

Diacrítics: teoria i termes

PERE FARRANDO

Professional i col·laborador de l'escola Eina
ORCID: 0000-0003-1030-8415
perefarrando@preedicio.com

Pere Farrando (Vic, 1968) és

llicenciat en filologia catalana per la Universitat de Barcelona i corrector de textos de professió. Va començar a publicar el 2008 amb *Sobre la cursiva* (Pagès) i ha continuat amb diversos articles, entre els quals destaquen «Josep M. Pujol i la composició del text» (*Llengua i Ús*, 53) i «De l'ortotipografia a l'editatge» (*Llengua i Ús*, 58). Ha col·laborat en el Màster en Tipografia Avançada de l'escola Eina de Barcelona i és especialista en composició del text amb InDesign. Al seu lloc web, www.preedicio.com, es poden descarregar la major part dels seus textos i consultar una bibliografia de i sobre Josep M. Pujol.



Resum

El present repàs de termes dels signes diacrítics comença amb una contextualització teòrica elaborada a partir dels ensenyaments de l'obra recent de Dimitrios Meletis. A continuació, es fa una definició provisional de signe diacrític i es comenten els termes dels diacrítics habituals, que tenen formes etimològiques o creades per especialització semàntica. Quan es considera escaient, es fa alguna proposta terminològica.

PARAULES CLAU: signe diacrític; escriptura; grafolingüística; grafemàtica; ortografia

Abstract

Diacritics: theory and terms

This survey of terms for diacritical marks starts out with a theoretical contextualization based on the recent work of Dimitrios Meletis. Next, we propose an interim definition of diacritical mark and we discuss the more usual diacritics, which adopt etymological or other forms created via semantic specialization. Lastly, as appropriate, we postulate some terminological solutions.

KEYWORDS: diacritical mark; writing; grapholinguistics; graphematics; spelling

TERMINÀLIA 23 (2021): 7-16 · DOI: 10.2436/20.2503.01.158
Data de recepció: 11/12/2020. Data d'acceptació: 01/02/2021
ISSN: 2013-6692 (impresa); 2013-6706 (electrònica) · <http://terminalia.iec.cat>

Preàmbul

Durant bona part de la redacció d'aquest article, la intenció era tractar primer el terme *signe diacrític* i després passar al repertori de diacrítics habituals. La primera part, però, va quedar desfasada quan, la primera quinzena de novembre del 2020, es va publicar el llibre *The nature of writing. A theory of grapholinguistics*, de Dimitrios Meletis, que es pot comprar en paper o descarregar gratuïtament en PDF i que pot ser vist sense recança com un gran pas endavant. No solament recull en una sola exposició una gran varietat de contribucions bibliogràfiques —molt destacadament les escrites en alemany, que no estaven tenint una circulació apropiada—, sinó que a més presenta un sistema ben travat per a l'estudi de l'escriptura i les escriptures, encara que estrictament parlant parteixi de línies ja formulades. No hi fa res que en algun moment ens sembli que estira massa tal fil o tal altre. El conjunt de l'obra es mereix el reconeixement.

1 Un tast de teoria i els seus termes

L'entramat d'elements que Meletis organitza com a proposta teòrica serveix per aprofundir tant en l'escriptura en general com en les escriptures en particular. L'autor —que afirma que segueix Martin Neef— postula que en l'escriptura operen tres mòduls —a qui li abelleixi en pot dir *components*—: el grafètic, el grafemàtic i l'ortogràfic. El mòdul grafètic (Meletis, 2020, p. 24-25) té a veure amb la dimensió material (visual i/o tàctil) de l'escriptura. Aquí els signes són tractats per ells mateixos, sense tenir en compte el seu valor lingüístic o informatiu, i no es consideren solament els signes «lletrítics», sinó també elements gràfics segmentalment superiors, com la línia, la columna i l'àrea d'escriptura, anomenada *espai holístic*. Al seu torn, el mòdul grafemàtic (Meletis, 2020, p. 26-28) s'ocupa de les relacions entre els signes i les unitats lingüístiques que representen (fonema, mora, síl·laba, morfema). I a continuació hi ha el mòdul ortogràfic (Meletis, 2020, p. 28-29), que no és necessari per naturalesa, perquè l'escriptura pot existir sense que la comunitat que la utilitza n'hagi compendiat les regles d'ús.

El desplegament teòric, Meletis l'acompanya d'un treball terminològic acurat. En el mòdul grafètic, els dos termes essencials són *graf* i *signe bàsic* o *forma de base* (traduccions possibles de l'anglès *basic shape*). En paral·lel a les nocions de so i fonema, el graf és la instanciació, la realització d'un signe gràfic, mentre que el signe bàsic és la noció abstracta del signe material. La convenció de Meletis per delimitar les expressions grafètiques en els textos és la barra vertical. Per tant, direm que el turc té el signe bàsic |1| (una i minúscula sense punt) i que l'alemany té |ß|.

Pel que fa al mòdul grafemàtic, podem celebrar que *grafema* passi a tenir una definició sòlida (Meletis, 2020, p. 26):

In a nutshell, in the present model, the grapheme is conceived of as a dyadic sign in the Saussurean sense, consisting of two inseparable constituents: the *signans* (pl. *signantia*), a graphetic unit (the most elementary of which is the basic shape), and the *signatum* (pl. *signata*), a linguistic unit such as a phoneme or a morpheme.

A més a més, el terme *caràcter*, que s'ha estat utilitzant en la bibliografia per denominar signes complexos d'escriptures com la xinesa, queda superat per la capacitat del grafema de tenir components subsegmentals. La convenció, ara —ja present en bibliografia anterior—, és la parella del més-petit-que i el més-gran-que, per exemple <1> i <ß>.

2 L'aplicació als signes diacrítics

Com explica Meletis, els diacrítics són una classe especial dins els signes bàsics del mòdul grafètic: no s'utilitzen mai sols, sempre en combinació. Són els únics signes que tenen la propietat de l'additivitat (*additivity*) (Meletis, 2020, p. 44):

Additivity is also graphetic in nature; it characterizes diacritics, i.e. smaller and dependent segments of basic shapes that attach to basic shapes within a single segmental space.

