Polish as a CV language
Complexity Scales and Licensing (CSL)
– a version of Government Phonology (GP)

Aim: provide a theory of speech sound self-organization in which phonological phenomena such as, *phonotactics*, *processes*, *typology*, *markedness*, *learnability* etc. follow from the general design of the internal structure of segments and their arrangement in the domain of phonological representation.

(1) Interactions (relations) in CVCV Phonology:

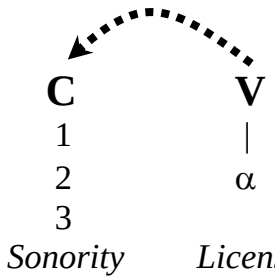
licensing { C ← V

government { C ↔ C
V ↔ V

effects depend on:
whether there is or isn't a governing relation (↔)
what Cs & Vs are made of (properties)
licensing strength (←)

(2) Licensing

Government & ***Licensing*** are the only two organizing forces in phonological representation.



- *Licensing*: interaction between consonants and vowels (C←V). Nuclei sanction the existence of the preceding onsets (what they are), and their relations (what they do).

- 'Normally', every onset receives licensing from its nucleus. Onset poses licensing demand on its nucleus. Absence, or weak licensing results in melodic depletion.

(3) Government

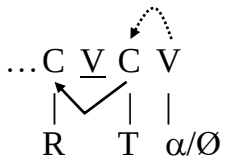
Government: interaction between consonants (C↔C) or vowels (V↔V). In consonants, government is a relation of dependency / control: less sonorous consonants (T) in a given configuration govern the more sonorous ones (R).

(4) Consonant clusters – *true* and *false (bogus)*

- ***True*** clusters involve 'government'. They are licensed as a whole by the following nucleus (they are parallel to compounds in morphology). They show phonotactic restrictions.

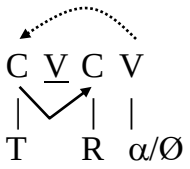
- ***False*** clusters do not involve 'government'. The intervening empty nucleus licenses its onset. The phonotactics in such clusters is more relaxed.

a. *true RT*



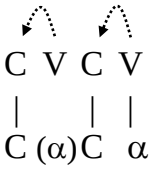
ba**r** **k**
żó**ł** **w**
da**r** **ń**
ba**ś** **ń**

b. *true TR*



t **r** a wa
k **r** o wa
bó **b** **r**
wia **t** **r**

c. *false*



k **r** a ~ kier
k **t** o
pu **d** **r** u ~ puder
swe **t** **r** a ~ sweter

(5) Conditions on Government between consonants

a.	***melodic sonority profiles*** (in which the governor, symbolized as (T), is sufficiently more sonorous than the governee (R). The labels T/R are not absolute, but relative, e.g.: /f/ is a governor (T) in *free*, and a governee (R) in *hefty*).
b.	***adjacency*** (the two consonants must be adjacent in the relevant sense: CC).
c.	***licensing*** (governing relations, like simplex segments, require licensing from the nucleus following such a segment or relation).

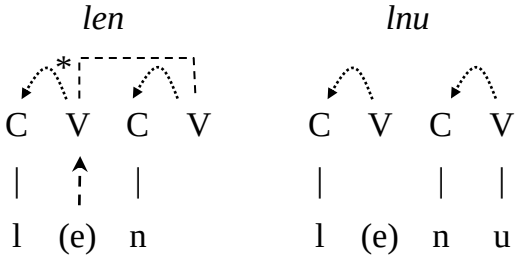
- Like with gravitation, government must be contracted if all conditions are fulfilled.

- Clusters in which government is involved are very much like compounds in morphology, they form an extended (single) domain, here for the purposes of licensing – thus clusters are more difficult to license (heavier).

(6) Conditions on *false* clusters

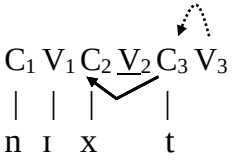
a. 'Ø's do not occur in sequences (*Ø–Ø)

b. 'Ø' is a licenser of the preceding structure

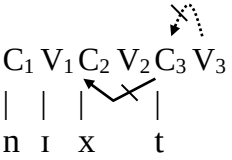


(7) Compensatory lengthening ([nixt] > [ni:t] > GVS > [nait])

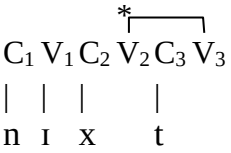
a. *original form*



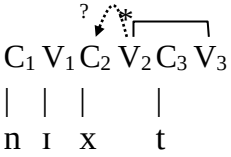
b. *weakened licensing*?



