

Conceptes clau per a elaborar la proposta de restitució tipogràfica de la Ela Geminada a Unicode.org

28.11.2007

Tots els conceptes exposats, enumerats de moment sense cap intencionalitat, utilitzaran el símbol [+] per descriure els elements que creiem que ens van a favor i el símbol [-] pels que creiem que ens van en contra.

1.

La Ela Geminada <L·L> és un signe (*glif*) únicament utilitzat al món per la llengua catalana.

[+] Els codis assignats finalment a aquest signe (caixa alta i caixa baixa) no entraran mai en conflicte amb cap altra llengua o ús quan esdevingui configurable en un teclat.

[+] Cal documentar històricament i filològicament el naixement i l'ús d'aquest signe.

[-] No creiem que hi hagi elements en contra per aquest concepte.

2.

Quan anomenem la Ela Geminada <L·L> hem de parlar de *signe* o utilitzar el tecnicisme *glif*, perquè la paraula *caràcter* i encara més *lletra* entren directament en conflicte amb les accepcions que la comunitat lingüística té per a les regles ortogràfiques.

[+] Necessitem tenir la majoria de la comunitat acadèmica de la llengua i dels estaments relacionats (Institut d'Estudis Catalans, Departament de Política Lingüística de la Generalitat, universitats...) a favor d'aquesta proposta i com a avals davant Unicode?

[-] El fet que en realitat la Ela Geminada no sigui una lletra de l'abecedari ens pot perjudicar?

[-] L'accepció en el món de la informàtica de la paraula *caràcter* és molt àmplia (malgrat que a molts lingüistes no els agrada). No utilitzar-la en el redactat pot quedar una mica extravagant? (Sempre ens queda la paraula *signe*, més neutral i molt utilitzada per Fabra.)

3.

Hauríem de saber si existeix cap cas paral·lel en alguna altra llengua que comparteixi les característiques especials de la Ela Geminada:

* **No és una lletra de l'abecedari.**

* **Tríglif (tres elements descomponibles en altres signes ja existents).**

* **Necessita integrar-se en un sol signe per estar ben composta (la composició separada dels tres elements no ens donarà mai una correcta construcció del *glif*).**

* **Comportament especial en la divisió de paraules (els tres elements 'l·l' es converteixen en dos 'l·l'—el punt volat desapareix— quan es parteix a final de línia).**

* **La composició de forma separada (ela + punt volat + ela) fa que la gent tendeixi a alterar-ne l'escriptura, bàsicament escrivint-la d'una d'aquestes tres maneres:**

a) **ela + punt + ela**

b) **ela + guió + ela**

c) **ela + *topo* + ela**

[+] Si realment no existeix cap cas paral·lel comparable, això ens fa singulars. Ens convé, però, apostar per la tesi basada en el fet que som una excepció a tenir en compte?

[+] Pel que fa al comportament en la divisió de paraules, ja tenim, amb l'alemanya *Eszett* (U+C39F), un cas paral·lel ben clar: la 'ß' passa a 's-s' a final de línia.

[-] Si realment existeix un cas paral·lel comparable, cal saber si la solució de codificació que Unicode li ha donat ens perjudica.

4.

És important que els *glifs* (signes) Ldot <L·> i ldot <l·> s'hagin començat a declarar obsolets (Unicode 5.1 Beta), perquè no s'han fet servir gairebé mai en composició digital i la seva codificació ha estat un malentès provocat per unes determinades circumstàncies històriques relacionades amb l'homologació de teclats i la màquina d'escriure elèctrica de margarida.

[+] Podem demostrar històricament que això ha estat així (la Ldot/ldot mai no va ser fabricada en linotípia o en tipus mòbils!)

[-] Com justifiquem aquesta badada? (Cal fer-ho?)

5.

Quan un signe ha estat declarat obsolet per Unicode, què passa amb el seu codi? Nosaltres només demanem la reutilització dels codis corresponents als signes que poden ser declarats definitivament obsolets <L·> i <l·>, o sigui:

U+013F

<http://unicode.org/cldr/utility/character.jsp?a=013F>

U+0140

<http://unicode.org/cldr/utility/character.jsp?a=0140>

[+] Podem trobar algun exemple paral·lel de reutilització del codi d'un *glif* obsolet?

[+] Llavors, quin sentit té que ens rebutgin aquesta proposta si sempre podem ser presents utilitzant el rang <U+E000> fins a <U+F8FF> de codis de lliure disposició?

Vegeu cas <LL> i <ll> amb Adobe Systems i l'ús de la *Private Use Area* dins la seva llista de xifrat:

<http://partners.adobe.com/public/developer/en/opentype/glyphlist.txt>

[+] Això no es pot dir, però qui són Google, Apple, Adobe o Microsoft per decidir com hem d'escriure el signes de la nostra llengua?

[-] Hi ha algun exemple en el qual el codi numèric també hagi estat retirat de la llista oficial de codificació?

6.

La nostra proposta en cap cas va en contra del manteniment de la compatibilitat amb els tres elements separats <Ela> + <Punt Volat> + <Ela>: <U+006C> + <U+00B7> + <U+006C> i <U+004C> + <U+00B7> + <U+004C>. És ben clar que necessitem aquesta compatibilitat (sobretot pel que fa als dominis d'Internet!).

[+] No entrem obertament en conflicte amb les normes d'Unicode ni de la ICANN, només volem tenir una ferma solució tipogràfica alternativa.

[+] En qualsevol document escrit o de comunicació visual la correcció tipogràfica és una prioritat.