Grafemàticament parlant, els diacrítics també expressen una realitat subsegmental (Meletis, 2020, p. 103):

Other examples of subsegmental graphematic functions include the |ö|-diacritic in umlaut graphemes such as <ö> and <ü> in German that graphematically marks a change of the vowel phoneme represented by the unmarked graphemes <o> and <u>, or the Japanese diacritic |^ˆ|, which, when it is added to hiragana graphemes, marks voicing of the initial consonant phoneme in the syllables (or moras) represented by the graphemes: <き> /ki/ becomes <ぎ> /gi/.

En un altre punt (Meletis, 2020, p. 261), esmenta grafemes amb diacrítics com <ç, ñ, š>, però no comenta enlloc els sistemes d'accentuació del català, l'espanyol i el portuguès, que tenen una casuística, lingüísticament parlant, lexemàtica. Alerta: Meletis ens diria que no podem utilitzar aquí un concepte de la lingüística, perquè l'escriptura és un sistema autònom, dotat d'uns recursos propis. Aleshores, com expliquem una alternança com <consciència>/<consciència>?

Com hem vist fa un moment, els signes diacrítics són un component subsegmental. Ara bé, la seva presència crea un grafema, no un al·lògraf. Per entendre'ns, fins i tot en la parella <fals>/<falç>, on <s> i <ç> representen igualment /s/, l'alternança és entre grafemes, no entre al·lògrafs d'un grafema, perquè formen part de constituents morfosintàctics diferents (adjectiu masc. sing. vs. substantiu fem. sing.) (Meletis, 2020, p. 117, nota 134). (Una alternança al·logràfica seria la de <fals>/<fals>.) En la mateixa línia, la parella <consciència>/<consciència> té una explicació grafemàtica.

Molt bé, però, aleshores, com expliquem l'element prosòdic que governa l'accentuació en les escriptures esmentades? Aquí entra en joc el mòdul ortogràfic de tal escriptura o tal altra, que actua com un factor extern a la sistematicitat que s'observa en els mòduls grafètic i grafemàtic. La prova és que l'accentuació (ortogràfica) no resol ni de bon tros tots els casos d'homografia que es presenten en els textos; vegeu la parella d'homògrafs i homòfons <deu> 'aflorament d'aigua' i <deu> 'nou més un', o la parella d'homògrafs <sou> 'moneda' o 'remuneració' i <sou>, forma flexionada del verb ésser. En un cas com l'accentuació en francès de <château>, que no respon a la prosòdia, l'explicació continua sent ortogràfica, a part del fet que <â> sigui per si mateix un grafema.

Observeu que en cap moment he parlat de *paraula grafemàtica*, concepte que Meletis tracta (*graphematic word*) (Meletis, 2020, p. 133 i s.), perquè és present en la bibliografia de la grafolingüística, però que no postula com a vàlid perquè no és universal.

3 Definició provisional de signe diacrític

Malgrat que l'obra de Meletis mereix passar de la primera lectura, crec que podem assajar una definició «interina» de *signe diacrític*, que escaigui a la seva naturalesa grafètica, grafemàtica i ortogràfica. Per començar, l'enfocament no pot ser fonetista. És la pega de la proposta del DEC, que sembla fins i tot que es refereixi solament al repertori de l'alfabet fonètic internacional (va introduïda per l'etiqueta «FON»): «Senyal convencional que hom afegeix a un signe gràfic i que representa una realització fonemàtica, per indicar alguna mena de modificació fonètica.» En segon lloc, la proposta no pot ser tan breu com la del DIEC2: «Signe gràfic que confereix a un altre signe un valor diferent.» És la mateixa redacció que divulga el Cercaterm (del Diccionari de lingüística del TERMCAT), que l'amplia amb una nota.

L'assaig de definició podria anar així: «Signe gràfic d'una escriptura que complementa algun signe de base, connectant-s'hi o no, i que modifica el valor d'aquest signe o respon a una convenció ortogràfica.»

Comentaris sobre la definició proposada:

1. «Signe gràfic d'una escriptura». Aquí s'ha de parlar d'«escriptura», i no d'«ortografia», que tindria un sentit restrictiu.

2. «que complementa algun signe de base». Aquí s'explica que el signe no té entitat per si sol. No crec que sigui apropiat utilitzar lletra en comptes de *signe de base*, perquè, com hem vist més amunt, en el japonès també hi ha un diacrític, i no es diu que el japonès tingui lletres.
3. «connectant-s'hi o no». Amb aquesta acotació, queda oberta qualsevol col·locació del diacrític en relació amb el signe de base, i certament se'n col·loquen en totes les posicions.
4. «i que modifica el valor d'aquest signe o respon a una convenció ortogràfica». Aquesta part és la més delicada de la definició. Dient «que modifica el valor», expliquem la naturalesa grafètica i grafemàtica del signe; dient que «respon a una convenció ortogràfica», admetem sistemes d'accentuació prosòdica i no prosòdica.

4 La codificació: convencions i termes

Si passem pàgina de la teoria de l'escriptura i volem examinar els diacrítics en la vida real, hem d'anar a parar a la seva codificació, és a dir, al tractament que se'n fa perquè siguin processables pels dispositius digitals, i la mare de totes les codificacions és Unicode. Cap altra taula de codis està concebuda i preparada per donar compte formalment de totes les escriptures del món. Encara n'hi ha que no hi són, com explica el lloc web de la Script Encoding Initiative (www.linguistics.berkeley.edu/sei/index.html), però Unicode té capacitat per codificar-les totes —si no perd el nord pel camí per la febre dels emojis, que un no sabia dir què hi fan, en l'estàndard. Una introducció sumària a Unicode molt ben construïda és la de James K. Tauber (2019), que la fa per parlar després del grec polític.

