Lexical Sources of Phonological Alternation: a role for Voting Bases

Canaan Breiss[1] & Donca Steriade[1]

[1]*Massachusetts Institute of Technology (USA)*
canaan@mit.edu, steriade@mit.edu

Summary: We explore the lexicon's influence on the shape of new derivatives, using data from Romanian derived verbs. The key notions in what follows are: (i) *Derivatives* (Ds), forms created by affixation to a word or root; (ii) *local Bases* (B_L), exponents of an immediate syntactic constituent of D; and (iii) *remote Bases* (B_R), forms that are lexically related to D, but distinct from B_L. For example, the English *sỳntactíc-ian* has *syntáct-ic* as its B_L; and *sýntàx* as a B_R.

We start from the observation that in some morphological systems, the phonology of some Ds is determined not by the shape of their B_Ls, as expected ([1]), but by that of a B_R. The stress of trochee-initial *sỳntactícian*, for instance, is a version of the trochaic B_R *sýntàx*, and different from that of its iamb-initial B_L *syntáctic*. An extensive pattern of B_R-influenced Ds has been reported for Romanian in [2]. The present contribution is an effort to extensively verify one aspect of that study's findings, to explore a different analysis, and to clarify through a *wug*-test if the pattern we found in the Romanian lexicon mirrors native speaker preferences.

Romanian verbs and palatalization: Romanian derived verbs are formed by suffixes -a, as in [ɨm-pʌjenʒen-á] 'to cover in spiderwebs', B_L [pʌjánʒen] 'spider'; -í, e.g. [ɨm-pʌdur-í] 'to cover in forests,' from [pʌdúre] 'forest'; and -uí, e.g. [ɨn-vʌl-uí] 'to veil', B_L [vʌl] 'veil'. While free in general, the suffix choice is restricted after velars, a fact due to the process of velar palatalization, which turns velars [k, g] into palato-alveolars [tʃ, dʒ] before front vocoids: [e, i, j]. A first analysis of these restrictions [2] builds on two facts. First, like most front suffixes in Romanian, the verbalizer -í, triggers velar palatalization: e.g. [ɨn-furtʃ-í] 'to bifurcate (intrans.)' from B_L [fúrk-ʌ] 'fork'. Second, the verbalizer -í attaches freely to velar-final bases like [fúrk-ʌ], causing velar palatalization *iff these bases already possess a palatalized allomorph in their inflectional paradigm*. Thus, the plural of [fúrk-ʌ] is [fúrtʃ-i]. If the base lacks a palatalized plural – because it lacks *any* plural, or because its plural suffix does not trigger palatalization – the verbalizer -í is avoided, and one of the alternatives, -á or -uí, is used instead. Thus, [tsark] 'fenced space', plural [tsárk-uri], no palatalization, gives rise to [ɨn-tsʌrk-uí] 'to fence in'. Nouns like [tsark], with invariant velars in their inflectional paradigm, rarely give rise to /-í/-suffixed verbs: forms like *[ɨn-tsʌrtʃ-í], with palatalized root allomorph found just in the derived verb, not in its base N, are rare and marginal, and no verbs like *[ɨn-tsʌrk-í], with surface [k-i] across the suffixal boundary, exist at all. Thus both markedness, i.e. avoidance of [ki], and faithfulness play a role in selecting the verbal suffix and in velar palatalization. Since verbs like [ɨn-furtʃ-í] 'bifurcate' don't refer to a plurality of participants, the B_L of the derived verb is not its plural. Thus plural [furtʃ-i] 'forks,' is a B_R since its presence in forming the verb is not syntactically justified. From these assumptions, it follows that the selection of the verbalizer /-í/ and the application of palatalization in forms derived with this /-í/ are licensed by a B_R, the palatalized plural stem of the base N.

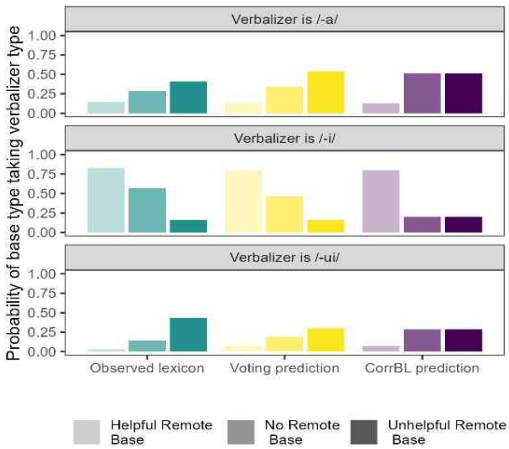
Analyses: We explored two analyses. One ([2, 3]) uses four ingredients: (i) a markedness constraint *KE, banning velars before front vocoids; (ii) a faith constraint ID B(ASE)- D(ERIVATIVE) requiring D's stem to correspond in its consonantism to that of *some* Base, B_L or B_R; (iii) a preference for D's stem to be in correspondence to its B_L (Corr B_L); and (iv) a preference to use the -í suffix in change-of-state intransitives (USE /-í/). The ranking ID-BD >> *KE >> USE /-í/ >> CORRB_L models the ban on palatalized items like *[ɨn-tsʌrtʃ-í] (given invariant [tsark]); the avoidance of /-í/ after bases with such invariant velars; and the default preference for the verbalizer /-í/ in derived intransitives, as in [ɨn-furtʃ-í]. A similar model invoking violable CorrB_L proves useful elsewhere ([3, 4]). Tableau A (bottom p. 2) presents a schematic analysis with select candidates in a standard categorical OT.