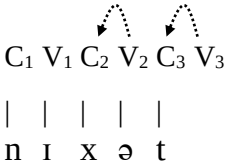
*c. *bogus cluster is out…*



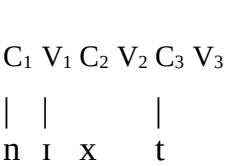
*d. *even if V_2 is a licenser*



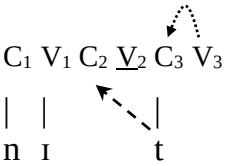
[?]e. *epenthesis as repair*



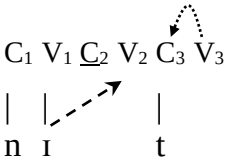
*f. *[x]-deletion*



[?]g. *gemination*

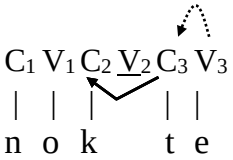


☺h. *compensatory lengthening*

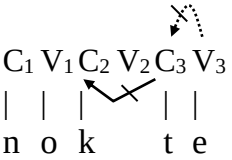


(8) Gemination in Italian ([nokte] > [notte])

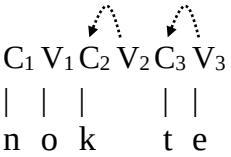
a. *original form*



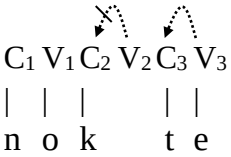
b. *weakened licensing*?



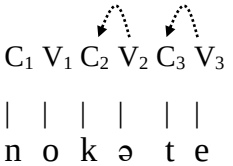
[?]c. *bogus cluster OK…*



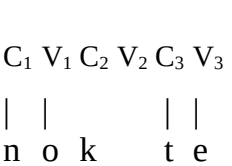
*d. *unless V_2 is not a licenser*



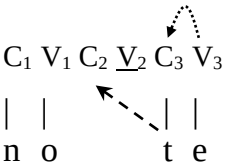
[?]e. *epenthesis as repair*



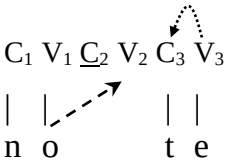
*f. *[k]-deletion*



☺g. *gemination*

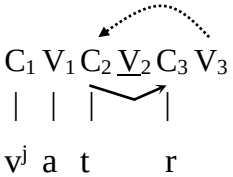


☺h. *compensatory lengthening*

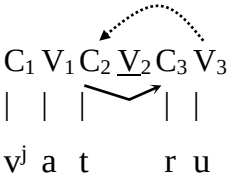


(9)Vowel-zero alternations in Polish, some structures…

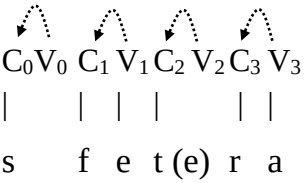
a. *wiatr*



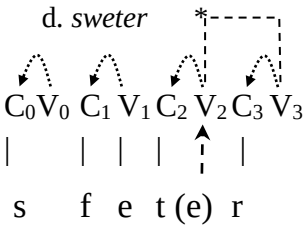
b. *wiatru*



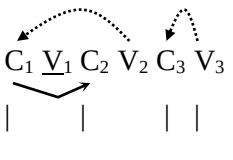
c. *swetra*



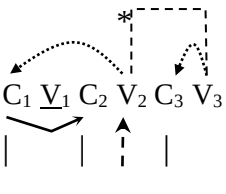
d. *sweter*



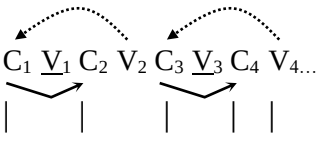
e. *krwi*



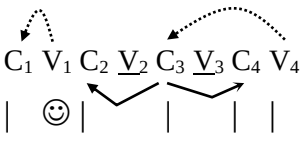
f. *krew*



g. *drgnąć*



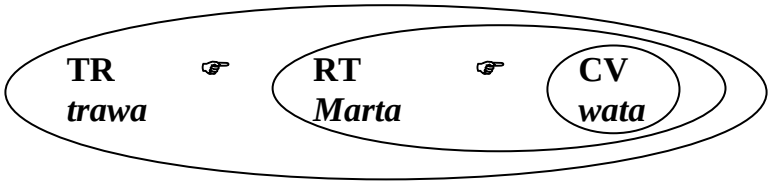
h. *pstry*



(10) Syllable structure typology

level	*structure*		*example (Eng)*	*example language, licensing strength settings*
I	$C\underline{\alpha}$	CV	ba**by**	English, Turkish, Desano
II	$R.T\underline{\alpha}$	CVC	wi**nte**r	English, Turkish
III	$TR\underline{\alpha}$	CCV	**tra**p	English