A l'hora de consignar els punts de codi d'Unicode, se segueixen algunes convencions. (Observeu que a partir d'ara utilitzaré la barra vertical, que hem vist que era pròpia de les expressions del mòdul grafètic.) Per exemple, la *ela* barrada del polonès, |ł|, és 0142 LATIN SMALL LETTER L WITH STROKE. No cal fer «u+0142», perquè el prefix «u+» és innecessari en textos divulgatius. En una obra ben composta tipogràficament, es pot utilitzar la versaleta per a la definició del signe, com fem aquí, en comptes de les majúscules d'Unicode, però aleshores es recomana que les xifres del punt de codi siguin antigues (0142).

Ja s'ha pogut veure més amunt que s'utilitza un caràcter comodí com a base per als signes diacrítics, el cercle puntejat: |◌| 25CC DOTTED CIRCLE. Moltes lletres no el tenen, però sí la Times New Roman.

Els terminòlegs han de ser conscients d'un fet: Unicode no té com a missió la normalització de terminologia. Un exemple d'això és la seva definició dels caràcters de la barra inclinada i la contrabarra (|/, \|), en què opta per uns termes secundaris, 002F SOLIDUS i 005C REVERSE SOLIDUS, en comptes dels més habituals slash i backslash.

Una altra característica d'Unicode és que assigna una categoria a cada punt de codi, de manera que a partir de la llista de categories (consultable a www.fileformat.info/info/unicode/category/index.htm) podem observar com estan classificats tots. La que ens interessa ara és l'anomenada «Mark, Nonspacing». Observeu que Unicode, com l'anglès en general, utilitza *mark*, que en català no hauria de ser *marca*, sinó *signe*. Noteu també l'adjectiu *nonspacing*, antitètic de *spacing*. Traduiria ambdós per *no espaiat* i *espaiat*.

Efectivament, els diacrítics poden ser signes espaiats o no espaiats, això és, dotats d'amplada horitzontal o no. Els que ens interessin ara són els no espaiats. Funcionen de tal manera que, en el moment d'inserir-los en el punt del text on el cursor està actiu, salten cap a l'esquerra i es col·loquen en la posició que els pertoca en relació amb la lletra de base. En realitat, totes les lletres amb diacrítics de les escriptures occidentals consten a Unicode també en la seva forma compacta, d'una peça, tal com estaven codificades en les taules de codis antigues i que Unicode va haver d'incorporar per no crear un problema de falta de compatibilitat a escala mundial (figura 1).

En Unicode, els punts de codi dels diacrítics no espaiats tenen definicions en què s'observen unes expressions productives. La primera és l'adjectiu *combining*, que tradueixo per *de combinació*. Per exemple, |◌̇| 0323 COMBINING DOT BELOW i |◌̂| 0342 COMBINING GREEK PERISPOMENI. La segona fórmula productiva és l'ús dels adverbis *above*, *below* i el nom *overlay* per denominar posicions en relació amb el caràcter de base. Per exemple, tenim |◌̘| 032D COMBINING CIRCUMFLEX ACCENT BELOW i |◌̃| 0334 COMBINING TILDE OVERLAY. En català, podríem optar tant per *sobreescrit* i *sotaescrit* com pels sintagmes *a sobre* i *a sota* (la meua tria aquí), i per *superposat* en el cas de *overlay*. La tercera constant en les definicions de caràcters és que les minúscules s'anomenen *small letter*, i les majúscules, *capital*. Per exemple, tenim |◌̈| 036A COMBINING LATIN SMALL LETTER H.

Per tant, tots els punts de codi que acabo d'esmentar poden quedar en català de la manera següent, amb

fidelitat a la definició d'Unicode sense que es perdi el to de llengua natural:

◌̇	punt de combinació a sota
◌̂	perispomeni de combinació
◌̘	circumflex de combinació a sota
◌̃	titlla de combinació superposada
◌̈	hac minúscula de combinació

5 Estratègia terminològica de les llengües

Pel que fa als noms dels diacrítics, les llengües o bé utilitzen noms amb base etimològica o bé creen termes per especialització semàntica, és a dir, hi assignen noms de la llengua general que descriuen la forma del signe. Per la banda dels termes etimològics, hi ha *circumflex*, que ve del llatí *circumflexus*; *titlla* és substantiu de *titllar*, que prové de *titulare* (igual que el verb *titular*); *dièresi* ve del grec *διαίρεσις*, però el francès i el portuguès opten per *tréma/tréma*, que prové del també grec *τρήμα*. Entre els casos de creació per especialització semàntica, tenim l'eslovac, que anomena la dièresi *dve brodky* ('dos punts'); el francès, de l'anella a sobre en diu *rond en chef*, i del punt a sobre, *point en chef*; més endavant veurem el *kreska* i l'*ogonek* del polonès, i el *háček* i el *kroužek* del txec.

A *The insects project*, l'hongarès Robert Kravjanszki (2016, p. 41) comenta que la seva llengua no té mots per a aquests signes i utilitza com a alternativa unes denominacions pròpies del llenguatge Postscript, com *egrave* per a |è|:

There are no Hungarian translations for those marks. In the rare cases that we do come across a discussion about them, we find inaccurate descriptions. For example, the dotaccents above i and j are called *dots* – the same name that is used for the full stop. This can also be found in the terms *kettőspont* (colon) and *pontosvessző* (semicolon). The former, written separately (*kettős pont*), stands for the dieresis. [...] In the absence of a standardised terminology, I will be using standard Postscript names for diacritical marks.

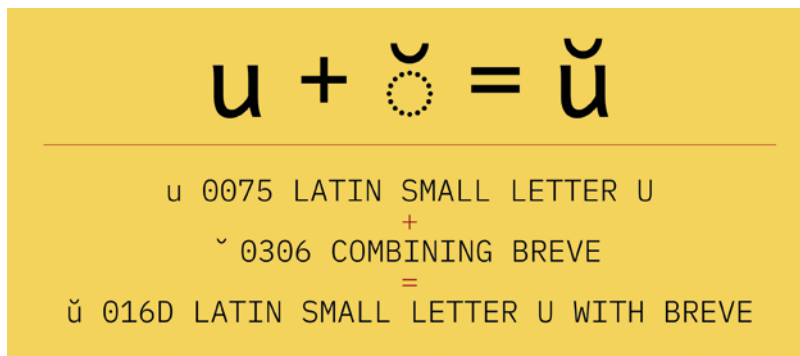


FIGURA 1. Les lletres amb diacrítics es poden construir de manera modular, com es veu en la primera part d'aquesta «equació», en què s'enumeren dos punts de codi, un per a |u| i un per al breu, |◌̃|. Tanmateix, Unicode conté també punts de codi de les formes íntegres, d'una peça, |ü|, per una raó de compatibilitat amb les taules de codis anteriors.