An alternative model, used in [6], starts from the idea that each form in a lexical paradigm exerts some attraction on the stem of the D, independently of improvements in D's markedness. There are multiple active faithfulness constraints (ID$_L$, ID$_R$), each expressing a preference for

correspondence between D's stem and a specific B, either the B_L or a B_R. Markedness (*KE), ganging up with these constraints, can result in the selection of B_R-based stems if a B_R improves the stem's markedness. Importantly, however, the two faithfulness constraints can gang up: when the B_L and B_R do not improve markedness, they should jointly discourage palatalization even more than in nouns lacking any B_R at all: therefore, they should select non/-í/ even more often.

Although the two theories make identical predictions about forms with palatalized B_Rs (e.g. pl. [fúrtʃ-i], called *helpful* B_Rs here) and about forms lacking any B_R (e.g. [vlʌg-uí] 'exhaust' based on singular-only [vlágʌ] 'force'), they make different predictions about the effect of B_Rs that resemble the B_L, e.g. [tsark], pl. [tsark-uri]. We call plurals like [tsark-uri] *unhelpful* B_Rs: if *two* bases 'voting' to preserve [k], thus *against* palatalization, have more effect than just one, such nouns could inhibit palatalization in the derived verb more than plural-less [vlágʌ].

Evidence for Voting Bases in the Romanian lexicon: Drawing on dictionary data (dexonline.ro), we find that the distribution of verbalizer allomorphs is sensitive not only to the presence of helpful B_Rs (with -í more likely in cases where the paradigm already has a palatal-final stem allomorph), but also that B_Ls with an unhelpful B_R (that is, a non-palatalized plural) are even less likely to take the non-í verbalizers, -á or -uí, which do not trigger palatalization, than nouns lacking any B_R. These data are plotted in the figure below, with predictions from the two theories derived from a Maximum Entropy grammar ([6]) implementing each analysis fully and faithfully (not shown). A χ-squared test with one degree of freedom finds that the CORRB_L model fits (right) fits the lexical data (left) significantly less well than the Voting Bases model (center); $p < 0.01$. An experiment with Romanian native speakers is in progress to test the generality of this pattern in words that lack a lexicalized corresponding verb. If speakers' allomorph selection is sensitive to the presence of both helpful and unhelpful Remote Bases, we will take this to support the lexical evidence we have advanced here to favor the Voting Bases model.



References

[1] Chomsky, N., & Halle, M. (1968). The sound pattern of English.

[2] Steriade, D. (2008). A pseudo-cyclic effect in Romanian morphophonology. *Inflectional identity*, *18*, 313-360.

[3] Steriade, D., & Stanton, J. (2020). Productive pseudo-cyclicity and its significance. *Talk at LabPhon*, *17*.

[4] Steriade, D., & Yanovich, I. (2015). Accentual allomorphs in East Slavic: inflection dependence. *Understanding Allomorphy*, 254-314.

[5] Breiss, C. (2021). *Lexical Conservatism in phonology: theory, experiments, and computational modeling*. UCLA

[6] Goldwater, S., Johnson, M., Spenader, J., Eriksson, A., & Dahl, Ö. (2003, April). Learning OT constraint rankings using a maximum entropy model. In *Proceedings of the Stockholm workshop on variation within Optimality Theory* (Vol. 111, p. 120).

Tableau A: CORRB_L analysis (select candidates); H = Harmony, candidate subscripts indicate Base-Derivative correspondence.

B_L [furk]$_L$-ʌ B_R [furtʃ]$_R$-i	ID-BD	*KE	USE /-i/	CORRB_L	B_L [tsark]$_L$ B_R [tsark]$_R$-uri	ID-BD	*KE	USE /-i/	CORRB_L
D_1: [furtʃ]$_L$ -í	*				D_1: [tsʌrtʃ]$_L$-í	*			
D_2: [furk]$_L$ -í		*			D_2: [tsʌrk]$_L$-í		*		
D_3: [furtʃ]$_R$ -í				*	D_3: [tsʌrtʃ]$_R$-í	*			*
D_4: [furtk]$_R$ -í	*	*		*	D_4: [tsʌrk]$_R$-í		*		*
D_5: [furk]$_L$ -uí			*		D_5: [tsʌrk]$_L$-ui			*	

Tableau B: Voting Bases analysis (select candidates) for illustration of violations and ganging only; H = Harmony

B_L [furk]$_L$-ʌ B_R [furtʃ]$_R$-i	ID-B_L	ID-B_R	*KE	USE /-i/	H	B_L [tsark]$_L$ B_R [tsark]$_R$-uri	ID-B_L	ID-B_R	*KE	USE /-i/	H
	1	1	1	1			1	1	1	1	
D_1: [furtʃ] -í	*				1	D_1: [tsʌrtʃ]-í	*	*			2
D_2: [furk] -í		*	*		2	D_2: [tsʌrk]-í			*		1
D_3: [furk] -uí		*		*	2	D_3: [tsʌrk]-ui				*	1