(11) Syllable markedness



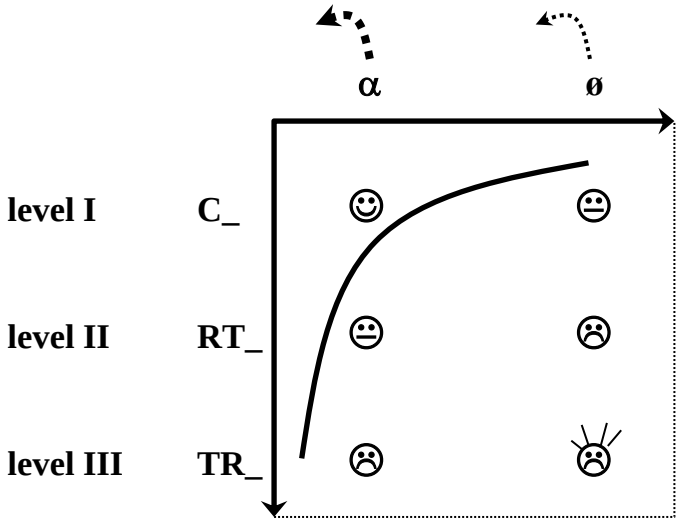
(12) Syllabic *typology* and *markedness* follow from *government* & *licensing*

I	II	III
a. Simple Licensing	b. Direct Government Licensing	c. Indirect Government Licensing
C V | | C α	C V C V | | | R T α	C V C V | | | T R α
c i t y	*mi l k y*	*t r ee*
CV open syllables	CV open syllables CVC coda-onset contacts	CV open syllables CVC coda-onset contacts CCV branching onsets

- The complexity scale corresponds to a scale of licensing strength – the given structural complexities are possible only if the nuclei are assigned the required licensing strength.
- ***Formal Complexity:*** Derives from the presence or absence of government and its type. Onset configurations exhibit three levels of formal complexity. Each formally defined level is like a 'quantal' region of stability, which poses different demand on the licenser – the nucleus that follows the respective configurations.
- Formal complexity scale corresponds to the required licensing strength.
- 'Ø' can license all levels of complexity
- structure licensed by 'Ø' implies the presence of the same structure licensed by a vowel 'α': ($\phi \supset \alpha$)

(13) The scale of licensers

(14) Syllabic space in Polish is defined by two vectors: scale of <u>licenser types</u> (differing in licensing strength) and <u>formal complexity</u> scale (scale of required licensing strength)



- Different structural configurations are relatively more or less marked due to their relative complexity and the relative strength of different types of licesers
- CV as in ***<u>city</u>*** is the the least marked because the simplest form is licensed by the strongest licenser. A final branching onset /...TRø/ in Polish ***<u>wiatr</u>*** [vʲatr] 'wind' is the most marked (but still grammatical!) because it is the most complex structure licensed by the weakest possible licenser.

(15) Relative markedness of contexts

a. context (traditional)	**effect**	**b. licensing scale**
_α	unmarked, no restrictions	_α
_ə	more marked, some restrictions	_ə
_#	most marked, severe restrictions	_ø

Prediction: If cross-linguistically the formal complexity scale (TR ⊃ RT ⊃ C) is indeed dependent on the licensing strength of nuclei, there should be an empirical reflection of that scale within individual systems when nuclei of different prosodic (and melodic) status are compared.

TR and RT clusters in Dutch

(16) Dutch TR clusters

a.

	^{ok}TRα	>	*$^{/ok}$TRə	>	*TRø
katrol ‘pulley’	[ka.trɔl]		*[ka.trəl]		*[katr]
duplo ‘duplicate’	[dy.plo]		*[dy.plə]		*[dypl]

Well-formed branching onsets are found in pretonic position in words like *fregat* [frəˈgat] ‘frigate’, *brevet* [brəˈvɛt] ‘patent’ (Kager 1989: 213).

b.