En l'altre extrem, l'eslovac Palo Bálik (2016, p. III-III4) dona termes dels diacrítics en la seva llengua (*dlžeň*, accent agut; *mākčeň*, hàtxek; *vo káň*, circumflex, i el ja esmentat *dve bodky*, dièresi).

El català es comporta com les altres llengües europees d'escriptura romana, o potser una mica millor: per als diacrítics amb què estem poc familiaritzats ja tenim adaptacions com màcron i hàtxek i la nova ortografia del 2017 ens va regalar *trenc* quan alguns ja creïem que no hi havia més remei que adoptar l'espanyol *cedilla*.

A continuació abordo uns quants diacrítics i proposo denominacions quan no en tenim (vegeu exemples de disseny de diacrítics a les figures 3 i 4).

5.1 Els accents obert i tancat

Els accents obert i tancat, o greu i agut, no presenten cap problema terminològic. En espanyol, al costat de *acento*, s'utilitza *tilde*, que té menys precisió, perquè el seu significat cobreix tant l'accent com la titlla. El polonès té unes quantes consonants amb accent tancat (figura 2): [ć, ń, ś, ź]. El terme per referir-s'hi és *kreska* ('ratlla, línia'). L'eslovac posa accent tancat en [ě, ř, ý]. L'hongarès fa ús de l'accent tancat doble: [ő, ű].

No crec que hi hagi cap conflicte a anomenar *accent tancat* o *agut* el signe posat sobre consonants, encara que la dicotomia obert/tancat deixi de tenir-hi sentit. En Unicode, aquestes consonants són codificades amb el terme *acute*: [ć] 0107 LATIN SMALL LETTER C WITH ACUTE.

Font: Arxiu de l'autor



FIGURA 2. Dues plaques d'un mateix carrer de Varsòvia que mostren com els diacrítics han sigut durant molt temps un punt dèbil en el disseny de lletra, sobretot els que s'apliquen a majúscules (a les imatges, [Ś]). La situació actual és molt millor, encara que no òptima. Moltes escriptures més enllà de les d'alfabet romà, grec i ciríl·lic presenten reptes a Unicode i encara més al format de lletra OpenType.

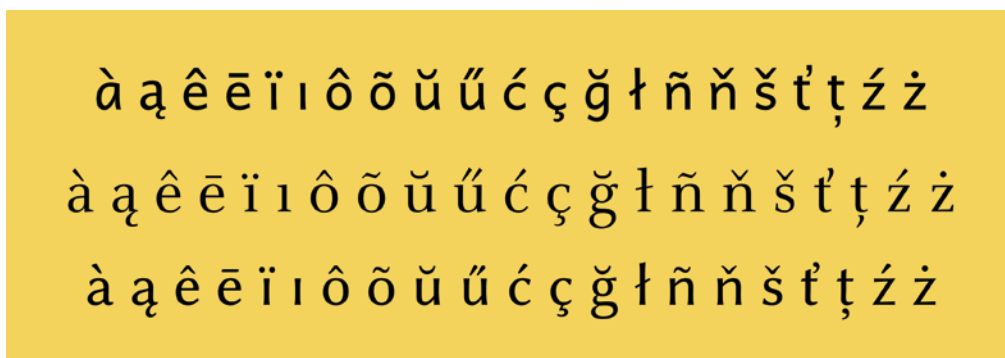


FIGURA 3. Mostres de disseny de diacrítics. En primer lloc, lletra Andika, de SIL, que té uns diacrítics proporcionalment grans. Es tracta d'una lletra destinada a l'alfabetització (software.sil.org/andika/). En segon lloc, lletra Celeste, de Christopher Burke. En tercer lloc, lletra Linux Biolinum, de gran repertori (2.401 glifs la rodona). En els tres casos, el hàtxek de la te s'assembla massa a l'apòstrof.

5.2 La dièresi

La dièresi tampoc presenta cap dubte terminològic. Unicode opta per *diaeresis*, però, així com de vegades aporta denominacions relacionades, en aquest cas no diu res ni del francès *tréma* ni de l'alemany *Umlaut*. Gràficament parlant, aquest darrer no es diferencia gens de la dièresi.

En català, es podria prendre en consideració que la dièresi i l'*Umlaut* fan funcions diferents (el signe alemany indica metafoia 'inflexió d'una vocal per influència d'una altra vocal del mot') i imitar el Larousse del francès, que entra *umlaut*. L'entrada catalana podria ser *úmlaut*, adaptació de l'original, plana i masculina.

Bringhurst (2005, p. 306) qualifica l'accent tancat doble de l'hongarès de *long umlaut* i diu: «*ű* is the long form of *ü*».

5.3 La titlla

La distribució de significats de titlla i vírgula que proposa el Manual d'estil (2009, p. 171, n. 5) és molt pràctica i dona al català capacitat de precisió:

En aquest MANUAL D'ESTIL, distingim la vírgula, que és una línia ondulada horitzontal situada a l'altura de l'ull mitjà de la lletra (~), de la titlla, que és una 'ratlleta que es posa sobre una lletra com a signe diacrític o d'abreviació, especialment el que s'usa en espanyol en la ñ o en el portuguès en alguns diftongs nasals'.

