THE EFFECT OF RECENT LOANWORDS ON SPELLING.
THE SPECIAL FUNCTIONS OF HYPHEN
IN CURRENT ORTHOGRAPHY

GABRIELA PANĂ DINDELEGAN[1]

Abstract. The author analyses the Romanian hyphen, one of orthographic marks that have the highest frequency and the most numerous values. The purpose of this article is to make an inventory of the general values of the Romanian hyphen, starting from the traditional ones and adding the latest ones, most commonly used in recent loan words and constructions. In the end, the author makes two prescriptive proposals, both regarding the use of Romanian hyphen in loan words and constructions, meant to regulate coherently and uniformly the new uses.

Keywords: orthography, punctuation mark, Romanian hyphen, values, prescriptive proposals.

1. PREAMBLE

The *hyphen* (also known in Romanian as *liniuță/linioară de unire/despărțire* 'linking/separating little line/dash') is an orthographic and punctuation mark with multiple and diverse uses in Romanian orthography. Apart from the traditional values, themselves very numerous, new values are added in relation with the use of hyphen in recent loanwords. The purpose of this synthetic presentation is to make an inventory of traditional values and to add the new ones, namely the ones in loanwords; at the end, two normative proposals will be put forth.

This article is dedicated to the highly esteemed Professor Camelia Stan, with the purpose of emphasising her contribution to the presentation of the orthographic and punctuation marks in DSL [The Dictionary of Language Sciences], a contribution which – a quarter of a century after its publication – is still intact in terms of scientific accuracy and relevance.

2. THE USE OF HYPHEN IN OLDER WORDS

Current normative papers identify the hyphen as an orthographic and punctuation mark (see DOOM[3]: 61–70), assigning it with the following important functions:

[1] Romanian Academy, g_dindele@yahoo.com.

● As an orthographic mark,

(i) it is used between words that are pronounced together (the same syllable), as an effect of phenomena in syntactic phonetics (synaeresis (*ne*[-]*am cunoscut* 'us.reciprocal=have met/we've met'), elision (*într*[-]*o seară* 'in=a evening/one night'), aphaeresis (*de*[-]*nțeles* 'of=understood/to understand') or as a result of the proclisis of a series of clitics (*l*[-]*ai văzut* 'him.cl=have seen/you saw him', *ne*[-]*am spus* 'us.reciprocal=have told/we told each other') or their enclisis (*văzând*[-]*o* 'seeing her', *spunându*[-]*i* 'telling him/her'); other non-stressed words are also pronounced in a single syllable (*că*[-]*s* 'that they are', *de*[-]*s* 'if they are', *mi*[-]*s* 'they are to me');

(ii) it is used in words derived with the prefixes *ne-*, *re-* (adding a negative, respectively iterative meaning to the word it attaches to), showing the deletion of the initial vowel of the root and the one-syllable pronunciation of the prefix and the root (*ne*[-]*nțeles* 'not understood', *re*[-]*ncărcat* 'recharged');

(iii) it is used in old compounds of various types: nouns (*mamă*[-]*soacră* 'mother-in-law', *floare*[-]*de*[-]*colt* 'edelweiss', *rochița*[-]*rândunicii* 'bindweed', *nu*[-]*mă*[-]*uita* 'forget-me-not', *lasă*[-]*mă*[-]*să*[-]*te*[-]*las* 'stick-in-the-mud', adjectives (*galben*[-]*verzui* 'yellow greenish', *nou*[-]*nouț* 'brand new'), adverbs (*târâș*[-]*grăpiș* 'with great difficulty'), showing a weaker degree of compounding than the compounds that are spelt in one word (*untdelemn* 'lit. wood butter/oil', *douăzeci* 'lit. two tens/twenty', *doisprezece* 'lit. two over ten/twelve', *dumneata* 'lit. your sirness/you.formal', *altcineva* 'lit. other someone/someone else');

(iv) it marks the different pronunciation tempo (fast or slower), transforming the hiatus in syntactic phonetics, in the case of fast tempo, in an ascending diphthong (*de abia* [de.a] vs *de-abia* [dea] 'barely', *de atunci* [de.a] vs *de-atunci* [dea] 'since then').

● As a punctuation mark,

(i) it is used between repeated words that form one unit (*se mișcară râuri*[-]*râuri* 'they moved like one river after another', M. Eminescu; *gata*[-]*gata să cadă* 'very nearly ready to fall');

(ii) it is used between juxtaposed numerals, expressing a numeric approximation (*au așteptat patru*[-]*cinci ore* 'they waited for four-five hours', *au alegat patru*[-]*cinci kilometri* 'they ran for four-five km');

(iii) it is used between words that set limits in space and time, expressing the path, respectively the duration (*cursa București*[-]*Viena* 'the Bucharest-Wien flight', *perioada 2*[-]*12 mai* 'the period/during 2-12 May').