(17) Homorganic Dutch RT clusters

a.			b.		
[damp]	*damp* ‘vapour’		[χɛrt]	*Gert* ‘first name’	
[daŋk]	*dank* ‘thanks’		[boelt]	*bult* ‘hunch’	
[avənd]	*avond* ‘evening’		[vɛrs]	*vers* ‘fresh’	

(18) Non-homorganic Dutch RT clusters

a. *ə-epenthesis obligatory* (RTø ⟹ RəTø)

[harəp]	*harp* ‘harp’
[kɛrək]	*kerk* ‘church’
[balək]	*balk* ‘beam’
[hɛləm]	*helm* ‘helmet’

b. *ə-epenthesis optional* (RTə ⟹ R(ə).Tə)

[kar(ə).pər]	*karper* ‘carp’
[kɛr(ə).kər]	*kerker* ‘dungeon’
[stal(ə).kər]	*Stalker* ‘Stalker’
[hɛl(ə).mər]	*Helmer* ‘first name’

c. *ə-epenthesis excluded* (RTα ⟹ R.Tα)

[har.pun]	*harpoen* ‘harpoon’
[kar.kas]	*karkas* ‘carcass’
[bal.kan]	*Balkan* ‘Balkan’
[hɛl.ma]	*Helma* ‘first name’

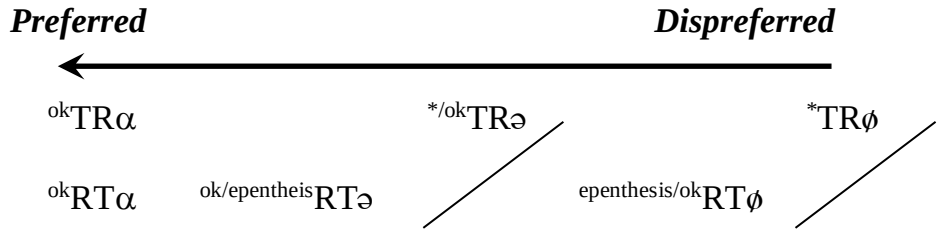
- the empirical scale above matches the scale of licensers (α > ‘ə’ > ø) in that the weaker the nucleus following the RT cluster the greater the chance that the cluster will be broken up.
- if the analysis of this scalar effect is correct, it supports the view that RT clusters are sanctioned in the ‘following syllable’ as it were.

(19) What is epenthesis?

a. T α R ϕ T α **b.** T α R ϕ T **'ə'** → T α R ə T **'ə'**

- the empty nucleus enclosed within a governing relation (19a) is 'locked' – it is invisible to phonological computation, it is not a licenser of its onset
- in (19b) the empty nucleus is called to do licensing, it is vocalised to be able to do this, no new V position is created, no resyllabification of any sort

(20) Preferred and dispreferred TR and RT clusters in Dutch

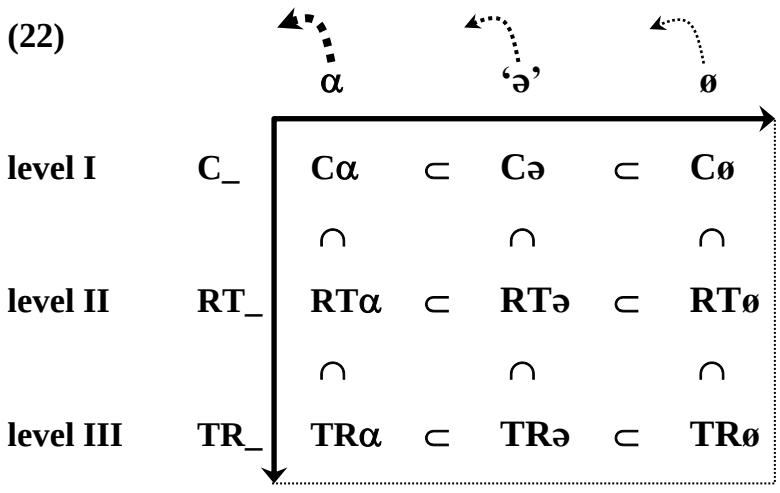


(21) Relative markedness of contexts

a. context (traditional)	**effect**	**b. licensing scale**
_α	unmarked, no restrictions	_α
_ə	more marked, some restrictions	_ə
_#	most marked, severe restrictions	_ø

Consequence II: The acquisition of the syllabic space

(22)



- the acquisition of the syllabic space consists in extending the two vectors in response to positive input.