Aquesta titlla és la correspondència directa de la tilde d'Unicode (|̃| 0303 COMBINING TILDE). Pel que fa a vírgula, Unicode solament esmenta *virgule* com a sinònim en la barra inclinada (|/| 002F SOLIDUS). Prové del llatí *virgula* i es va estendre com a pausa intermèdia a la baixa edat mitjana, efectivament amb una forma semblant a la barra inclinada actual. Parkes (1993, p. 46): «The *virgula suspensiva* became almost as common a mark of punctuation as the *punctus*, and was used primarily to indicate medial pauses.» El DIEC2

en dona una definició genèrica en l'accepció 1, «Ratlleta curta i prima», i la considera sinònima de coma en l'accepció 2; en aquesta, sembla clar que s'hi hauria d'afegir una acotació que digués alguna cosa com «utilitzada en la puntuació medieval».

La fitxa de titlla del Cercaterm replica la definició del DIEC2: «Ratlleta que es posa sobre una lletra com a signe diacrític o d'abreviació, especialment el que s'usa en espanyol en la ñ o en el portuguès en alguns diftongs nasals.» Tots dos recursos haurien de deixar de dir que «es posa sobre una lletra» i substituir-ho per alguna cosa com ara «es combina amb algunes lletres», perquè en Unicode hi ha una titlla de combinació a sota (0330), una de superposada (0334), una de vertical a sobre (033E) i una de doble a sobre (0360). Un exemple de titlla superposada és la de l'alfabet fonètic internacional (AFI), per exemple |ʈ|, que indica velarització.

5.4 El circumflex i el hàtxek

Així com *circumflex* és paraula internacional, *circumflex invertit* mereix algun comentari. En català ja tenim *hàtxek* a la Neoloteca (|̂| 030C COMBINING CARON). Curiosament, la fitxa no ens dona l'alternativa *circumflex invertit*, que entre nosaltres proposava Josep M. Pujol (2001, p. 28, n. 13):

El mot *háček* (diminutiu de *hák* 'ganxo') és el que designa aquest signe en txec, i s'utilitza també en anglès (sovint sense accent i al costat del menys freqüent *wedge* 'tascó') i francès. No he pogut trobar enlloc fins ara l'origen del mot anglès *caron*, preferit per Unicode però que manca als diccionaris de referència d'aquesta llengua [...]. Una denominació descriptiva neutra seria *accent circumflex invertit*.

La presència de l'anglès *caron* actualment en molts textos es pot deure al fet que és l'opció d'Unicode. És la forma que utilitza Bringhurst al seu glossari (2005, p. 304), on fins i tot parla de «*caroned letters*».

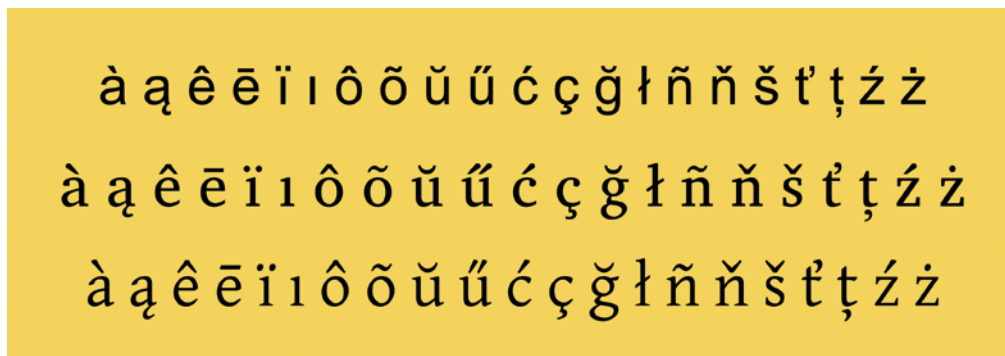


FIGURA 4. Més mostres de diacrítics. Primera lletra, Arial. El circumflex, el hàtxek i els accents simple i doble queden tallats. Segona lletra, Charis, de SIL, molt àmplia (3.692 glifs la rodona) (software.sil.org/charis/). Tercera lletra, Gentium Plus, també de SIL, més estilitzada i encara més àmplia (4.359 glifs la rodona) (software.sil.org/gentium/). Aquesta versió (Plus; l'altra s'anomena Basic) solament té estils de rodona i cursiva.

El Hàtxek té la particularitat que en txec i en eslovac —llengües molt pròximes de països veïns— té una forma diferent quan s'adjunta a certes minúscules i a una majúscula: [ḍ, ḷ, ṭ, ḷ]. El disseny d'aquesta variant hauria de ser prou diferent de l'apòstrof perquè el lector no tingués dificultat per distingir-los.

En Unicode hi ha també el circumflex i el Hàtxek de combinació a sota (032D, 032C).

5.5 El màcron i el breu

En català, ja tenim màcron i breu a la Neoloteca. Pujol (2001, p. 28, n. 9) es va avançar i ja utilitzava l'adaptació màcron. Pel que fa al breu, s'hi referia amb el llatíisme *nota brevis* (2001, p. 28, n. 10):

A manca de denominació catalana fujo d'estudi amb la terminologia llatina de sant Isidor «Figura est, cuius nota syllabae agnoscuntur. [...] Nam nota brevis inferior semicirculus est [...]» (Or. 1.17.29). Els classicistes anglesos utilitzen correntment el terme *breve*, derivat de l'acusatiu singular neutre de *breuis-e* 'breu', però el 1755 ja en parlava John Smith [...], que en diu *short* (i *long* del màcron) [...].

Les definicions de màcron i breu de la Neoloteca tornen a ser restrictives quant a la posició: «se situa sobre algunes lletres». En Unicode hi ha màcrons i breus a sobre i a sota, normals i invertits, i simples i dobles.

Fan ús del breu el romanès, |ǔ|; el turc, |ğ|, i l'esperanto, |ŭ|.

5.6 L'anella

Pujol (2001, p. 28) anomena *anell sobrescrit* el diacrític |◌̊| 030A COMBINING RING ABOVE. Aquí m'inclino per *anella* (sembla que *anell* tendeix més a ser un cercol que envolta alguna cosa en forma de tub). En tot cas, *sobrescrit* hauria de ser *sobreescrit* o *a sobre*. També existeix l'anella de combinació a sota (0325).