● As a conventional mark,

It is used differently from one professional field to another. In linguistics, for instance, it is used:

(i) to mark the elements lacking autonomy, often bound inflection components (see, for example, verbal suffixes *-ând*, *-at*, the vocative desinence *-o*, the noun plural desinence *-uri*, the desinence and the article *-ul*);

(ii) as a spelling marker to split words into syllables and to split a long(er) word at the end of a line[2] (*car*[-]*te* 'book', *jert*[-]*fă* 'sacrifice', *in*[-]*e*[-]*gal* 'unequal').

---

[2] To distinguish between *split into syllables* vs *split at the end of a line*, which share some features, but are not identical, see in DTG, s.v. *despărțire* 'split'.

I will insist on the special role of hyphen in compounding and zero-derivation.

● In older compounds,

(i) the hyphen distinguishes between compounds that are strongly connected, and do not accept separation through a hyphen (*acasă* 'at home', *demult* 'a long time ago', *altfel* 'different', *dânsul* 'he', *untdelemn* 'vegetable oil'), and those less bound, which are hyphenated (*mama-soacră* 'mother-in-law', *dis-de-dimineață* 'early in the morning', *după-amiază* 'afternoon', *astă-vară* 'last summer'), thus the hyphen marks the different degrees of compounding (closed vs hyphenated compounds);

(ii) the hyphen separates compounds from analysable phrases (see *lup-de-mare* 'wolf fish', a compound meaning a type of sea animal, vs *lup de mare* 'sea wolf', an analysable phrase indicating a metaphor for a sailor, *după-amiază* 'afternoon', adverbial phrase (*Doarme după-amiaza* 'He sleeps in the afternoon') vs *după amiază* 'after noon', preposition + noun (*Până la amiază e bine, dar către amiază și după amiază se face prea cald* 'Until noon the weather is fine, but towards noon and after noon it gets too hot').

● In zero-derivation, the hyphen is used for special cases, integrating both in inflection and syntax (by attaching the corresponding flectives) any word in meta-language undergoing nominalisation (*la-uri* 'ats', *pe-uri* 'ons', *și-uri* 'ands', *că-uri* 'thats', *nu-uri* 'nos', *i-uri* 'is'; e.g.: *Sunt prea multe bla-bla-uri* 'There are too many blablas', *Lipsesc pe-urile* 'The ons are missing'); see also more "scholarly" unique autonymic use: „*Mi s-a părut, am mai zis, un număr slab, alcătuit din gemete ultime, din <Ce să mai...>-uri lehămetite*" 'I thought it was, as I said before, a weak performance, made of ultimate moans, of disgusting *what to do*s' (Paul Goma). Note here that not all zero-derivations are hyphenated, being used with only those regarding nominalisations in meta-language (compare with frequent zero-derivations of the type: *aproapele* 'close.one.DEF', *binele* 'good.DEF', *nimicul* 'nothing.DEF', *verdele* 'green.DEF', *frumosul* 'beautiful.DEF', *intrândul* 'entry-ând.GER.DEF').

3. THE USE OF HYPHEN IN RECENT LOANWORDS

In recent loanwords, the hyphen is used with four types of values (3(i) –3(iv)):

(i) it renders the use as in the language of origin (see from English *baby-sitter*, *drive-in*, *duty-free* or French *deux-pièces*, *en-titre*, *tête-à-tête*, with a hyphen, but unhyphenated, as in their language of origin (English), in *cheeseburger*, *cheesecake*, *workshop*);

(ii) it is used to connect Romanian flectives to roots of recent loanwords, in which case it functions as a marker of the different degree of adjustment/localisation of the loan, showing full adjustment, if unhyphenated (see *consomméuri* 'consommé.pl', *partyuri* 'parties', *standuri* 'stands', *targeturi* 'target.pl'), vs partial adjustment, if hyphenated (see *cookie-uri* 'cookies', *deadline-uri* 'deadlines', *like-uri* 'like.pl on social media', *mouse-uri* 'computer mice', *smoothie-uri* 'smoothies');

(iii) rarely, it has the graphic role to distinguish between two words (*e-mail* 'electronic mail' vs *email* 'enamel');

(iv) it marks the inflectional and syntactic integration (by attaching the appropriate flectives) of acronymic compounds (*CNP-ul* 'the SSN', *NUP-uri* 'not to commence criminal prosecution.pl', *PIB-uri* 'GDPs', *PIN-uri* 'PINs', *PSD-ul* 'the social democrat party', *RMN-uri* 'MRIs', *TAB-ul* 'the TAB', *USB-uri* 'USBs', etc.) or by attaching affixes to acronymic compounds (*IT-ist* 'IT person', *non-UE* 'non-EU', *pro-UE* 'pro-EU');

