CYCLICITY WITHOUT CONTAINMENT IN ROMANIAN PERFECTS

DONCA STERIADE

MIT

Abstract. The Romanian perfect exhibits a form of directional paradigm uniformity: verbal perfect forms adopt the stress and segmental characteristics of the perfect participle. An analysis of this pattern of paradigmatic identity is proposed, which has broader implications for the theory of the phonological cycle.

Keywords. cycle, paradigm uniformity, Base-Derivative correspondence

1 Introduction

This study analyzes a pattern of similarity among Romanian verb forms, previewed in (1).

(1) Stem identities in the Romanian perfect[1]

		PERFECT			NON-PERFECT	
		3sg. perf.	1st pl. perf.	Participle	3.sg. indic.pres.; gerund	
(a)	'fall'	[*kʌzú*]	[*kʌzú*]-*rʌm*	[*kʌzút*]	[*kád*]-*e*	[*kʌz*]-*índ*
	'burn'	[*árs*]-*e*	[*árs*]-*e*-*rʌm*	[*árs*]	[*árd*]-*e*	[*arz*]-*índ*
(b)	'hold'	[*ʦinú*]	[*ʦinú*]-*rʌm*	[*ʦinút*]	[*ʦín*]-*e*	[*ʦin*]-*índ*
	'put'	[*pús*]-*e*	[*pús*]-*e*-*rʌm*	[*pús*]	[*pún*]-*e*	[*pun*]-*índ*

The left side of table (1) shows four sets of perfect forms with identical stress and segmentally similar stems. The perfect participles display an unpredictable difference between an *-s* suffix, as in *ars,* and a *-t* suffix, as in *kʌzút.* Some properties of participial stems are transmitted to the tensed perfect: witness *pús, púse* vs. *ʦinút, ʦinú*. The non-perfect forms, shown on the right side of (1), have stress and segmentals that differ from the perfect. In non-perfect verb forms, stress alternates and generally abides by the constraints applicable to morphologically simple forms, as argued below. The perfect differs.

I will show that the stress and segmental composition of all perfect forms can be predicted from the perfect participle, which follows the accentual pattern of simple words. This requirement of stem identity that governs the entire paradigm

[1] Data sources for this study include Lombard & Gâdei 1981; Pană Dindelegan (ed.) 2013; and dexonline.ro, an extensive, searchable lexical database. The IPA symbols used here are mapped to Romanian graphs as follows: [ʌ, i̯, ʃ, ʦ] = <ă, î, ş, ţ>, [ʧe, ʧi> = <ce, ci >, [ke, ki] = <che, chi>.

of the Romanian perfect is analyzed here as an instance of cyclic inheritance, in which the participle functions as the Base (= cycle n) and tensed perfect verbs are generated as its Derivatives (= cycle n+1). Unlike in standard cyclic cases, the Base in these perfect paradigms is not contained, morphologically or syntactically, in its Derivatives. A modified theory of the cycle (Stanton and Steriade 2018) does justice to this and comparable other cases.

2 Romanian perfects

Romanian perfect paradigms consist of a participle (PPf) and three sets of verbal forms (VPf): a simple perfect comparable in its aspectual value to the French *passé simple*, a pluperfect and an analytic perfect, comparable to the *passé composé*, consisting of the PPf plus an auxiliary. Of interest here are the synthetic verb forms, the simple perfect and pluperfect.

Table (2) presents two complete paradigms, accompanied by a morphological parse of the tensed perfects, one for each of the types shown. '$\text{PERF}_{1\text{-}3}$' is a reference to different perfect exponents; *se* is the pluperfect suffix, *rʌ* is a marker of plurality in perfects; AGR refers to all other person-number endings[2].

[2] The function of *-e-* in sigmatic perfects is to block impossible C clusters, as in *ars-e-se*, and to promote anti-homophony, as in 3rd sg. *ars-e,* which would otherwise merge with *ars*, the PPf.

(2) Identities in the stems of perfect forms of two strong verbs

infinitive	*kʌd-eá* ROOT-THEMEV 'fall'		*árd-e* ROOT-THEMEV 'burn'	
PPf	*kʌz-ú-t* ROOT-PERF$_1$-PERF$_2$		*ár-s* ROOT -PERF$_3$	
VPf:	1	[*kʌz-ú*]-*j* [*kʌz-ú*]-*rʌ-m*	[*ár-s*]-*e-j*[3]	[*ár-s*]-*e-rʌ-m*
simple	2	[*kʌz-ú*]-*ʃj* [*kʌz-ú*]-*rʌ-tsj*	[*ár-s*]-*e-ʃj*[3]	[*ár-s*]-*e-rʌ-tsj*
perfect	3	[*kʌz-ú*] [*kʌz-ú*]-*rʌ*	[*ár-s*]-*e*	[*ár-s*]-*e-rʌ*
VPf:	1	[*kʌz-ú*]-*se-m* [*kʌz-ú*]-*se-rʌ-m*	[*ar-s*]-*é-se-m*	[*ar-s*]-*é-se-rʌ-m*
pluperfect	2	[*kʌz-ú*]-*se-ʃj* [*kʌz-ú*]-*se-rʌ-tsj*	[*ar-s*]-*é-se-ʃj*	[*ar-s*]-*é-se-rʌ-tsj*
	3	[*kʌz-ú*]-*se* [*kʌz-ú*]-*se-rʌ*	[*ar-s*]-*é-se*	[*ar-s*]-*é-se-rʌ*
ROOT- PERF$_1$-(se$_{PLUPERF}$)-(rʌ $_{PLURAL}$)-AGR ROOT- PERF$_3$-e- (se$_{PLUPERF}$)-(rʌ $_{PLURAL}$)-AGR				

Like (1), the data in (2) shows that stress is on the same stem syllable in VPfs as in the participle. (2) also illustrates an exception to the general pattern of accentual identity in the perfect: stress changes in verbal forms to avoid identical *sese* strings. Thus, pluperfects like *arsésem* contain *sése,* not **ársesem*, the expected form given *árserʌm.* This limited deviation from identity between the PPf and the VPf forms will play a role in the analysis. (1) and (2) also show that perfect stems

[3]For this paradigm type, the 1st and 2nd sg forms admit a variant with final stress: *arséj, arséʃj*. This is the only option reported by Zafiu (2013, 33). The variant with root stress, *ársej, árseʃj,* is reported elsewhere (Lombard & Gâdei 1981,135) and has been encountered by the present writer numerous times, including in rimes like *rʌmásej* (remain-Vpf-1sg) - *kásej* (house-Datsg), where the unreduced vowels and the requirements of riming identity guarantee the location of stress.

are segmentally identical, aside from the PPf suffix *-t*, in forms like *kʌz-ú-t̲*. This *-t* is systematically missing in VPf forms for reasons explored below.

The verbs in (1-2) are two of the cca 250 strong verbs originating in the 2nd and 3rd Latin conjugations. The vast majority of Romanian verbs descend from the Latin 1st and 4th conjugations, in *-ā, -ī*. As in Latin, the perfects of these verbs, two of which appear in (3), preserve the theme vowel of the present. By contrast, the strong verbs lose their present theme vowels in the perfect.

(3) Perfect forms of two weak verbs: 'hear' and 'praise'

infinitive	*auz-í* ROOT-THEMEV 'hear'		*lʌud-á* ROOT-THEMEV 'praise'	
PPf	*auz-í-t* ROOT-THEMEV -PERF$_2$		*lʌud-á-t* ROOT - THEMEV-PERF$_2$	
VPf:	1	[*auz-í*]-*j* [*auz-í*]-*rʌ-m*	[*lʌud-á*]-*j*	[*lʌud-á*]-*rʌ-m*
simple	2	[*auz-í*]-*ʃj* [*auz-í*]-*rʌ-tsj*	[*lʌud-á*]-*ʃj*	[*lʌud-á*]-*rʌ-tsj*
perfect	3	[*auz-í*] [*auz-í*]-*rʌ*	[*lʌud-ʌ́*]	[*lʌud-á*]-*rʌ*

Weak verbs like those in (3) follow in the perfect the same patterns of identity as those seen in (1-2). The unique exception from identity is lawful. It appears in the 3rd sg.pf. of verbs like *lʌud-á*, which are realized with a change of theme vocalism, [*lʌud-á*]→ [*lʌud-ʌ́*], to avoid homophony with the imperfect 3rd sg. [*lʌud-á*].