En espanyol, Jorge de Buen (2017, p. 35) entra *anillo*. El txec la utilitza en |ů| i l'anomena *kroužek* ('anellet'); el suec i el danès la posen en |å|. També s'utilitza en el símbol de l'àngstrom (212B).

5.7 El punt

Hi ha uns quants idiomes europeus que fan ús del punt de combinació a sobre (◌̇| 0307 COMBINING DOT ABOVE): lituà, |ė|; polonès, |ż|; maltès, |ċ, ġ, ž|. El cas més singular és el turc, que tracta el punt de |i| com un diacrític i arriba a fer |ı̇|, i |İ̇|. També existeix el punt de combinació a sota (0323).

En català, tenim l'esment de *punt alt* que fa el *Manual d'estil* (Mestres et al., 2009, p. 170), però l'enumera dins del grup dels signes auxiliars, no el dels diacrítics. És probable que es refereixi a una puntuació antiga i medieval, del sistema d'Eli Donat anomenat *distinctio-*

nes, del segle IV, que va perviure en manuscrits medievals (Parkes, 1993, p. 13, pàssim). En torno a parlar a l'apartat «El punt volat i el cas del català».

Molts diccionaris d'anglès entren *tittle* 'punt o ratlleta sobre una lletra com a diacrític o per indicar una abreviació'. És el signe de |i, j| en la lletra gòtica medieval, en què era molt difícil distingir lletres com |n, u, i, j| (Buen, 2017, p. 208-209).

En anglès existeixen *overdot* i *underdot* (Bringhurst, 2005, p. 315 i 319). En polonès, el punt a sobre s'anomena *kropka* i en l'escriptura manual es pot trobar tant en forma de punt, |ż|, com de ratlleta superposada, |z| (Małecka i Oslislo, 2016, p. 80).

5.8 El trenc

El terme *trenc* el va introduir el 2017 la nova ortografia de l'Institut d'Estudis Catalans (OIEC, p. 28). Unicode opta per l'espanyol *cedilla* (◌̣| 0327 COMBINING CEDILLA), que de fet es troba en altres llengües (Buen, 2017, p. 61): en francès *cédille*, en portuguès *cedilha*, etc.

Alerta: Jorge de Buen (2017, p. 61) diu erròniament que el romanès utilitza el trenc en |s, t|. És una conseqüència de les vicissituds que ha viscut la codificació del romanès. El 1987, aquesta llengua va quedar vinculada a la taula de codis ISO 8859-2, que incloïa les dues consonants amb trenc |ș, ț|, en comptes de les correctes amb coma a sota. Els sistemes operatius de Microsoft i Apple no resolien el problema. El 1999, la versió 3 d'Unicode va introduir les formes correctes amb coma a sota i a Romania es va publicar l'estàndard SR 13411, que les convertia en oficials. Actualment hi ha aplicacions que executen |ș, ț| quan se selecciona l'idioma romanès, però en molts entorns la selecció d'idioma no és una opció. Es pot llegir la cronologia d'aquestes vicissituds, molt més enrevesada que el que comento aquí, a Kit (2019).

5.9 La coma

No veig cap pega a anomenar *coma* el diacrític que Unicode anomena *comma*, perquè el signe en efecte té forma de coma petita. Existeixen les comes de combinació a sobre (0312), a sota (0326), a sobre inversa (0314), a sobre girada (0312) i a sobre a la dreta (0315).

Segons Bringhurst (2005, p. 319), la coma a sota és una variant del trenc que es va fer popular en el període d'ús de la màquina d'escriure, perquè el trenc hi solia mancar i era fàcil teclejar una coma com a alternativa. També afirma (2005, p. 304) que en letó i romanès «the undercomma is preferred». Evidentment, pel que hem vist en l'apartat anterior, *preferred* no és la paraula correcta, tampoc per al letó. La particularitat d'aquest idioma és que en |g| (solament en la minúscula) posa la coma a sobre, i per tant fa |ḡ Ḣ|, però |ḡ Ḣ, ḡ Ḣ, ḡ Ḣ|.

5.10 L'ogònek

El català no té cap terme per a aquest diacrític (|◌̇| 0328 COMBINING OGONEK). Małecka i Oslislo (2016, p. 79) afirmen que *ogonek* 'cueta' és «the only Polish word in the international typographic terminology». Pujol (2001, p. 28, n. 12) encara no es va llançar a fer l'adaptació *ogònek*. Entre dir-ne *cua* i agafar prestat i adaptar el terme polonès (paraula masculina i plana), m'inclino per això segon, igual com ja s'ha fet amb *hàtxek*.

Fan ús de l'ogònek el polonès, |ą, ę|, i el lituà, |ą, ę, į, ū|.

5.11 La sotaratlla

Bringinghurst (2005, p. 313 i 319) parla de *lowline*, que seria la sotaratlla normal, i *underscore*, la sotaratlla de combinació. Unicode fa el contrari, perquè la de combinació és |◌̣| 0332 COMBINING LOW LINE.

Les llengües europees no utilitzen la sotaratlla com a diacrític, però segons Bringinghurst (2005, p. 319) sí que es troba en llengües africanes i indigenoamericanes.

El *Manual d'estil* (Mestres et al., 2009, p. 170) esmenta com a signes auxiliars la ratlla baixa o sotaratlla i, d'una manera gens comprensible al meu entendre, el subratllat, com si fos un caràcter i no un atribut tipogràfic.

5.12 La barra o ratlla superposada

Introdueixo aquest terme sense estar segur que calgui normalitzar-lo com a signe diacrític. Es tracta de la ratlla transversal de |ł, Ł| del polonès. Małecka i Oslislo (2016, p. 85 s.) s'hi refereixen amb l'anglès *slash* i la tracten de diacrític, perquè grafèticament i grafemàticament ho és. També és el cas de |ø|, del danès, el feroès i el noruec, i la lletra *eth*, |ð|, de l'islandès i el feroès.