- each configuration implies the grammaticality of the less marked configurations, that is, anything to the left and up
- the presence of a given string, eg. TRϕ, in Polish *wiatr* 'wind', tells us about the licensing properties of 'ϕ' and implies the presence of less demanding structures, RTϕ, Cϕ, and the presence of TR licensed by stronger licensers TRə, TRa. In fact, TRϕ implies the presence of everything else in the syllabic space! (good for learnability)
- each input structure strengthens the unmarked nature of Cα, because it is always at the end of the implication chain.
- *both theoretically and for the purposes of acquisition, a potential problem for the model constitute systems in which there are TRs but no RTs (Malayalam, Late Common Slavic).*

Consequence III: Shifts in licensing strength

- The licensing strength of nuclei is partly an abstract setting, which however has a concrete reflection in the amount of structure that is licensed (Complexity Scale: TR $\supset$ RT $\supset$ C), and partly depends on the melodic representation of the nucleus (α – 'ə' – ϕ).

  - **phonological shifts**
  - **register / dialectal (micro-variation)**
  - **historical change**

(23) How to become a stronger licenser?

a. Through **systemic strengthening**. Abstract, because the nucleus itself is not modified melodically (what is a stronger 'ə' or 'ϕ'?), yet concrete in that the licensed structure testifies to its strength, e.g. FOD blocking in Polish, *kod, dób*.
b. Through **melodic strengthening**. E.g. $\phi \rightarrow$ ə in Dutch epenthesis, sometimes an empty nucleus is vocalised word-finally to license its onset content, e.g. Malayalam /kaaṯ/ > [kaaṯə] 'ear', converging with register differences, e.g. [paalə] /[paal] 'milk'.

(24) How to become a weaker licenser?

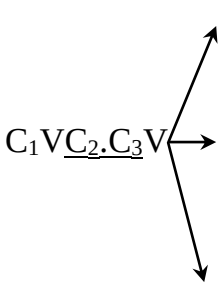
a. Through **systemic weakening**. Result: havoc in syllable structure
b. Through **melodic weakening**. E.g. vowel reduction $\alpha \rightarrow$ ə, syncope, apocope ə $\rightarrow \phi$
Weak licensers which are a new arrival in a system always cause structural changes

The puzzles of Late Common Slavic

- Law of open syllables
- Elimination of coda liquids in TART (but Polabian, Kashubian, Pomeranian)
- Non-elimination of coda liquids in TURT
- Loss of jers and reintroduction of closed syllables

Some of these puzzles can be understood better with Complexity Scales and Licensing Strength.

(25) Law of open syllables



$C_1V\ C_3V$	the simplification of clusters (C_2 = obstruent) e.g. **živ-ti* > *ži-ti* ‘live’
$C_1\tilde{V}\ C_3V$	the rise of nasal vowels (C_2 = nasal) e.g. PS **mĕnsă* ‘meat’, **rănkā* ‘hand’ > LCS **męso*, **rọka*
	(C_2 = liquid)
$C_1C_2\ V\ C_3V$	liquid metathesis, e.g. PS *pá**rg**u > Pl. **pr**óg
$C_1V\ C_2V\ C_3V$	pleophony *(polnoglasie)* > Russ. p**oró**g ‘threshold’
$C_1V\underline{C_2.C_3}V$	nothing!? PS *u̯irxu > Russ. verx, Pl. wierzch ‘top’

The phenomenon did not only affect coda-onset clusters but also the loss of final consonants, elimination of diphthongs, introduction of prosthetic consonants in place of empty onsets initially, e.g. Pl. *jeść* ‘eat’ and *jagnie* ‘lamb’ (cf. Lith. *ésti* ‘eat’ and Latin *agnus* ‘lamb’, respectively). It seems that also branching onsets were affected (the rise of non-ethymological jers), e.g. PS **ǫ**gl**ĭ* > Pl. wę**giel** / wę**gl**a ‘coal, nom.sg./ gen.sg.’.

After the process was complete, most syllables in Slavic were open, that is, they ended in a vowel and had almost no internal codas (CVCV).
While there were almost no codas (true?), complex clusters of the branching onset type were present. In this respect, C in the scheme CVCV, stands for a consonant or consonant cluster of rising sonority, with the exception of s+C.