The stem identity seen in (1-3) is limited to the perfect. Non-perfect stems differ segmentally from each other and from the perfect. Their stress alternates:

(4) Alternations in non-perfect forms: ‘fall’ and ‘hear’

pres.ind.	1	*kád*	*kʌd-é-m*	*a.úd*	*a.uz-í-m*
	2	*káz-j*	*kʌd-é-tsj*	*a.úz-j*	*a.uz-í-tsj*
	3	*kád-e*	*kád*	*a.úd-e*	*a.úd*

I show next that the accentual mobility in the non-VPf forms is the effect of rankings holding generally in the language. What will have to be explained is the contrast between the accentual invariance characteristic of the perfect and the regular accentual mobility observed outside the perfect.

3 The stress system outside the perfect

Outside the perfect, stress is largely predictable in Romanian. Most words are stressed on the penult, unless the final is heavy, in which case final stress is the rule: see constraints and rankings in (5.a-c). Three additional options are attested, but disfavored[4]: (i) antepenult stress in words with light finals and penults, like *kámerʌ* ‘room’; (ii) penult stress when the final is heavy, e.g. *úmʌr* ‘shoulder’; and, even less commonly, (iii) final stress on a light syllable, e.g. *halvá* ‘halvah’. I analyze all these deviations from the general pattern by letting a lexically indexed

[4] Vasiliu 1965 provides lexical counts on stress in roots, from which stress in fully inflected words can be inferred. See also Chițoran 2001, Steriade 1985.

constraint, IDENTSTRESS$_{LEX}$ (Pater 2000) outrank some of the M constraints, as in (5.e). Pre-antepenult stress and antepenult stress in words with closed penults or finals are virtually impossible in native words and nativized loans. I use the conjunction of WSP and *LAPSER to analyze this, (5.g-h). I use grid-based constraints (Gordon 2002) and constraint conjunction, in (5.g), but neither is critical to the main argument.

(5) Stress constraints and rankings for mono-morphemes

a. NONFINALITY (NF): one * for any final stress.

b. WEIGHT-TO-STRESS (WSP): one * for every stressless heavy syllable.

c. WSP >> NONFINALITY >> STRESSRIGHT

d. *LAPSER: one * for each final pair of stressless syllables, 00#

e. IDENTSTRESS$_{LEX}$ >> *LAPSER, WSP# >> NONFINAL >> IDENT STRESS

f. *EXTLAPSER: one * for 000#.

g. *LAPSER-WSP one * any string that violates both LAPSER and WSP.

h. *EXTLAPSER, *LAPSER-WSP >> IDENT-STRESS$_{LEX}$ >> WSP, *LAPSER

In non-perfect forms, the rankings proposed above cause accentual alternations, as in *mút-ʌ, véd-e* ‘displaces/sees’ vs. *mut-ʌ́m, ved-ém* ‘we change/we see’.

(6) a. *mút-ʌ* ‘changes’ b. *mut-ʌ́m* ‘we change’

Root: mut-	NONFIN	STRESSR
mutʌ́	*!	
☞ mútʌ		*

Root: mut-	WSP	NONFIN
☞mutʌ́m		*
mútʌm	*!	

In verb roots with lexical stress on a non-final root syllable, *LAPSER-WSP causes alternations between antepenult stress, in V-final forms like *mʌ́tur-ʌ,* and final stress in words ending in VCC_0, like *mʌtur-ʌ́m, mʌtur-íńd* .

(7) a. *mʌ́tur-ʌ* ‘he/she sweeps’ b. *mʌtur-ʌ́m* ‘we sweep’

mʌ́tur-	IDSTRESS$_{LEX}$	*LAPSER
mʌtúrʌ	*!	
☞mʌ́turʌ		*

mʌ́tur-	*LAPSER-WSP	IDSTRESS$_{LEX}$
mʌ́turʌm	*!	
mʌtúrʌm		**!
☞mʌturʌ́m		*

The distribution of stresses in Romanian is identical to that of Spanish, as analyzed by Harris 1983. Our analyses differ in an interesting way. Harris uses segment extrametricality to generate two of the three marked stress patterns of Spanish-Romanian. In his analysis, antepenult stress results from a final vowel being extrametrical; penult stress in words with heavy finals is due to an extrametrical consonant. When applied to Romanian, Harris’s proposal will parse *kámerʌ* as *káme.r-*<ʌ> - with angle brackets marking extraprosodicity - and *úmʌr* as

úmʌ<r>. When nothing is extrametrical, the default pattern emerges. The lexical exceptions to stress in verbs bear on the difference between analyses. Verb roots are frequently followed by monosyllabic endings. On our analysis, any syllable in the root can be lexically stressed. This marked stress is protected, modulo higher constraints, by IDENTSTRESS$_{LEX}$. This idea is reflected in the rankings in (5.h). It generates all and only the patterns of exceptions found in single forms.

Consider now the attested patterns of accentual alternations. Those found in verbs are seen in (8). They include an unmarked alternation pattern, (8.a), in which stress moves between a final VCC$_0$ and a penult followed by a light final, plus two marked types, in (8.b-c). (8.b) illustrates verbs with marked antepenult stress; (8.c) displays verbs with fixed penult stress. Finally, row (d) illustrates a conceivable but impossible pattern, to be compared with the attested ones: antepenult stress before a light final alternating with penult stress before a heavy final[5].

(8) Normal and exceptional stress alternations in present tense verbs

	Stress distribution	Ø ending	_V ending	-VC ending	
(a)	default: alternating penult/final stress	arʌ́t	arát-ʌ	arʌt-ʌ́m	'show'
(b)	marked: 1st root σ/final stress	mʌ́tur	mʌ́tur-ʌ	mʌtur-ʌ́m	'sweep'

[5] Pre-antepenult stress is made impossible, in our analysis, by undominated *EXTLAPSER.

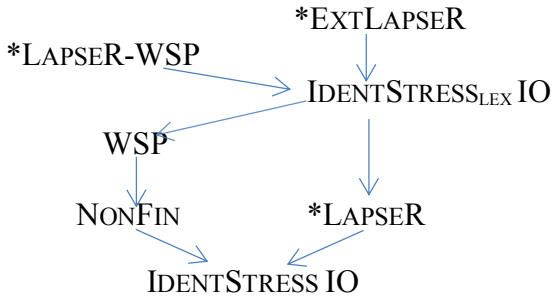
(c)	marked: fixed stress on 2nd root σ	adúk	adúʧ-e	adúʧ-em	'bring'
(d)	unattested: 1st root σ/2nd root σ	*bʌ́tur	*bʌ́tur-ʌ	*bʌtúr-ʌm	

The analysis proposed in (5) generates all the attested patterns of accentual alternations, (8.a-c), and only those. Patterns (8.a) and (8.b) were derived in (6) and (7); pattern (c) follows from the same rankings, if roots like *adúk, adúʧ-em* contain a lexically stressed second syllable. The fact that (8.d) is impossible has been previewed in (7), and follows from the assumption that any accentually irregular root has *some* underlying stress. If the lexical stress is on the second root syllable, we expect (8.c); if on the first, we expect (8.b). The shift in (8.d) can't be generated from one lexically stressed syllable, in first or second position; nor from multiple lexical stresses.

