaa joidenkin näppäilyjaksojen tulosta, mistä ei kuitenkaan aiheutune käytännössä mainittavia ongelmia. Kun nykyisin painetaan peräkkäin esimerkiksi sirkumfleksi-tarkenäppäintä ja c-kirjainnäppäintä, tuloksena on ^c, kun siitä ehdotuksen mukaan syntyy ĉ. Merkkijonon ^c synnyttämiseksi tulee tarkenäppäimen painalluksen jälkeen painaa välilyöntiä, kuten useimmat käyttäjät yleensä tekevät nykyisinkin.

Lisämerkit ja -tarkeet on esitetty liitteen 1 taulukossa ja liitteen 2 kaaviokuvassa.

Rakenteen avoimuus

Ehdotettu näppäimistökaavio on rakenteeltaan avoin. Ehdotuksessa on suoraan määritelty vain olennaisin osa tuetusta merkkivalikoimasta.

Nykyisen UCS/Unicode-käytännön mukaisesti ei ehdotuksen toteuttava näppäimistöajuri voi tarkistaa, onko jokin perusmerkin ja tarkkeen tai tarkkeiden yhdistelmä käyttäjän tarkoitukseen soveltuva. Tarkistukset jäävät sovelluksen ja käyttäjän hoidettaviksi. Kun käyttäjä kirjoittaa epätavallisia merkkejä, hänen tulee katsoa, syntyykö oikeannäköinen tulos. Sovelluksen tulisi tarkistaa, ettei se ota käsiteltäväkseen muita kuin hyväksytyjä merkkejä. Esimerkiksi väestökisterissä käytettävä merkkivalikoima on rajoitettu (joskaan siihen ei tulisi kohdistua läheskään nykyisen kaltaisia rajoituksia).

Suunnitteluperiaatteiden johdonmukaisen noudattamisen vuoksi muita yhdistelmämerkkejä kuin å, ä ja ö ei ole sijoitettu näppäimistökaavioon. Joitakin tällaisia merkkejä olisi kuitenkin mahdollista sijoittaa vaarantamatta rakenteen avoimuutta, mikäli tähän todetaan olevan suurta tarvetta. Tällaisena poikkeuksena voisi tulla kyseeseen ü-kirjain.

Liitteessä 1 esitetään taulukossa näppäimistön toiminnot, uudet nykyisiä tarkemmin.

Liitteessä 2 esitetään kaaviokuva merkkien sijoittelusta näppäimiin. Tarkoitus ei ole, että kaikki esitetyt merkit kaiverrettaisiin näppäimiin.

Liite 1:

Eri näppäinten toiminnot:

<i>Pos.</i>	<i>Perus-toiminto</i>	<i>+ Shift</i>	<i>Alt Gr -toiminto</i>	<i>Alt Gr + Shift</i>	<i>Huomautuksia</i>
E00	Poikki-viivan lisäys	½	§		Toiminnon siirto + lisätty lisäystoiminto. Ellei E00 käytössä, ks. C09!
E01	1	!			
E02	2	"	@	” (201D)	Lisätty lainausmerkki.
E03	3	#	£	» (00BB)	Lisätty lainausmerkki.
E04	4	¤	\$	« (00AB)	Lisätty lainausmerkki (norjaa ja tanskaa varten).
E05	5	%		‰ (2030)	Lisätty promillen merkki.
E06	6	&			
E07	7	/	{		
E08	8	([
E09	9)]		
E10	0	=	}	° (00B0)	Lisätty asteen merkki.
E11	+	?	\		
E12	Yt: ´ akuutti-korkom.	Yt: ` gravis-korkomerkki	Yt: ´ sarvi (031B)	Yt: ´ yläp. koukku (0309)	Lisättyjä yhdistäviä tarkkeita [Yt] (erityisesti vietnamia varten).
D01	q	Q			
D02	w	W			
D03	e	E	€		
D04	r	R			
D05	t	T	þ (00FE)	Þ (00DE)	Lisätty thorn-kirjain.
D06	y	Y			
D07	u	U			
D08	i	I	ı (0131)	İ (0130)	Lisätyt i:n pistevaihtoehdot.
D09	o	O	œ (0153)	Œ (0152)	Lisätty oe-ligatuuri.
D10	p	P			
D11	å	Å	Yt: ° yläp. ympyrä (030A)	Yt: ˇ lyhyysmerkki (0306)	Lisättyjä tarkkeita.
D12	Yt: ˇ hattu (030C)	Yt: ^ sirkumfleksi	Yt: ~ tilde	Yt: ¨ treema	Toiminnon siirto + lisätty tarke.
C01	a	A	ə (0259)	Ə (018F)	Lisätty švaa-kirjain.
C02	s	S	ß (00DF)		Lisätty kaksois-s-kirjain.
C03	d	D	đ (00F0)	Đ (00D0)	Lisätty eth-kirjain.
C04	f	F			
C05	g	G			
C06	h	H			
C07	j	J			
C08	k	K	κ (0138)		Lisätty kra-kirjain.

C09	l	L			Ellei E00 mukana näppäimistöissä, vaihtoehtoinen sijoitus poikkiviivanlisäystoiminnolle Alt Gr -positiossa.
C10	ö	Ö	ø (00F8)	Ø (00D8)	Lisätty tanskalainen ö.
C11	ä	Ä	æ (00E6)	Æ (00C6)	Lisätty tanskalainen ä.
C12	'	*	' (2019)	Yt: ´, alap. pilkku (0326)	Lisätty lainausmerkki + tarke.
B01	<	>			
B02	z	Z	Ʒ (0292)	Ʒ (01B7)	Lisätty ezh-kirjain.
B03	x	X			
B04	c	C			
B05	v	V			
B06	b	B			
B07	n	N	ŋ (014B)	Ŋ (014A)	Lisätty äng-kirjain.
B08	m	M	μ (00B5)		Lisätty mikro-merkki.
B09	,	;	Yt: ´, sedilji (0327)	Yt: ¸ ogonek(0328)	Lisättyjä tarkkeita.
B10	.	:	Yt: ´, alap. piste (0323)	Cd: ¨ yläp. piste (0307)	Lisättyjä tarkkeita + erityistoiminto rivinkeskiselle pisteelle.
B11	-	–	– (2013)	Yt: ¯ pituusm. (0304)	Lisätty lyhyt ajatusviiva + tarke.
A	Välilyönti				Myös yhdistävä välilyönti (NBSP) ja mahdollisesti kapea yhdistävä välilyönti; toteutus avoin.

Näppäimistökaaviossa on tarkoituksella jätetty useita mahdollisia positioita käyttämättä, vaikka näihin voitaisiin esittää hyvinkin luonnollisia sijoituksia. Lisäsjoitusten tarve vaihtelee käyttäjän, tekstin luonteen ja käytetyn ohjelman mukaan. Ne onkin katsottu parhaaksi jättää avoimiksi peruskaaviossa, jotta eri käytännöille jää tarpeeksi toteutustilaa.