(26) Elimination of coda liquids

	TART >	***trot / tret***	***trat / tret***	***torot / teret***	***tort / tret***
Gloss	*PS*	*Polish*	*Czech*	*Russian*	*Polabian*
‘cow’	*kár**u̯**ā	**kr**owa	**kr**áva	k**oró**va	kor**v**o
‘threshold’	*pá**rg**u	**pr**óg	**pr**áh	p**oró**g	po**rg**
‘frost’	*má**rz**u	**mr**óz	**mr**àz	m**oró**z	mo**rz**
‘fortification’	*gâ**rd**u	**gr**ód	**hr**ad	g**óro**d	go**rd**
‘shore’	*bê**rg**u	**br**zeg	**bř**eh	b**ére**g	**br**ig
‘milk’	*me**lk**á	**ml**eko	**ml**éko	m**olo**kó	**ml**ákă

Liquid metathesis is a potential problem to the Complexity Scale and Licensing model as a result RT clusters were eliminated and TR clusters created (less marked replaced with more marked, or is it?).

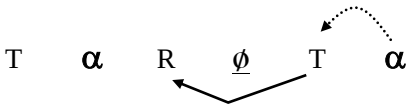
(27) Two facts about Late Common Slavic

a) short u/i became weak (jers) during Law of Open Syllables, and were subsequently lost
u/i → ъ/ь → ø
b) Late Common Slavic witnessed some dramatic changes with respect to prosody, a new quality emerged which spanned the Slavic dialects with various degrees of intensity:
a bisyllabic trochaic foot as a prosodic organizer (Bethin 1998).

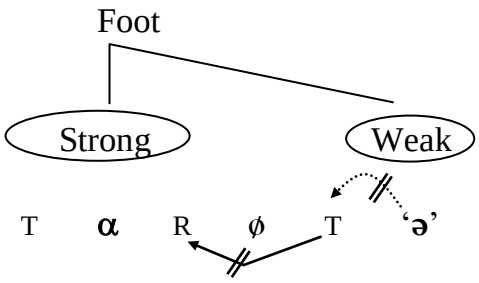
- on both counts, the nucleus directly following the RT clusters was weakened.

Analysis:

(28) initial stage



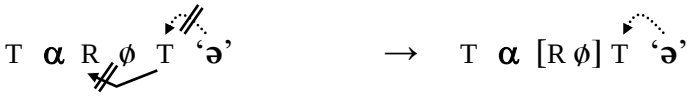
(29) the emergence of the trochaic foot



- prosodically and melodically weak licensers are a new arrival
- the inability of 'ə' to license RT (cf. Dutch) resulted in a number of repairs

(30) Law of open syllables due to weak licensing?

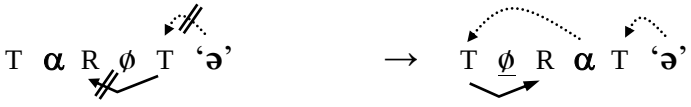
a. ***cluster simplification***



b. ***epenthesis***

c. ***redefinition of the licensing potential of 'ə'? or lack of weakening?***

d. ***metathesis***



Result: if Law of open syllables was indeed due to weakening of licensers, then the loss of jers and the reintroduction of 'closed' syllables is no longer a paradox. Both phenomena are strictly connected. Loss of jers was an inevitable next step. From then on, the FEN could only be strengthened – cf. Modern Polish.

(31) Chain of events:

emergence of trochaic foot > weakening of V_2 > RT cluster repairs

epenthesis TαRTə → TαRᵊTə due to weak licensing in second syllable (pleophony)

Q: why ***metathesis***? (strong binary foot?)

Q: Did metathesis create TRs at the cost of RTs? Yes.

Q: Is this problematic for the model? Normally it would but not in this case!

All that metathesis is showing is that in the desire to retain a binary foot the systems preferred TRα to RTə (TRα > RTə). The preference would be impossible only if the licensers were of the same type (Cf. TURT!)

Q: Why no metathesis in some Polabian forms?
This language exhibited a phenomenon which is sometimes referred to as progressive accent shift (e.g. **vórna* > *vornó* 'crow'), while the predominant direction of accent movement in LCS dialects was in the opposite direction (Bethin (1998: 159). The stress system of that language has also been analysed as a fixed final stress, with iambic metrical organization.

V_2 in Polabian could be strong!

Q: any modern reflection of LCS liquid metathesis? yes, Irish. Irish facts show that metathesis is linked to epenthesis and nuclear strength.