To understand how Harris's analysis might generate (8), recall that his proposal is to derive exceptional stress by marking a word-final segment extrametrical. In verbs, this word-final segment may belong to an ending, but the extrametricality condition must be entered in the lexical entry of roots, to distinguish type (a) roots from type (b) or (c). Setting this odd feature aside, a Harris-style analysis can't generate patterns (b) or (c), because neither of them displays a constant pattern of segment extrametricality: e.g. *mʌ́tur, mʌ́turʌ* in (8.b) can be generated by extrametricality of the last segment, but that also predicts **mʌtúrʌm*. Similarly,

adúk and *adúʧe* get the correct stress if nothing is extrametrical, but *adúʧem* requires that final /m/ be extrametrical. The only case that is easily derived by Harris's analysis is the unattested pattern (d): **bʌ́tu.<r>, *bʌ́tu.r<ʌ>, *bʌtúrʌ<m>*. We conclude from this that the better analysis of marked and default stress is based on the ranking schema M_1 >> IDENTSTRESS$_{LEX}$ IO >> M_2. That analysis is summarized below:

(9) Ranking summary



4 Perfect correspondence: accentual evidence

We return to the pattern of accentual similarity between perfect stems. The analysis in (9) helps identify a key difference between VPfs and PPfs. The verbal forms systematically violate active Markedness constraints: forms like [*kʌz-ú*] violate NONFINALITY; pluperfects like [*kʌz-ú*]-*se-rʌ,* [*kʌz-ú*]-*se-rʌ-m,* violate *LAPSER and *LAPSER-WSP. PPfs, by contrast, are regular in terms of (9): all

end in a heavy syllable and all are stressed on that heavy final. From this difference between the accentually unmarked participles and the accentually marked tensed perfects, we infer that the perfect similarity pattern is directional: stress in the participle is computed according to (9); it is then transferred to the VPfs, which deviate from (9) in order to preserve their accentual similarity to the PPf.

A simple mechanism generates this form of directional identity: the PPf is the Base (B) of the entire perfect paradigm. Its shape is generated in a first step, comparable to a cycle 1. VPf forms are its Derivatives (D): they use the surface form of the PPf, as the input into their evaluation. The constraint requiring accentual identity among correspondent perfect forms, IDENT STRESS (BD), ranks above all Markedness constraints in (9).

(10) a. Deriving the Base, PPf: *kʌz-ú-t* 'fallen'

UR: /kad-u-t/	WSP	NONFIN
kázut	*!	
☞ kʌzút		*

b. Deriving Ds: *kʌzú* 'fell3sg', *kʌzú-se-rʌ-m* 'fell-plupf-1pl'

B: [kʌz-ú-t] 'fallen'	IDENT STRESS (BD)	NONFIN
[kʌ́zu]	*!*	
☞ [kʌzú]		*

B: [kʌz-ú-t] ‘fallen’	IDENT STRESS (BD)	*LAPSER-WSP
[kʌzu]-se-rʌ́-m	*!	
☞ [kʌzú]-se-rʌ-m		*

The contrast between *mʌtur-ʌ́-m*, (7.b), with stress shifted to the final, and *kʌzú-se-rʌ-m,* (10.c), with stem stress, emerges from the common ranking schema characterizing one OT theory of cyclic phonology (Benua 1997): FAITH (BD) >> M >> FAITH (IO). In the present case, the schema is instantiated as: IDENT STRESS (BD) >>*LAPSER-WSP >> IDENT $\text{STRESS}_{\text{LEX}}$ (IO).

A deviation from accentual identity in the perfect was mentioned earlier, in the case of pluperfects in *sese*, e.g. PPf [*ár-s*] vs. [*ar-s*]-*és-e* ‘had burned’. Here stress is shifted off its expected position in the root. The trigger of this shift is an identity avoidance constraint: a change of stress is used to differentiate two otherwise identical syllables, to reduce their similarity. I lack evidence for the full scope of this constraint in Romanian. The version in (11.a) prohibits only identical *CeCe* strings. More plausibly, any sequence of strictly identical CVs is avoided[6].

[6] Zukoff 2015 motivates an identical constraint in Ponapean. For Romanian, searches in dexoline.ro for identical CeCe strings show that they are absent word internally, outside the onomatopoeic and child-directed lexica. Items like *zíle-le* ‘the days’ are clitic groups.

(11) a. OCP(*CeCe*): one * for each identical *CeCe* string.

b. OCP (*CeCe*) >> IDENT STRESS (BD)

c. Deriving pluperfect *ars-é-se* ‘had burnt-3sg’

B: [árs] ‘fallen’	OCP (*CeCe*)	IDENT STRESS (BD)	*LAPSER
[árs] -e-se	*!		*
☞ [ars] -é-se		* (á-a)	

For *sese*-pluperfects other than the 3rd sg., this analysis is still insufficient. The stress shift to the first *se* is predicted by the present analysis in the 3rd sg. because the second *se* is a final light. But other *sese*-forms have an incentive to shift stress differently, because their Markedness score improves if stress lands on the second *se*. (12) shows how the attested 1st pl. form [*ars*]-*é-se-rʌ-m* loses to penult-stressed *[*ars*]-*e-sé-rʌ-m*. The 3rd pl. [*ars*]-*é-se-rʌ* will similarly lose to *[*ars*]-*e-sé-rʌ,* due to a critical *LAPSER violation.

(12) An initial failed attempt to derive *ars-é-se-rʌ-m* ‘burnt-plupf-1pl’

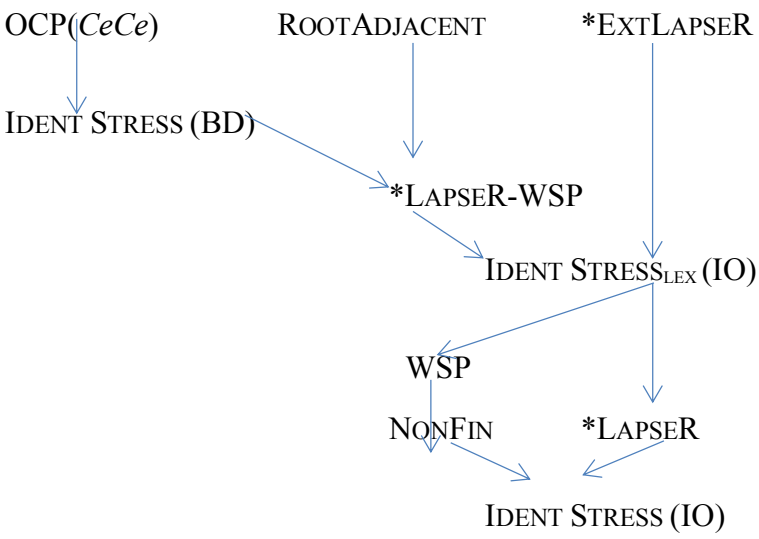
B: [árs] ‘fallen’	OCP (*CeCe*)	ID STRESS BD	*LAPSER-WSP
[árs]-e-se-rʌ-m	*!		*
[ars]-e-se-rʌ́-m	*!	* (á-a)	
[ars]-é-se-rʌ-m		* (á-a)	*!
!☚ [ars]-e-sé-rʌ-m		* (á-a)	

The simplest remedy is to stipulate that stress must remain in proximity of the root, not separated from it by any syllable. A constraint implementing this idea is ROOT ADJACENT, which bans candidates whose stress is separated by a syllable or more from the root. ROOT ADJACENT eliminates in (13) the winner of (12), allowing the attested form to emerge as optimal. A summary ranking follows.