(v) it is used with neologic compounds with two or three components connected by the compounding vowel "o" (see the examples with two components: *economico-financiar* 'economic and financial', *financiaro-bancar* 'financial and banking', *medico-legal* 'medical and legal/forensic', *medico-farmaceutic* 'medical and pharamceutical', *politico-administrativ* 'political and administrative', and with three components: *buco-maxilo-facial* 'oral and maxillofacial'). The presence of the vowel "o" is the sign of losing the autonomy of the first (two) component(s) (the components ending in "o" can never be used by themselves), which calls for the use of the hyphen[3]; it is used with neologic compounds with various types of structures, many of which directly translated from other languages, keeping the hyphen from the language of origin (such as: *dicționar-tezaur* 'thesaurus-dictionary', *dosar-plic* 'evenlope-folder/file folder', *firmă-căpușă* 'company-tick/parasite company', *grup-țintă* 'target group', *mamă-natură* 'mother nature', *master-plan* 'master plan', *termen-limită* 'deadline', *timp-record* 'record time').

4. COMMENTS

There are loans with two (or three) hyphens, marking the two functions under (i) and (ii) (see *box-office-ul* 'box office.DEF', *reality-show-uri* 'reality shows', *tête-à-tête-uri* 'tête-à-têtes') or with one taken from the language of origin, and the flective *-uri* (to mark the plural) bound directly (*exit-polluri* 'exit polls', *sex-shopuri* 'sex shops', *ski-passuri* 'ski passes', *ski-jeturi* 'ski jets').

From the same pattern with the compounds, where the presence vs absence of the hyphen indicates the different degree of bounding the components (see §2 above), DOOM[2-3] requires conventionally the rule to use the hyphen with integration flectives, especially with the enclitic definite article, and the plural desinence *-uri*, namely: the flective attaches directly in case the root ends in a letter that has a corresponding pronounced sound (*frappéuri* 'frappes', *screeninguri* 'screenings', *statusuri* 'statuses', *stenturi* 'stents', *swinguri* 'swing.PL', *șlemuri* 'slam.PL', *targeturi* 'target.PL'), but the same flective attaches with a hyphen in case the root of the loan word ends with a letter that is pronounced differently or not pronounced in Romanian (*badge-uri* 'badges', *cookie-uri* 'cookies', *deadline-uri* 'deadlines', *debate-uri* 'debates', *like-uri* 'likes (on social media)', *mouse-uri* 'computer mice', *online-ul* 'online.DEF', *partie-uri* 'parties', *review-uri* 'reviews', *scotch-uri* 'scotch' or 'adhesive tape', *search-uri* 'search.PL', *shake-uri* 'shake.PL', *slide-uri* 'slide.PL', *smoothie-uri* 'smoothies', *smartphone-uri* 'smartphones', *tête-à-tête-uri* 'tête-à-têtes').

Another new function is to mark to subclass of acronymic compounds. In the case of acronyms, the hyphen is used together with a different orthographic convention: marking with capital letters the initials of the compounding elements. The two orthographic conventions are in place for both recent acronyms (*FMI-ul* 'IMF.DEF', *ISBN-uri* 'ISBN.PL',

---

[3] Note the inconsistent solution provided in DOOM[3], which, under the same circumstances of connecting the components with "o", uses the hyphen (as in the examples provided above, taken from DOOM[3]) or, on the contrary, does not use the hyphen, in examples such as: *sociocultural* 'social and cultural', *socioeconomic* 'social and economic', *sociolingvistic* 'social and linguistic, *socioprofesional* 'social and professional', *corticosteroid* 'corticosteroid', *insulinodependent* 'insulin-dependent', *imunodeficiență* 'immunodeficiency', *toxoplasmoză* 'toxoplasmosis', *aerosoloterapie* 'aerosol therapy', *balneofizioterapie* 'balneo-physiotherapy', etc.

RMN-uri 'MRI.PL', *SMS-uri* 'SMS.PL', *STAS-uri* 'state standard.PL', *TVA-ul* 'VAT.DEF', *SWIFT-ul* 'SWIFT.DEF', etc.), newly included in DOOM³, and less recent acronyms (see: *CFR-ul* 'Romanian Railways.DEF', *PMR-ul* 'Romanian Labourer's Party.DEF', *RFG-ul* 'Federal Republic of Germany.DEF', *RDG-ul* 'German Democratic Republic.DEF', *URSS-ul* 'Union of Soviet Socialist Republic.DEF'). Rarely, the normative dictionary allows for free variation, with two spellings, one with a capital letter and a hyphen, while the other – with a small case letter and unhyphenated, marking, for the latter, the loss, in user's mind, of the special status of the word in use. (*VIP-uri/vipuri* 'VIPs'). In the case of words derived from acronymic roots, the hyphen is used for recent derived words, deemed as less integrated (of the type *IT-ist* 'IT person' or *pro-UE* 'pro-EU'), but it is not used for words derived from older acronyms, which are considered as already integrated (such as *ceferist* 'railway person', *pesedist* 'PSD person', *pesedizare* 'making something PSD-like', *uslaș* 'person in/related to Antiterrorist Fight Special Unit', *utecist* 'member of the Young Communists Unit').