[-] Dupliquem el criteri de cerca de paraules amb Ela Geminada. Els cercadors hauran de treballar tant en la cerca de 3 signes (el criteri que tenim ara) com en un de sol (el criteri que proposem i que és més senzill per a l'algorisme).

[-] Hi ha àmbits en els quals la compatibilitat tècnica preval per damunt de la correcció tipogràfica.

Un cas clar és dins l'actual món dels dominis d'Internet amb caràcters fora de l'ascii, on la ICANN ha establert un mecanisme de codificació alternatiu anomenat *Punycode*: <https://www.dw-formmailer.de/index.php?action=convert>, on, per exemple, l'adreça accentuada

<http://www.fundació.cat> també pot escriure's d'aquesta manera: <http://www.xn--fundaci-r0a.cat>, de forma que amb aquest mètode s'assegura la universalitat d'una URL en qualsevol teclat.

7.

La nostra proposta no és nova. És, només, la recuperació d'un signe que ja tenim en tipografia clàssica (sobretot usat dins el període 1915-1985 amb tipus mòbils i linotípia) i que, inexplicablement, ha quedat ignorat en tipografia digital.

[+] Vegeu el catàleg de tipus per a linotípia de la Mergenthaler-Linotype (Nova York, 1940):
http://femfum.com/PDF/ElaGeminada/LinotypeCutting_SpecimensCatalans.pdf

[+] Vegeu exemples de tipus mòbils (ploms de l'Editorial Barcino i llibres de pòlisses de la tipogràfica Bauer):

http://femfum.com/PDF/ElaGeminada/5JPL_PonenciaElaGeminada.pdf

[+] Vegeu el cas de les tipografies digitals d'Adobe (Courier i esperem que d'altres), en les quals s'inclou la Ela Geminada sencera:

http://femfum.com/PDF/ElaGeminada/comparativa_Courier002004.pdf

[-] Caldria saber si existeix cap caràcter que hagués estat fabricat en linotípia i s'hagués deixat de generar en tipografia digital, i, en conseqüència, no li hagués estat assignat mai cap codi Unicode.

8.

Podem demostrar tipogràficament que una correcta composició de la Ela Geminada —tal com marca la norma— no serà mai possible per mitjà dels tres caràcters actuals.

[+] La solució per compondre correctament els 3 caràcters passaria per poder controlar al 100% l'espaiat entre ells, i això és un impossible evident davant la multiplicitat d'aplicacions i criteris, molt variables en la composició.

[+] Si podem saber amb certesa que el punt volat no només s'utilitza per compondre la Ela Geminada (on es pot veure això clarament a Unicode?

<http://unicode.org/cldr/utility/character.jsp?a=00B7>), això ens pot afavorir per demostrar la inviabilitat del control d'espaiat, de la dimensió, de la correcció òptica i del posicionament vertical/horitzontal.

[+] La construcció del corrector tipogràfic d'Eles Geminades per a documents PDF en què treballem també pot ajudar a demostrar el que diem.

9.

Dins l'àmbit de la indexació i la cerca de paraules en documents electrònics caldria poder demostrar els avantatges d'un sol caràcter enfront dels inconvenients actuals (3 caràcters sovint alterats en el punt volat).

[+] Cal enumerar els problemes actuals en la indexació i en la cerca dins els diversos formats de documents electrònics oberts (HTML, PDF, XML, etc.) i els motors de cerca més populars (Google, Yahoo, Acrobat, Spotlight, IndexServer, etc.).

[+] La Viquipèdia catalana (<http://ca.wikipedia.org/wiki/Portada>) és un excel·lent exemple del caos en la cerca de paraules que duen la Ela Geminada, i que hem pogut experimentar en un recent buidat que ens ha fet un dels administradors del sistema de cara a repassar continguts del nostre diccionari.

10.

Cal explicar que aquesta proposta és el pas obligat per a assolir l'objectiu final de poder tenir una tecla directa en tota mena d'estris electrònics que puguin ser configurats en llengua catalana.

[+] L'adveniment dels teclats tàctils de pantalla ens ho facilitarà. Vegeu l'iPhone:

<http://www.apple.com/iphone/usingiphone/keyboard.html>

o el teclat Optimus:

<http://www.artlebedev.com/everything/optimus-concept/>

11.

Cal tenir molt present el pes específic de la Ela Geminada dins l'àmbit de la llengua escrita: més de 1.100 arrels diferents i més de 3.500 paraules la duen (estem elaborant un diccionari, que utilitzarem per al corrector tipogràfic, pentinant molt diverses fonts).

[+] Es pot concloure que té una presència més gran que signes com la 'z' i la 'x', i, per descomptat, que la 'k' i la 'w' (forasteres en català).

[+] Continuarà creixent gràcies als neologismes.

[+] És l'únic element tipogràfic distintiu de la identitat de la llengua catalana (juntament amb al .CAT per a la www).

12.

Caldria tenir un representant per a la llengua catalana a Unicode.org

[+] Formar part de l'organisme Unicode hauria de ser un objectiu estratègic per assegurar la presència internacional de la nostra llengua dins els estàndards en què la tecnologia es basa en l'àmbit global.

[-] Fins ara, en l'àmbit institucional i en aquells temes en els quals la llengua ha d'anar lligada a productes tecnològics, s'ha seguit una estratègia molt parcial i s'ha negociat per separat amb empreses i per productes (Adobe, Apple, Microsoft, Nokia, etc.), i els resultats sempre han estat intermitents.