(13) Deriving pluperfect *ars-é-se-rʌ-m* 'had burnt-1pl'

B: [ars]	OCP(*CeCe*)	ROOTADJ	ID STRESS (BD)	*LAPSER-WSP
☞ [ars]-é-se-rʌ-m			*	*
[ars]-e-sé-rʌ-m		*!	*	

(14) Perfect correspondence



5 Alternatives

The present analysis explains an asymmetry noted earlier: the surface structure of the PPf transparently explains its stress - e.g. *kʌz-ú-t* has predictable stress on a heavy final - but the identical stress of corresponding VPfs lacks surface phonotactic justification. Thus [*kʌz-ú*] 'fell-3sg' surfaces instead of the expected *[*kʌ́z-u*], as does [*kʌz-ú*]-*rʌm,* instead of expected *[*kʌz-u*]-*rʌ́m,* like *mʌtur-ʌ́m.* The preceding section has laid out the beginnings of an account that explains this asymmetry: the participle is the base of the VPfs, the tensed perfect forms, so the latter get their stress from the former, and they pay a Markedness price – e.g. violations of *LAPSER-WSP, for their identity to their Base. I explore next two alternative accounts of this identity, one based on McCarthy's (2005) Optimal Paradigms (OP) theory, and one based on the idea of lexically stressed morphs.

We begin with the latter. It is conceivable that unexpected stresses like [*kʌz-ú*] 'fell-3sg', [*auz-í*] 'heard-3sg' are due to the presence of lexically stressed theme vowels, *ú* and *í*. If so, the stress in [*kʌz-ú-t*] 'fallen', [*auz-í-t*] 'heard' is doubly motivated, both by lexical stress and by the weight of the final, while stress in [*kʌz-ú*], [*auz-í*] is due entirely to the theme vowel's underlying accent, and not to B-D correspondence. We should examine then the evidence for lexical stress on theme vowels. The theme vowel -*ú*- of [*kʌz-ú*] is limited to the perfect, where no

independent consideration suggests that it has, or lacks, inherent stress. More revealing are the weak verbs of the type *auz-í*, *lʌud-á*, *arʌt-á,* seen earlier in (3) and (8). These have the same theme vowels in the perfect and the present. Their perfect paradigms were seen in (3). Representative present forms follow.

(15) Present of weak verbs

infinitive	*auz-í* ROOT-THEMEV 'hear'		*arʌt-á* ROOT-THEMEV 'show'	
present	1	[*aúd*] [*auz-í*]-*m*	[*arʌ́t*]	[*arʌt-ʌ́*]-*m*
	2	[*aúz*]-*j* [*auz-í*]-*ʦj*	[*arʌ́ʦ*]-*j*	[*arʌt-á*]-*ʦj*
	3	[*aúd*]-*e* [*aúd*]	[*arát*]-ʌ	[*arát*]-ʌ

This data shows that, aside from the infinitive and the perfect, the theme vowels are stressed only if they happen to occur in predictably stressed positions: e.g. present [*auz-í*]-*ʦj* or [*arʌt-á*]-*ʦj*, both with stress on heavy finals. The same theme vowels can surface in stressless form: the [ʌ] in the 3rd persons of the present [*arát*]-ʌ, [*láud*-ʌ] is identifiable as the theme vowel -*a*- of the 1st conjugation. In such forms, the theme vowel is regularly unstressed by the system in (14), and is regularly reduced. The final -*e* of 3rd sg. *aúd-e* could similarly be a stressless, reduced avatar of the theme vowel -*i*-. The theme vowels delete in the present of the weak conjugations, if followed by vowel-initial endings, some of which then proceed to disappear in turn: e.g. /aud-i-u/, /arat-a-u/, /arat-a-i/ are the underlying

forms of 1st sg. *aúd* and *arʌ́t* and of 2nd sg. *arʌ́tsj* [7]. Significantly, stressed vowels don't otherwise delete in Romanian[8]. All this suggests that the theme vowels of weak verbs are not invariably stressed. Lexical stress is then not a credible basis for an account of the accentual identities in the perfect, quite aside from the fact that it would account for only a fragment of the perfect identities we analyze.[9]

The remaining alternative to our analysis uses the theory of Optimal Paradigms

---

[7] An account along these lines is defended by Feldstein (1994). The *-u* ending of the 1st sing. is justified in Steriade 1985. Significantly, Feldstein finds evidence for some underlyingly stressed morphemes (the imperfect -á-) but not for underlyingly stressed theme vowels. A question that remains open is how to reconstruct the opacity inherent in Feldstein's analysis - V_1-deletion in hiatus followed by deletion of the trigger V_2 - in the present account.

[8] See Steriade 1985 on stress-conditioned vowel deletion and gliding in Romanian.

[9] The infinitive is also stressed on a final light, in most verb types: e.g. *arʌt-á, auz-í*. A possible reason for these final stresses is that these infinitives are truncated from regularly penult-stressed forms like *arʌt-á-re, auz-í-re,* old infinitives which now function as verbal nouns. A truncation account of stress in the current infinitives is defensible synchronically and requires no mention of lexically stressed vowels. An alternative explanation is that final stress in infinitives is a means to avoid homophony to other paradigm cells. For verbs like *arʌt-á* and *auz-í*, the regular penultimate stress will generate, when we plug in the reduction processes affecting atonic syllables, *arát-ʌ* and *aúz-j*, but these forms are already in use as, respectively, 3rd sing and 2nd sing. indicative presents. Paradigm-internal anti-homophony has significant other effects in the Romanian conjugation. Either way, the upshot is, again, that no evidence supports lexically stressed theme vowels.

(OP; McCarthy 2005; cf. related proposals in Kenstowicz 1998), according to which a set of non-directional correspondence constraints promote similarity between the stems of forms that comprise a lexeme's inflectional paradigm. The OP constraint set includes MAX (OP), DEP (OP), IDENT F (OP), for any feature, including stress. Their function is to verify that each pair of stems in an inflectional paradigm is identical for the property named in the constraint. Entire candidate paradigms are evaluated simultaneously. For each constraint C, violations of C incurred by individual members are summed over each candidate paradigm. This includes violations of the OP CORR constraints. Each such violation represents a pair of paradigm members whose stems differ in the relevant respect.

A successful OP alternative to the current analysis could invalidate the key claim of this study, that asymmetrical Base-Derivative correspondence obtains even when the Base is not a constituent in its Derivatives. In OP analyses, there are no Bases and no Derivatives. There are only members of the same paradigm seeking to converge upon the optimal compromise between stem identity across the paradigm and Markedness/IO Faithfulness satisfaction. If such an analysis is right, our claim about B-D correspondence without containment cannot be sustained.

In (16), I present a successful OP analysis of the perfect paradigm of 'fall', seen earlier in (2). Three candidate paradigms are worthy of consideration. The paradigm in (16.a) displays regular stress on heavy finals, and otherwise on

penults, but suffers disqualifying violations of OP IDENT STRESS, a constraint defined in (17). The paradigm in (16.b) has uniform initial stress, but too many of its forms suffer from lapse. Finally, (16.c) is accentually uniform, like (16.b), and reduces Markedness violations to a minimum. It wins.

(16) OP analysis of stress in a perfect paradigms: *kʌzút* 'fallen'

	/kʌzu-/	OP IDSTRESS	*LAPSER	NF
a.	kʌzút; kʌzúj, kʌzúʃj, kʌ́zu kʌzurə́m, kʌzurə́tsj, kʌzúrə; kʌzusém, kʌzuséʃj kʌzúse, kʌzuserə́m, kʌzuserə́tsj, kʌzusérə	47*!		9*
b.	kʌ́zut, kʌ́zuj, kʌ́zuj kʌ́zuʃj, kʌ́zurəm, kʌ́zurətsj, kʌ́zurə; kʌ́zusem, kʌ́zuseʃj kʌ́zuse, kʌ́zuserəm, kʌ́zuserətsj, kʌ́zuserə		8*!	
☞c.	kʌzút; kʌzúj, kʌzúʃj kʌzú, kʌzúrəm, kʌzúrətsj, kʌzúrə; kʌzúsem, kʌzúseʃj kʌzúse, kʌzúserəm, kʌzúserətsj, kʌzúserə		3*	4*

(17) OP IDENT STRESS: In each pair of paradigm-internal correspondent forms, W_1-W_2, assign a * for each V in W_1 that has a correspondent V' in W_2 such that V and V' differ in stress.