El que ocorre és que Unicode no conté cap *stroke* que es pugui combinar amb |l, o, d| per donar aquest resultat ni tinc constància que les ortografies de les llengües esmentades considerin que utilitzen cap diacrític. La classificació de Gaultney dels diacrítics per la seva morfologia gràfica (Gaultney, 2002, p. 4) inclou dos *solidus* d'Unicode que podrien servir per compondre les minúscules i majúscules en qüestió (|◌̆| 0337 COMBINING SHORT SOLIDUS OVERLAY, |◌̇| 0338 COMBINING LONG SOLIDUS OVERLAY), però en la resta del seu treball els ignora i, en efecte, quan els combino amb |l, o, d|, no creen les lletres apropiades.

Crec que les expressions *ela*, *o* i *de barrada* són pràctiques i es poden utilitzar al costat de *ela*, *o* i *de* amb *ratlla transversal*. En Unicode encara hi ha tres lletres barrades més: |ħ, ı, đ| (0127, 0166, 0110). El maltès utilitza |ħ|.

5.13 El punt volat i el cas del català

El català continua tenint el tema coent de la *ela* geminada, en bona part per culpa de la naturalesa múltiple del punt volat. Per dir-ho ràpid, en l'ortografia del català i l'occità gascó és un signe diacrític, en Unicode és un signe de puntuació (|·| 00B7 MIDDLE DOT, categoria Punctuation, Other) i en el disseny de lletra tipogràfica forma part de la lligadura de la *ela* geminada.

Terminològicament, en català, el punt volat és definit com un diacrític: ho llegim a l'OIEC (p. 28), al DEC i al DIEC2 (entrada *punt*). Al Cercaterm hi ha la fitxa de l'adjectiu *volat* sense definició, procedent del *Lèxic multilingüe de la indústria*. Hi manca una fitxa per a *punt volat*.

Estrictament parlant, l'adjectiu *volat* s'aplica a una posició superior en la vertical de la línia, que toqui per dalt o sobrepassi l'altura de majúscules (vegeu |x^a, x⁽¹⁾|). Per tant, la seva definició hauria de marcar distància amb la de *punt volat*, que en realitat es refe-

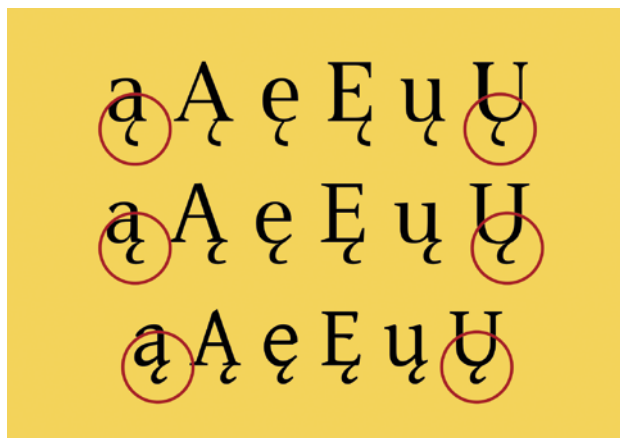


FIGURA 5. Exemples de disseny de l'ogònek. La primera lletra, Spitzkant, erra de manera flagrant: l'ogònek ha d'enllaçar bé amb el serif, i en lletres de base corbada, com la |e| i la |U|, ha de situar-se cap a la dreta de l'eix central. Les altres dues lletres, Celeste i Gentium Plus, ho fan molt millor.

reix a altura intermèdia. Una altra qüestió és com fem diferència entre l'adjectiu *volat* i el substantiu i adjectiu *superíndex*. La definició del DEC del primer, certament, val per als dos termes: «Dit de la lletra o del signe, generalment de cos més petit que el que hom empra en el text, que hom col·loca a la part superior de la ratlla.» El primer suggeriment que se m'acut és que es defineixi *volat* com una posició en l'altura de línia —que escaigui a expressions com *lletra volada*, *signe volat*, *xifra volada*— i *superíndex* com un tipus de signe de l'escriptura.

Pel que fa a punt *volat*, crec que hauria de tenir com a mínim tres accepcions, i si haguéssim de fer cas de la pàgina que hi dedica la Viquipèdia, alguna més i tot. Proposo les següents:

1. Signe diacrític en algunes llengües.
2. Signe de puntuació en algunes llengües.
3. Signe alternatiu al de multiplicació.

La segona accepció podria cobrir el punt separador de mots que hi va haver en llatí durant una franja de temps i que Parkes (1993, p. 10) anomena en anglès *interpunct*, com també el punt intermedi de les distincions d'Eli Donat que he esmentat a l'apartat sobre el punt a sobre. Perquè s'entengui bé, es tractava d'un joc de tres punts (Parkes, 1993, p. 13):

A low point (the *subdistinció*) was used to indicate a minor medial pause [...], where the sense (*sensus*) is incomplete; a medial point (the *media distinció*) indicated a major medial pause [...], where the sense is complete but the meaning (*sententia*) is not; and a high point (the *distinció*) indicated a final pause [...], where the *sententia* is completed.

Bibliografia

- BÁLIK, Palo (2016). «Designing Slovak diacritics». A: MAŁECKA, Agnieszka; OSLISLO, Zofia, (coord.) (2016). *The insects project: Problems of diacritic design for Central European languages* [en línia]. 2a ed. Katowice: Academy of Fine Arts, p. 92-115. <http://theinsectproject.eu/wp-content/uploads/2017/05/Insects-Project-2nd-edition_screen.pdf> [Consulta: 20 setembre 2020].
- BLAŽEK, Filip (2016). «Czech diacritics: From Hus to Unicode». A: MAŁECKA, Agnieszka; OSLISLO, Zofia (coord.) (2016). *The insects project: Problems of diacritic design for Central European languages* [en línia]. 2a ed. Katowice: Academy of Fine Arts, p. 12-35. <http://theinsectproject.eu/wp-content/uploads/2017/05/Insects-Project-2nd-edition_screen.pdf> [Consulta: 20 setembre 2020].
- BRINGHURST, Robert (2005). *The elements of typographic style*. 3a ed. rev. Vancouver: Hartley & Marks.
- BUNN, Jorge de (2017). *Diccionario de caracteres tipográficos*. Gijón: Trea.
- DANIELS, Peter T.; BRIGHT, William (1996). *The world's writing systems*. Oxford: Oxford University Press.
- DEC = *Diccionari.cat* ([2008]) [en línia]. Barcelona: Enciclopèdia Catalana. <<http://www.diccionari.cat>> [Consulta: 19 setembre 2020].
- DIEC2 = INSTITUT D'ESTUDIS CATALANS (2007). *Diccionari de la llengua catalana* [en línia]. 2a ed. Barcelona: Institut d'Estudis Catalans: Edicions 62: Enciclopèdia Catalana. <<https://dlc.iec.cat>> [Consulta: 19 setembre 2020].