5. CONCLUSIONS AND NORMATIVE PROPOSALS

Aside from the values identified in traditional orthography, recent normative papers add new values, with the hyphen being used to mark the latest loans, either by providing information on the degree of loan integration, or by marking the special class of acronymic compounds, or by marking the neologic compounds bound with the use of the compounding vowel "o". For the first two classes, the hyphen marks certain special types of words, felt by the speaker to belong to a different linguistic code than the one in use[4]. For the latter, the hyphen marks those compounds whose components are not autonomous.

In the orthography of contemporary Romanian, with the recent values related to the class of loanwords, with a high productivity and frequency, the use of the hyphen increases implicitly.

The hyphen, as a marker of mixing traditional and recent orthographic values, has the effect to increase the degree of ambiguity of punctuation marks, bringing in difficulties in contextual establishing of the value and, implicitly, difficulties to use effectively (users frequently ask themselves: "Which is the correct spelling in Romanian: *check-inul*, *check-in-ul* or *checkinul* 'checkin.def'?").

The first proposal is to generalise the use of the hyphen in the case of neologic compounds connected with the compounding vowel "o", regardless of the number of components (*socio-economic* 'social and economic', *struțo-cămilă* 'ostrich-camel/hybrid; neither one, nor the other', *balneo-fizio-terapie* 'balneo-physiotherapy'), except for the prefixes and prefix-like elements ending in "o" (*aerodynamic* 'aerodynamic', *ecosystem* 'ecosystem', *etnobotanice* 'ethnobotanicals', *eurobarometru* 'Eurobarometer').

The second proposal is to generalise the use of the hyphen in case of attaching the integration/localisation flectives (the desinence and the definite article *-ul* and the plural desinence *-uri*), regardless of the pronunciation of the final sound of the root (*selfie-uri* 'selfies', *bordeaux-uri* 'bordeauxes', *like-uri* 'likes (on social media)', but also *link-uri* 'links', *poncho-uri* 'ponchos', *tiramisù-ul* 'tiramisu.DEF').

---

[4] For the quality as "special" words, felt as belonging to a different linguistic code, see Pană Dindelegan 2002/[2018]. In the cited article, this special class includes the zero-derivations in meta-language, be they hyphenated or not (*Lipsesc și-urile* 'The *și*s "ands" are missing'); see §2 above.

REFERENCES

DOOM[2], I. Vintilă-Rădulescu (coord.), 2005, *Dicționar ortografic*, *ortoepic și morfologic al limbii române*, 2nd edition, reviewed and added, București, Editura Univers Enciclopedic.

DOOM[3], I. Vintilă-Rădulescu (coord.), *Dicționar ortografic*, *ortoepic și morfologic al limbii române*, 3rd edition, reviewed and added, București, Editura Univers Enciclopedic Gold, 2022.

DSL, A. Bidu-Vrănceanu, C. Călărașu, L. Ionescu-Ruxăndoiu, M. Mancaș, G. Pană Dindelegan, *Dicționar de științe ale limbii*, Editura Științifică, 1997; 2nd edition, added, Editura Nemira, 2001, 2005.

DTG, G. Pană Dindelegan (coord.), *Dicționar de termeni gramaticali și concepte lingvistice conexe*, București, Univers Enciclopedic, 2023.

Pană Dindelegan, G. 2021, „Comportamentul acronimelor în DOOM[3]", *Limba română*, LXX, 3–4, 447–454.

Pană Dindelegan, G., 2002/[2018], „Formații substantivale recente și rolul *clasificatorilor* în actualizarea contextuală", in G. Pană Dindelegan (ed.), *Aspecte ale dinamicii limbii române actuale*, București, Editura Universității din București, 31–46, republished in Gabriela Pană Dindelegan, *Dinamică și variație. Studii lingvistice*, Bucharest, Editura Academiei Române.

Vintilă-Rădulescu, I., 2007, „Beția de cratime", in C. Stan, R. Zafiu, A. Nicolae (eds), *Studii lingvistice. Omagiu profesoarei Gabriela Pană Dindelegan la aniversare*, București, Editura Universității din București, 457–466.