The test of this type of analysis comes when some constraint promoting a failure of identity among paradigm members dominates OP CORR. In the present case, one dominant constraint is OCP(*CeCe*), which blocks strictly identical *sese* sequences. To satisfy OCP (*CeCe*) and ROOTADJACENT, the first *se* must be stressed. (18) reveals that, when the dissimilarity-inducing constraint OCP (*CeCe*) outranks OP IDENT, the optimal candidate is one that minimizes the numbers of pairwise dissimilarities by shifting as many stresses as possible to the post-root syllable. The resulting forms are all accentually identical, with just one item, monosyllabic PPf *árs*, inevitably left with root stress. This winning candidate is very different from the actual paradigm of such perfects, represented by (18.c): real VPfs deviate from identity to the PPf only when forced by OCP(*CeCe*). This suggests that the number of violations of OP IDENT is in fact irrelevant. Of the other candidates in (18), (a) represents the paradigm with regular stress, on heavy finals and otherwise on penults; (b) represents the candidate that fully satisfies OP IDENT STRESS, by keeping stress on the root syllable. Each of them is eliminated by the top two constraints of the analysis. The significant part, though, is that the actual winner, (18.c) is also eliminated, by OP IDENT.

(18) Failed OP analysis of stress in a perfect paradigms: *árs* 'burnt'

	/ars-/	OCP (*CeCe*)	OP IDENT STRESS	*LAPSER	NF
a.	árs; ars-éi, arsé-ʃj, árse arse-rʌ́m, arse-rʌ́tsj, arsé-rʌ; arse-sém, arse-séʃj arsé-se, arse-serʌ́m, arse-serʌ́tsj, arse-sérʌ	6*!	39*		9*
b.	árs; árs-ej, árse-ʃj, árse árse-rʌm, árse-rʌtsj, árse-rʌ; árse-sem, árse-seʃj, árse-se, árse-serʌm, árse-serʌtsj, árse-serʌ	6*!		12*	*
c.	árs; árs-ej, árse-ʃj, árse árse-rʌm, árse-rʌtsj, árse-rʌ; arsé-sem, arsé-seʃj arsé-se, arsé-serʌm, arsé-serʌtsj, arsé-serʌ		42*!	6*	*
!☚ d .	árs; ars-éj, arsé-ʃj, arsé arsé-rʌm, arsé-rʌtsj, arsé-rʌ; arsé-sem, arsé-seʃj arsé-se, arsé-serʌm, arsé-serʌtsj, arsé-serʌ		12*	3*	3*

I noted in fn. 3 the existence of accentual variation in the stress of perfects like *árse*: variants like *ars-éj, arsé-ʃj* exist for the 1st and 2nd sg, alongside *árs-ej, árs-*

efj. They can be generated using additional constraints that outrank IDENT STRESS (BD). Importantly, such variants don't support an OP analysis either. Whatever constraint determines their stress, its interaction with OP IDENT generates the wrong forms for the rest of the paradigm. The right generalization is that the stress of each VPf form remains identical to that of its PPf, independently of what the rest of the paradigm does. Constraints like OCP (*CeCe*), which shift stress away from the Base have a local effect on individual forms, not a global effect on the paradigm. This pairwise relation between the participle and individual VPf forms can only be derived in a directional B-D analysis.

This discussion explains why the second alternative to our B-D analysis is not viable. What it does not invalidate is an OP analysis of fragments of the perfect paradigm. If we consider just the tensed perfect, or just the pluperfect, in isolation from other perfect forms, the accentual uniformity within these microparadigms is complete, and thus analyzable in OP. But what is striking about the Romanian perfect is the nearly complete accentual - and, as we shall see, segmental - identity of perfect forms *across* the PPf-VPf divide and across the perfect-pluperfect boundary. OP doesn't help generate that.

6 Containment and the Romanian perfect

Thus far I have shown that directional B-D Correspondence is the right mechanism to generate the pattern of identity described here. The next step in the overall argument is to show that the Base of the perfect paradigm, the PPf, is not contained in its VPf Derivatives.

In a surface phonological sense, this is a directly observable fact. In the paradigm of PPf *kʌzút*, the VPfs are *kʌzú*, *kʌzúrʌm, kʌzúserʌm,* etc*.,* not **kʌzút*, **kʌzútrʌm, *kʌzútserʌm*. Let's restate the significance of this fact for the analysis. The perfect marker *-t* of *kʌzút* causes its *u* to be stressed, by making heavy the final syllable. Then this *-t* is indirectly responsible, via IDENT STRESS (BD), for the stresses of VPfs like *kʌzú*, *kʌzúrʌm, kʌzúserʌm*, etc: without a requirement of identity to regularly stressed *kʌzút*, the stresses of these VPf forms would be elsewhere. Since *-t* itself is missing in VPfs, this shows that *kʌzút,* the full form whose stress is cyclically transmitted to the VPf forms, is not phonologically contained in them.

But we had set out to establish a distinct point, which is more directly relevant to the broader conclusion of this study: the Base is not *syntactically* embedded in its Derivatives in these perfect paradigms. This section completes that argument. I show that *-t* can't occur in VPfs like *kʌzú*, *kʌzúrʌm* because a syntactic feature expressed by *-t* isn't compatible with the syntactic structure of VPf forms. If any

syntactic feature of the PPf isn't contained in a VPf, it follows that the former isn't embedded in the latter. Once we establish this, a non-containment-based account of two central facts will have to be provided: PPfs and VPfs are in correspondence; and, in this relation, the PPfs are the Bases.

6.1 Classes of strong perfects and perfect exponents

There are six varieties of PPfs, three from weak verbs and three from strong verbs. In such forms, the suffixes *-t, -s* and *-u* (in 19.a-b) are the perfect markers whose distribution we investigate, while *-a-* and *-i/ɨ-* (in 19.c) are aspectually neutral theme vowels used in the present and perfect of weak verbs. The numbers in (19.a) are counts of strong verbs from the lists in Lombard and Gâdei 1981.

(19) a. Classes of strong perfects and lexical counts

	Class 1 (N: 32)	Class 2 (N: 186)	Class 3 (N: 21)
Infin.	*kʌd-eá* <*kʌd-é* 'fall'	*árd-e* 'burn'	*fiérb-e* 'boil'
PPf	*kʌz-ú-t*	*ár-s*	*fiér-t*
VPf	[*kʌz-ú*]-*j*, [*kʌz-ú*]-*rʌ-m, ...*	[*ár-s*]-*e-j*, [*ár-s*]-*e-rʌ-m, ...*	[*fiér-s*]-*e-j*, [*fiér-s*]-*e-rʌ-m, ...*

b. Distribution of perfect markers in the strong verbs

	PPf			VPf		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
-t	√ (*kʌz-ú-t*)		√ (*fiér-t*)			
-u	√ (*kʌz-ú-t*)			√ (*kʌz-ú*)		
-s		√ (*ár-s*)			√ (*ár-s-e*)	√ (*fiér-s-e*)

c. Classes of weak perfects and examples

	-a verbs	*-i* verbs	*-ɨ* verbs
Infin.	*lʌud-á* 'praise'	*auz-í* 'hear'	*hotʌr-ɨ́* 'decide'
PPf	*lʌud-á-t*	*auz-í-t*	*hotʌr-ɨ́-t*
VPf	[*lʌud-á*]-*j*, [*lʌud-á*]-*rʌ-m, ...*	[*auz-í*]-*j*, [*auz-í*]-*rʌ-m, ...*	[*hotʌr-ɨ́*]-*j*, [*hotʌr-ɨ́*]-*rʌ-m, ...*

The suffix distribution in (19) suggests several elements of analysis. First, the suffix *-t* never occurs in VPf forms, whether weak or strong, only in PPfs. We explain this fact if *-t* is an exponent of both [+perfect] and of a lexical category feature, say [+adjective], which is incompatible with *-t*'s occurrence in VPfs.