5.14 Altres diacrítics

En l'escriptura del grec també s'utilitzen diacrítics. En la ciríl·lica, molt pocs. Tots consten en la llista de la categoria Mark, Nonspacing que he esmentat al principi, així com molts més que fan falta per al vietnamita, el turc, l'AFI i sistemes de romanització d'altres escriptures.

6 Repàs final

En aquestes pàgines, he començat esmentant la publicació recent del llibre *The nature of writing*, de Dimitrios Meletis, i n'he apuntat uns quants termes fonamentals. He assajat d'aplicar la teoria de Meletis a la noció de signe diacrític *i*, a continuació, he fet una definició provisional d'aquest terme intentant que no entri en conflicte amb els ensenyaments de l'autor.

En la part pràctica de l'article, he introduït el paper d'Unicode en la codificació de les escriptures i les convencions i termes que s'hi utilitzen. Tot seguit, he comentat els dos tipus de termes sobre diacrítics que utilitzen les llengües d'escriptura romana: etimològics i creats per especialització semàntica. En l'últim tram, he comentat els diacrítics habituals i algunes qüestions sobre els seus termes, i he proposat solucions per a les denominacions quan m'ha semblat apropiat. ✨

- GAULTNEY, J. Victor (2002). *Problems of diacritic design for Latin script text faces* [en línia]. <<https://gaultney.org/jvgtype/typedesign/diacritics/ProbsOfDiacDesignLowRes.pdf>> [Consulta: 20 setembre 2020].
- KIT PAUL, Cristian (2019). «Romanian diacritic marks». A: *Kit blog* [en línia]. <<https://kitblog.com/2019/en/romanian-diacritic-marks/>> [Consulta: 19 setembre 2020].
- KRAVJANSZKI, Robert (2016). «The case of Hungary». A: MAŁECKA, Agnieszka; OSŁISŁO, Zofia (coord.) (2016). *The insects project: Problems of diacritic design for Central European languages* [en línia]. 2a ed. Katowice: Academy of Fine Arts, p. 36-61. <http://theinsectproject.eu/wp-content/uploads/2017/05/Insects-Project-2nd-edition_screen.pdf> [Consulta: 20 setembre 2020].
- Larousse = *Dictionnaire de français* ([2020]) [en línia]. París: Larousse. <<https://www.larousse.fr/dictionnaires/francais/>> [Consulta: 19 setembre 2020].
- MAŁECKA, Agnieszka; OSŁISŁO, Zofia (2016). «Polish diacritics: the history and principles of design». A: MAŁECKA, Agnieszka; OSŁISŁO, Zofia (coord.) (2016). *The insects project: Problems of diacritic design for Central European languages* [en línia]. 2a ed. Katowice: Academy of Fine Arts, p. 62-91. <http://theinsectproject.eu/wp-content/uploads/2017/05/Insects-Project-2nd-edition_screen.pdf> [Consulta: 20 setembre 2020].
- MELETIS, Dimitrios (2020). *The nature of writing: A theory of grapholinguistics*. Brest: Fluxus Éditions. També disponible en línia a: <<http://www.fluxus-editions.fr/gla3.php>> [Consulta: 10 desembre 2020].
- MESTRES, Josep M. [et al.] (2009). *Manual d'estil: La redacció i l'edició de textos*. 4a ed. Vic: Eumo; Barcelona: Universitat de Barcelona: Associació de Mestres Rosa Sensat: Universitat Pompeu Fabra. També disponible en línia a <<https://estil.llocs.iec.cat>> [Consulta: 10 desembre 2020].
- MESTRES, Josep M.; GUILLÉN, Josefina (2001). *Diccionari d'abreviacions: Abreviatures, sigles i símbols*. 2a ed. rev. i ampl. Barcelona: Enciclopèdia Catalana.
- OIEC = INSTITUT D'ESTUDIS CATALANS (2017). *Ortografia catalana* [en línia]. Barcelona: Institut d'Estudis Catalans.
- PARKES, Malcolm B. (1993). *Pause and effect: An introduction to the history of punctuation in the West*. Berkeley; Los Angeles: University of California Press.
- PUJOL, Josep M. (2001). «L'escriptura tipogràfica, 1. La font digital». *Llengua i Ús: Revista Tècnica de Política Lingüística* [en línia], núm. 22, p. 24-35. <<http://www.raco.cat/index.php/LlenguaUs/article/view/127881>> [Consulta: 19 setembre 2020].
- ROGERS, H. (2005). *Writing systems: A linguistic approach*. Malden, MA; Oxford: Blackwell.
- TAUBER, James K. (2019). «Character encoding of classical languages». A: BERTI, M. (ed.) (2019). *Digital classical philology: Ancient Greek and Latin in the digital revolution*. [S. l.]: De Gruyter Saur, p. 137-158.

Llista alfabètica de termes catalans introduïts a l'article

accent tancat o agut (també per a consonants)	contrabarra	signe bàsic o forma de base
anella	dièresi	signe diacrític
barra o ratlla transversal	graf	signe espaiat
barra inclinada	grafema	signe no espaiat
barra vertical	grafemàtica	sobreescrit o a sobre
breu	grafolingüística	sotaescrit o a sota
cercle puntejat	hàtxek o circumflex invertit	sotaratlla
circumflex	màcron	superposat
coma	majúscula	titlla
de combinació	minúscula	trenc
component grafemàtic	ogònek o cua	úmlaut
component grafètic	ortografia	vírgula
component ortogràfic	punt	
	punt volat	