(32) Modern Irish has both TARTA → TRATA and TRATA → TARTA

a. ***stress related, liquid glues to strong licensing (cf. Polabian)*** ***TRə < RTα**

[ˈprɑʃigʲ] – [pərˈʃaχ]	*praisigh / praiseach*	'porridge, DATsg./NOMsg.'
[ˈbrɑdigʲ] – [bərˈdɑχ]	*bradaigh / bradach*	'thieving, GENsg./NOMsg.'

b. ***epenthesis related, liquid glues to stronger licensing (cf. LCS)*** **TRα > *RTə**

i.

initial form	***Munster Irish***	***Connacht Irish***		
/bolgəm/	boləgəm	blogəm	*bolgam*	'mouthful'
/tʹirʹməj/	tʹirʹim	tʹrʹuməj	*tirim / tiormaigh*	'dry'/ 'to dry'

ii.

****Proto-Slavic***	***hypothesised LCS***	***modern Polish***		
**kárų̑ā*	?korᵊva	krova	*krowa*	'cow'
**párgu*	?porᵊgu	pruk	*próg*	'threshold'
**gârdu*	?gorᵊdu	grut	*gród*	'fortification'

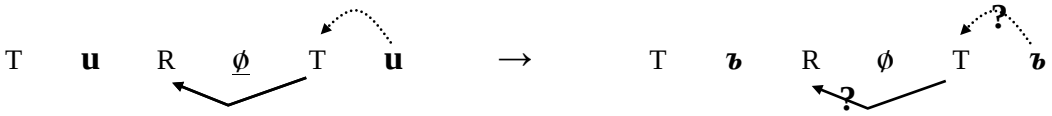
Final LCS puzzle: why no elimination of coda liquids in TURT (TURTU)?

(33) **TURTU** > ʔTЪRTЪ > ***tart*** ***tart / trat*** ***tṛt***

Gloss	*PS*	*LCS*	*Russian*	*Polish*	*Serbo-Croatian*
'top'	*u̯irxu		verx	wierzch	vr̂h
'death'	*sumirti		smert'	śmierć	sm̀r̀t
'wolf'	*u̯ilku		volk	wilk	vûk
'long'	*dulgu		dolg	dług	dûg

Analysis:

(34) **initial stage** → **weakening of u/i to jers *ъ/ь (='ə')***



(35) a. ***epenthesis*** ʔv ə r ə x ə must be assumed

b. ***pleophony*** *v ə r ə x ə no colour to copy no pleophony

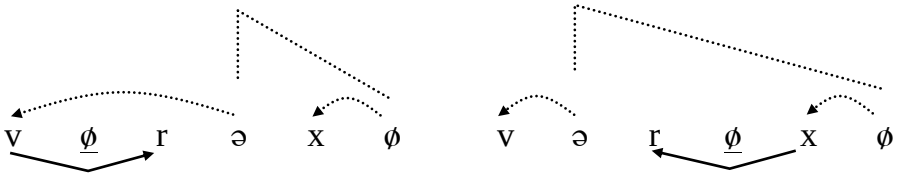
c. ***metathesis*** *v ϕ r ə x ə impossible: *TRə > RTə

(36) Modern Bulgarian liquid metathesis

Singular	***Definite sg.***	***Plural***	***Gloss***
vrəx	vərxə́t	vərxové	'top'
grəb	gərbə́t	gərbové	'back'
grəd	grədtá	gərdí	'bosom'

- In the Singular form, Modern Bulgarian has a slightly different choice than LCS because the final nucleus is empty. So it is TRə vs. RTϕ. Not TRə vs. RTə. When 'ə' appears after RT in the Definite sg. or Plural, then there is no choice! The liquid prefers RT.
- There is interesting dialectal variation in the singular but not in the Definite sg. This is what we predict.

(37) a. F b. F



(38) a. TRə 〉 RTϕ vrəx (Bulgarian)
b. TRə 〈 RTϕ vərx (Bulgarian dialectal, Polish, Russian)
c. TRə = RTϕ vṛx? (Bulgarian dialectal, Serbo-Croatian)?

- Bulgarian data can be viewed as a case of competition between the potential configuration *TR*ə and *RT*ϕ. And, to the extent that Bulgarian chooses the former rather than the latter (*TR*ə 〉 *RT*ϕ), it appears to re-live the dilemma of the LCS dialects which opted for metathesis of liquids in TARTə forms, or the dilemma of modern Irish where [boləgəm] loses to [blogəm].

When nuclei strengthen...