Second, the suffix *-s* occurs in both VPfs and in PPfs, as seen in the strong classes 2 and 3, (19.a-b). We explain this if *-s* is an exponent of [+perfect] only, and thus useable in both adjectival and verbal forms. Similarly, the suffix *-u* occurs in both VPf forms and PPf forms, as seen in the strong class 1. We infer that it too is an exponent of [+perfect] only, like *-s*.

Third, the suffixes -*u* and -*s* don't co-occur: there are no **kʌz-ú-s* type perfects. We have explained this by attributing to -*u* and -*s* identical exponence functions. There is no need for both in one form.

Fourth, the suffixes -*u* and -*t* do co-occur in PPfs, as in *kʌz-ú-t.* Here, -*t* would be a sufficient exponent, but -*u's* function is to satisfy segmental phonotactics and to keep stress off the root, a point developed in §5.3.

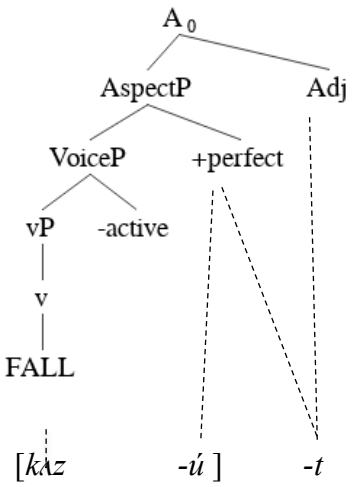
Next, the suffixes -*s* and -*t* don't co-occur either. This is partly explained by the syntactic features our analysis attributes to them. In VPfs, -*t* is unusable, *qua* adjectival suffix. For PPfs, -*s* expresses a proper subset of the features expressed by -*t*. By the Subset Principle (Halle 1997), we expect then only -*t* to occur in PPfs, and certainly not a combination of -*t* and -*s*. What remains unexplained is what makes -*s* a possible PPf marker at all, for some strong verbs: one might expect -*t* to always replace it, as in Class 3 perfects. The analysis proposed here offers no synchronic answer to this question. All sigmatic PPfs are inherited archaisms, rather than innovations. Synchronically, they must be analyzed by means of lexically indexed exponence constraints.

To sum up, this section has proposed that -*t* is an exponent of the features [+perfect] and [+adjective], while -*s* and -*u* express only [+perfect]; verb roots must be lexically indexed in order to use the [+perfect] exponent -*s*. This feature assignment explains several distributional generalizations, including the fact of

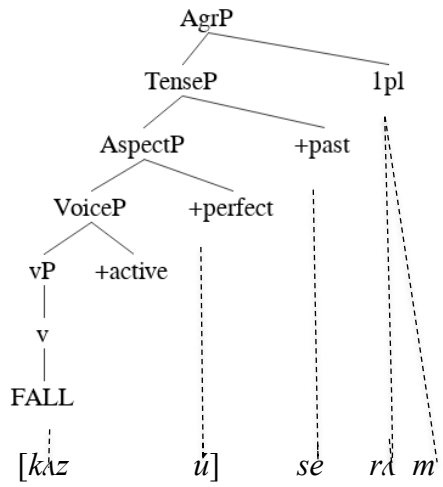
central interest here: the *-t* of PPfs like *kʌzút* can't be present in corresponding VPfs like *kʌzú*, *kʌzúserʌm*, while the *-s* of participles like *árs*, must be present in VPfs like *árse, árserʌm.* (20) helps visualize how syntactic structures map to exponents under this proposal. Dotted lines connect each morph to the nodes it is an exponent of. Perfect stems are in brackets.

(20) Exponents of two perfect forms: a PPf and a VPf (1st pl pluperfect)

(a)



(b)



Our next step is to clarify what causes sets of perfect forms, such as {*kʌzút; kʌzú*, *kʌzúserʌm…*}, to enter into correspondence at all. The answer will be the hypothesis that such instances of paradigm uniformity are due to constraints that place in correspondence cognate sets of forms sharing a syntactic feature, or a

bundle of syntactic features, independently of syntactic embedding. In the present case, the feature requiring correspondence is [+perfect].

6.2 Perfect Correspondence

The constraint causing stem identity among Romanian perfect forms is (21)

(21) CORR(PERFECT): If two syntactic structures S_1 and S_2 are lexically related and contain the value [+perfect], the stems of the word-sized exponents of S_1 and S_2 stand in correspondence.

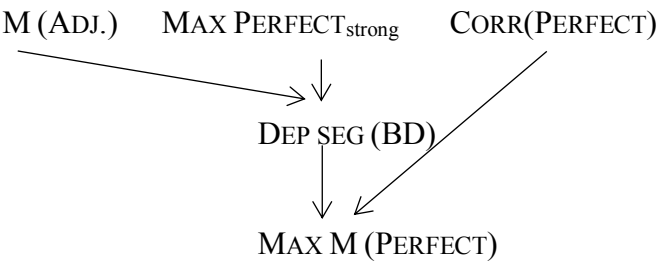
When sets of perfect forms enter in correspondence in order to satisfy (21), constraints like IDENT are activated. Under certain rankings, they will generate paradigms with phonologically identical or similar stems. The version of CORR(PERFECT) in (21) requires correspondence between *stems*. We define *stem* in this context as the smallest contiguous string of morphs that includes exponents of all the syntactic nodes shared by S_1 and S_2. How a perfect stem maps to a set of syntactic nodes was illustrated in (20), and reveals that the syntactic nodes corresponding to stems placed in correspondence by (21) need not be identical: the voice features of PPfs and VPfs differ.

CORR(PERFECT) causes not only accentual but also segmental identities between perfect stems. We revisit the data in (19) to illustrate one of these. In Class 3 strong perfects – e.g. *fiert, fierse, fierserʌ* – the adjectival *-t* must be replaced in VPfs by *-s*, because bare VPfs like **fiere, *fiererʌ,* without any perfect suffix, are unacceptable. The table in (19.b) reveals a related generalization: *every* strong VPf contains some overt perfect suffix. This is not the case in weak verbs, as (19.c) and (3) show: weak PPfs like *auzít* 'heard' correspond to VPf sets that lack any perfect marker, like *auzí* 'heard-3sg'. Alternative VPfs like **auzú,* using the perfect suffix *-u*, or simple perfects like **auzís(e),* using *-s* – the latter comparable to attested strong paradigms like {*skris, skríse, skríserʌ*…} 'written, wrote' – are impossible in the weak verbs. The analysis must characterize two complementary generalizations: it must exclude all strong perfects lacking any [+perfect] affix, while also insuring that weak VPfs lack [+perfect] morphemes. The first step is to introduce a constraint, based on Wolf's (2008) MAX MORPH schema, banning perfect structures that lack exponents of the perfect.

(22) MAX M (PERFECT): a * for every instance of [+perfect] that lacks an overt exponent.

An undominated version of (22) will be indexed to the class of strong verbs in (19.a). This accounts for the first generalization: all strong perfects contain an exponent of the perfect. To explain the absence of perfect markers in weak VPfs, i.e. the impossibility of weak paradigms like *{*auzít, auzú, auzúr*ʌ…} or *{*auzít, auzíse, auzíser*ʌ…} (where **auzíse* etc. is a simple perfect) we appeal to CORR(PERFECT). In weak paradigms, the perfect stem of PPfs differs from that of VPfs minimally, only as required by the need to exclude *-t* from tensed forms. In impossible paradigms like *{*auzít, auzú, auzúr*ʌ} or *{*auzít, auzíse, auzíser*ʌ} (the latter two as simple perfects) the stems of participial and verbal forms differ more from each other: *-t* is not just missing from all VPfs, but is also replaced by *-u* or *-s*. To analyze the weak paradigms then, we rank CORR(PERFECT) and DEP SEG (BD) above the general version of MAX M (PERFECT). The constraint DEP M (ADJ.), also based on Wolf's (2008) proposals, completes the analysis: it bans VPfs containing exponents of [+adjective] like *-t*.

(23)



Below I derive individual VPf members, from a weak and a strong perfect paradigm, to illustrate how (23) characterizes some of the generalizations about perfect exponence presented in this section. As before, I assume that the PPf is the Base of each paradigm. Perfect stems are in brackets. The effect of MAX PERFECT$_{strong}$ is shown in (24.a), a 3rd pl. VPf from a strong verb. The contribution of CORR(PERFECT) is seen in (24.b), a 3rd pl. VPf from a weak verb: CORR(PERFECT) activates DEP SEG (BD), which excludes two losing candidates. Without CORR(PERFECT), DEP SEG (BD) would not block satisfaction of MAX PERFECT, and forms like *[*auzí-s*]-(*e*)-*rʌ*, second candidate in (24.b), would win.

(24) a. Deriving one VPf member of a strong perfect paradigm, class 3.

B: [fiér-t] 'boiled'	DEP M (ADJ.)	MAX PERFECT$_{strong}$	DEP SEG (BD)
[fier-t]-e-rʌ	*!		
[fierb]-e-rʌ		*!	*
[fier]-e-rʌ		*!	
☞ [fier-s]-e-rʌ			*

b. Deriving a VPf member of a weak perfect paradigm.

B: [auzí-t]^{i} 'heard'	DEP M (ADJ.)	CORR(PERF)	DEP SEG (BD)	MAX PERF
[auzí-t]^{i} -(e)-rʌ	*!			
[auzí-s]-(e)-rʌ		*!		
[auzí-s]^{i} -(e)-rʌ			*!	
[auz-ú] ^{i}-rʌ			*!	
☞ [auzí]^{i}-rʌ				*

6.3 Excursus: perfect-infinitive correspondence

We have yet to explain the double perfect exponence seen in PPfs like *kʌz-ú-t, dur-ú-t, pʌr-ú-t,* of the strong Class 1. Our analysis states that the perfect suffix *-u* expresses a proper subset of the syntactic features expressed by *-t*. We should ask then why *-u* is used at all in the *-u-t* participles, where *-t* makes it redundant: why *kʌzút, pʌrút, durút* and not **kázt, *dúrt, *párt,* just like *fiért*?

A further correspondence effect provides an answer. If the infinitive of a verb has iambic stress (as in *kʌd-eá, pʌr-eá, dur-eá*), the PPf has iambic stress as well (as in *kʌz-út, pʌr-út, dur-út*); and similarly for weak verbs, e.g. *auz-í, auz-í-t*[10]. The structure ROOT-V-t insures final, iambic stress in the PPf, in virtue of (14). Without the *-u* in PPfs like *kʌz-ú-t*, stress would be on the root, and would

[10] One exception is *adáus* 'added', with lexical stress in the PPf but regular final-stressed infinitive *adʌug-á.* (VPfs in Lombard & Gâdei 1981 for *adʌug-á* follow this old PPf: *adáuse* 'added-3rd sg'.) All other verbs with lexical stress in the weak conjugations behave as described in the text: e.g. *méstek* 'I chew', showing lexical stress on the first syllable, but *mestek-á.* The perfect forms *mestek-át*, *mestek-ʌ́* etc. follow the infinitive. What must be explained in such paradigms is the final stress in infinitives: given the lexically stressed root we expect **méstek-a* or, with reduction, **méstek-ʌ.* The hypothesis outlined in fn. 10, that infinitives are truncated derivatives of verbal nouns in *-re* and preserve their stress, explains these forms. *EXTLAPSER blocks **méstek-a-re,* forcing stress to advance to the penult in the *-re* noun. The infinitive preserves that stress.

mismatch the stressless root in the infinitive of this verb, *kʌd-eá*. Conversely, if the infinitive *has* root stress – because the verb root contains a lexical stress – the PPf generally has root stress too: *fiérb-em, fiér-t* (**fierbút*); *árd-em, ár-s* (**arz-út*); *pún-em, pú-s; dúʧ-em, dús*, etc. Had these PPfs included *-u,* in addition to *-t* or *-s,* stress would shift to the suffix, mismatching the infinitive (*púne, *punút*), or else WSP would be violated (*púne, *púnut*). There are exceptions, mostly in *k*-final roots (*fáʧe, fʌkút*, perhaps because *kt,* as in expected **fákt,* would regularly become *pt*), but the general correlation seems clear.

A preliminary suggestion for the mechanism behind this correlation is a preference for rhythmic identity between all non-tensed forms of a root, i.e. infinitives and participles. This will explain the double exponence in strong PPfs like *kʌz-ú-t.* Comparable chains of correspondence are found in Latin verb paradigms and their non-verbal derivatives (Steriade 2013, 2015). Romanian appears to have inherited such abstract patterns of rhythmic identity, if not their overt Latin manifestation.

7 Base Priority in the perfect and its source

In §2-3 I have presented an argument that the accentual identity among perfect stems in Romanian is directional: the stress of the PPf follows the general stress rules of the language and determines, via B-D correspondence, the stress of the

VPfs. This results in accentual anomalies in the stress of the latter. The preceding section has contributed to the same directional hypothesis. It has proposed that the choice to include the *-u* marker in strong PPfs like *kʌzút* – or not to include it, as in *árs* or *fiért* – stems from correspondence between the non-tensed forms of a root, the infinitive and the PPf. Once it is made, this choice of *-u* vs. no *-u* is transmitted to the VPfs, again via B-D correspondence. The direction is the same in both cases, from PPf to VPfs. The directional B-D analyses of stress identity offered in §3 have thus received independent justification.

Section §5 has shown that this directional identity can't be attributed to the Base being contained in the Derivative: PPfs are not embedded in cognate VPfs, in either a syntactic or a morphological sense. Hence, syntactic containment is not the source of the B-D asymmetry studied here.

What *is* then the source of this directional effect? We should distinguish grammar-external factors, which favor the selection of some forms as Bases in certain paradigms, from the grammatical reflex of the asymmetry between Bases and Derivatives. As far as the latter goes, the proposed mechanism can be (25), the statement that certain complex expressions have derivational priority over others.

(25) The exponents of PPfs are generated prior to those of VPfs.

(25) is comparable to the assignment of some derivatives to Level 1 and of others to a later Level. Statements like it are presupposed in Stratal OT analyses of level ordering and cyclicity (Kiparsky 2000), in analyses of directional paradigm uniformity effects elsewhere (Hall and Scott 2007) and in other instances of cyclic effects where Base Priority does not stem from syntactic containment (Steriade 1999, 2008, 2013, 2016).

It is clear that statements like (25) over-generate if nothing limits the pairs of expressions they can link, but it is not clear that this concern is best addressed by designing grammars in a different way. Albright (2010) has identified relative *informativity* as the property responsible for selecting one Base among a set of paradigm cells. The most informative potential Base B in a paradigm is one that would trigger the smallest number of phonological and morphological neutralizations, if B's stem were generalized to all cells. More informative Bases are favored, because they allow larger numbers of contrasts to surface across paradigms. Albright shows that this factor selects the right Base in a number of diachronic developments that create uniform paradigms. Relative informativity is a grammar-external factor, because there is no intrinsic connection between some paradigm cell being more informative than others and that cell being designated as the Base of a paradigm, in a statement like (25) or in any other format.

A different grammar-external factor that can differentiate potential Bases is token frequency. The more frequent the cell – or the higher the average token frequency of a cell type across paradigms –the better known its properties will be to learners, and thus the more likely it is that learners will extend its properties to other cells. Token frequency plays only an indirect role in Albright's model, and only this indirect role is validated by his findings. However, the frequency difference between Romanian PPfs and VPfs is so large at present that it may eliminate the VPfs from the competition for basehood. The simple perfect is in the process of being replaced in the literary language by the analytic perfect, which uses the PPf. The remaining synthetic VPf paradigm is the pluperfect, an uncommon tense-aspect combination. An illustration of the frequency disparity between VPfs and PPFs is provided by comparing Google hits for VPf and PPfs in 10 common roots. Results are in (26). The verbs were chosen so as to avoid homophony between the perfect forms and any paradigm members. This is why only one verb of the productive *-i* conjugation is included: the 3rd sg. perf. (e.g. *auzí*) is identical with the infinitive and is spelled identically with the 2nd sg. present. As (26) shows, VPf percentages range between a high of 15%, in one verb, and more common figures approaching 0.