(39) The right edge in Polish

a. RT# in Polish

wt, lt, rt, nt, wp, lp, rp, mp, ɲp, wk, lk, rk, ŋk, l͡ts, r͡ts, ɲ͡ts, wt͡ɕ, rt͡ɕ, wt͡ʃ, lt͡ʃ, rt͡ʃ, nt͡ʃ, ɲt͡ʃ, lx, rx, mx, wf, rf, st, ʃt, sp, ʃp, ɕp, sk, ɕt͡ɕ, ʃt͡ʃ, sf, ʃf, pt, kt, tʃt, p͡ts, kt͡ʃ, pt͡ʃ, rɲ, rn, rm

b. TR# in Polish[1]

tr, pr, kr, fr, kl, pl, tl, tw, kw, tf, kf, tʃ, kʃ, xʃ, pʃ

(40) The distribution of vowel – zero alternations at right edge in Polish

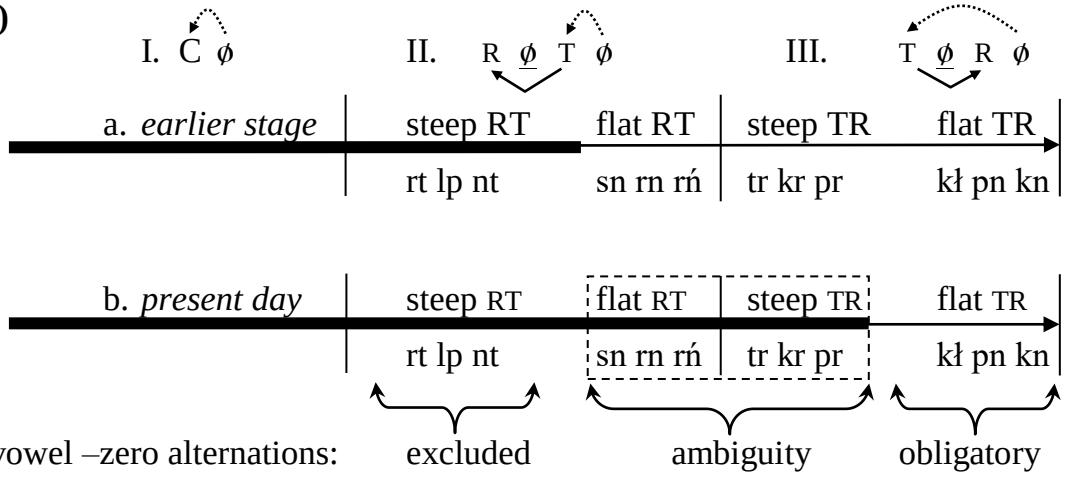
a. flat TR ***v – ø*** ***obligatory***
e.g. *ogień / ognia* 'fire, nom.sg. / gen.sg.'
okien / okno 'window, gen.pl. / nom.sg.'

b. steep TR ***v – ø*** ***common, ambiguity present***
e.g. *sweter / swetra* 'jumper, nom..sg. / gen.sg.'
wiatr / wiatru 'wind, nom.sg. / gen.sg.'

c. flat RT ***v – ø*** ***common, ambiguity present***
e.g. *darń* 'sod', *cierń* 'thorn'
dureń / durnia 'fool, nom.sg. / gen.sg.'

d. steep RT ***v – ø*** ***excluded***
e.g. gwałt 'rape', *czart* 'devil'

(41)



[1] Examples are: *wiatr* 'wind', *bóbr* 'beaver', *akr* 'acre', *szyfr* 'cipher', *cykl* 'cycle', *pejotl* 'peyotl', *plótł* 'he waffled', *mókł* 'he got wet', *modlitw* 'prayer, GENpl.', *tykw* 'gourd, GENpl.', *patrz* 'look, IMP.', *roziskrz* 'incite, IMP.', *wichrz* 'stir up, IMP.', *wieprz* 'pig'.

(19) Types and properties of nuclei in Polish (universal types and properties)

	type	*properties*
a.	**V** | α	- blocks interonset government - full licenser - distribution lexical / arbitrary and free
b.	**V** α	- blocks interonset government (may not be locked by IO) - may be a licenser; licensing properties the same as for empty nucleus if melody unassociated; licensing properties the same as for full vowel if melody linked - distribution lexical, and conditioned (it must be a licenser, and it must not be followed by another empty nucleus *Ø – Ø, or else the melody is linked); occurs only word-internally
c.	**V**	- does not block interonset government (may be locked by IO) - may be a licenser - distribution / lexical / partly arbitrary / partly predicable, and conditioned (it must be a licenser, and it must not be followed by another empty nucleus); occurs internally (IEN) and finally (FEN)