(26) Percentages of VPf forms of 10 verbs from the total number of perfect forms: Google hits in millions, unless indicated otherwise

	încuia	*mânca*	*auzi*	*încrede*	*rămânea*	*desface*	*cumpăra*	*înțelege*	*vedea*	*cădea*
total perfect	778k	96	42	437k	110	2	24	25	28	6
% VPf	0.02	0	0.06	0	0.01	0.03	0.02	0	0.06	0.15

The frequency disparity is even more substantial than (26) suggests. The PPf is phonologically homogeneous, as all its gender-number inflected forms have accentually identical stems, without the benefit of leveling. By contrast, as seen in (16) and (18), stress in the VPfs will alternate in the absence of some form of paradigmatic leveling, landing on the root, or the aspect markers, or the endings, depending on what affixes follow the root. This enhances the frequency disparity between candidates for the Base stem: the unique stem of the PPf with its cumulative frequency competes against each one of the three or four accentually diverse stems to be expected in the VPfs.

If we set frequency aside, it is unclear that informativity favors the PPf as the base of the perfect paradigm. In most strong verbs, the verb root is directly followed by *-t* or *-s*, participial suffixes that cause extensive neutralizations through the loss of the last consonant, as in *fiert* < *fierb-t, ars* < *ard-s, kurs* < *kurg-s, pus* < *pun-s*. Establishing the relative informativity of the PPf compared to other potential bases

would require a difficult comparison to a hypothetical set of VPfs, whose stems would have to be computed independently of the participle and of each other. In the absence of a clear way to carry this out, it seems safe to assume that the token frequency disparity is a large enough factor in favor of the PPf to dwarf any possible VPf advantage in informativity.

To summarize, the proposal is to model the B-D asymmetry between participles and tensed perfect forms by the derivational statement in (25). According to (25), VPfs are generated in a second derivational step, post PPf. The constraint CORR (PERF) and the identity conditions it triggers are necessarily activated only at this second stage, and can thus affect only the shape of VPfs. The preliminary proposal is to let extra-grammatical factors, like comparative token frequencies of various candidate Bases, dictate the derivational order in (25) and comparable cases.

The most important point in this discussion is that an analysis like the one advocated here is equally available to standard cases of cyclic inheritance, where Bases *are* contained in their Derivatives. The difference between those patterns and the one studied here (or in Hall and Scott 2007; or in the diachronic developments reconstructed by Albright 2005, 2010) is that (a) derivational priority statements like (25) are unnecessary in standard cyclic cases, but required here; and (b) a preference can be detected in other instances for B-D correspondence to involve only Bases that are contained in their Derivatives. The

Romanian perfect case shows, along with much other evidence, that such a preference for B-D correspondence under containment is violable: any constraint penalizing non-nested correspondent pairs must rank below CORR(PERFECT) in Romanian. Once we recognize the violability of this containment condition, B-D correspondence can receive an identical analysis as paradigmatic uniformity. In the Romanian perfect case, we have shown that this unification is not just conceptually satisfying, but empirically necessary: §4 has shown that non-directional mechanisms fail to generate the uniform paradigms described here.

Going beyond the data analyzed in this paper, Romanian morphology is a rich source of phenomena involving cyclic inheritance without containment. Some were studied elsewhere (Steriade 2008, Stanton and Steriade 2018), while others await formal analysis. Taken together, they suggest that the constraint limiting BD correspondence to nested pairs of Bases and Derivatives ranks very low throughout the grammar of Romanian.

References

Albright, Adam. 2005. "The morphological basis of paradigm leveling." In *Paradigms in Phonological Theory*, ed. by Laura Downing, T. Alan Hall, and Renate Raffelsiefen, 170-210. Oxford: Oxford University Press.

Albright, Adam. 2010. “Base-driven leveling in Yiddish verb paradigms.” *Natural Language and Linguistic Theory* 28 (3): 475-537.

Benua, Laura. 1997. *Transderivational Identity: Phonological Relations Between Words*. Ph.D. thesis. UMass Amherst, MA

Feldstein, Ronald. 1994-1995. “Accentul verbului în limba română” *Dacoromania* 1 (1-2): 75-96.

Gordon, Matt. 2000. “A factorial typology of quantity insensitive stress.” *Natural Language and Linguistic Theory* 20 (3): 491-552.

Hall, Tracy A., and John H. Scott. 2007. “Inflectional paradigms have a base: evidence from s-Dissimilation in Southern German dialects.” *Morphology* 17: 151-178.

Halle, Morris. 1997. “Distributed morphology: Impoverishment and fission.” In *MITWPL 30: Papers at the Interface,* ed. by Benjamin Bruening, Yoonjung Kang, and Martha McGinnis, 425-449. Cambridge, Mass.: MIT Press.

Harley, Heidi, and Elizabeth Ritter. 2002. “Person and number in pronouns: a feature-geometric analysis.” *Language* 78 (3): 482-526.

Harris, J. W. 1983. *Syllable structure and stress in Spanish: a non-linear analysis*. Cambridge, Mass.: MIT Press.

Kenstowicz, Michael. 1998. “Base identity and uniform exponence.” In *Current Trends in Phonology: Models and Methods*, ed. by Jacques Durand, and Bernard Laks, 363-394. Salford: University of Salford.

Alf Lombard, and Constantin Gâdei.1981. *Dictionnaire morphologique de la langue roumaine*. Editura Academiei Republicii Socialiste România and C.W.K Gleerup, Lund.

McCarthy, John. 2005. “Optimal paradigms.” In *Paradigms in Phonological Theory*, ed. by Laura Downing, T. Alan Hall, and Renate Raffelsiefen, 170-210. Oxford: Oxford University Press.

Pană Dindelegan, Gabriela. ed. 2013. *The Grammar of Romanian*. Oxford: Oxford University Press.

Pater, Joe. 2000. “Nonuniformity in English Stress: The role of ranked and lexically specific constraints.” *Phonology* 17 (2): 237-274.

Stanton, Juliet, and Donca Steriade. 2018. *Stress and the cycle, in English and elsewhere*. In progress.

Steriade, Donca. 1985. “Glides and vowels in Romanian.” In *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*. ed. by Claudia Brugmann, and Monica Macauley, 47-64.

Steriade, Donca. 1999. 1999. “Lexical conservatism in French adjectival liaison.” In *Procedings of the 25th meeting of the Linguistic Symposium on Romance*

Languages, ed. by Barbara Bullock, Marc Authier, and Lisa Reed, 243-271. Amsterdam: Benjamins.

Steriade, Donca. 2008. "A pseudo-cyclic effect in Romanian Morphophonology." In *Inflectional Identity*, ed. by Asaf BAchrach and Andrew Nevins, 314-359. Oxford: Oxord University Press.

Steriade, Donca. 2013. "The cycle without containment: Latin Perfects." to appear.

Steriade, Donca. 2016. "The morpheme vs. similarity-based syncretism: Latin t-stem derivatives." In *The Morphome Debate*, ed. by Ana Luís, and Ricardo Bermúdez-Otero, 112-174. Oxford: Oxford University Press.

Vasiliu, Emanuel. 1965. *Fonologia limbii române*, Bucureşti: Editura Ştiinţifică.

Wolf, Matthew. 2008. *Optimal Interleaving: Serial Phonology-Morphology Interaction in a Constraint-Based Model.* Ph.D. dissertation, University of Massachusetts Amherst.

Zafiu, Rodica. 2013. "Mood, Tense and Aspect." in Pană Dindelegan (ed.) 2013, 24-64.

Zukoff, Samuel. 2015. "Stress Restricts Reduplication." *Proceedings of the Annual Meetings on Phonology 2014.* Washington, DC: Linguistic Society of America.