

DISTRIBUTIONAL GAPS IN SLAVIC INITIAL CLUSTERS ARE ACCIDENTAL

- (1) Typology of word-initial clusters (T=Obstruent, R=Sonorant), e.g. Clements (1990)

	#TR	#RT	example
a. no initial clusters	no	no	e.g. Ticuna (native Indian, Colombia)
b. #TR-only	yes	no	English, French etc.
c. anything-goes	yes	yes	modern occidental Semitic, Berber, Slavic
d. #RT-only	no	yes	does not exist

Introduction

- (2) purpose
- it is generally held that the distribution of consonants in word-initial clusters is the result of grammatical activity
==> distributional gaps are systematic, not accidental
 - difference:
 - systematic gaps exist because the missing clusters are ill-formed: grammar does not tolerate them.
==> new words bearing such clusters cannot enter the language.
 - accidental gaps exist because there happens not to be any lexical entry with the missing clusters. Grammar does not object against these clusters, and hence
==> new words bearing them can enter the language without problem.
 - example:
a word like Mcyri "poem by Lermontov" could never become an English word.
 - there are two kinds of anything-goes languages:
 - those where **all** logically possible #RT clusters indeed exist in real words.
Example: Moroccan Arabic (cf. illustration below)
 - those where some #RT clusters exist in real words, but some others do not occur in any word. Examples: Russian, Czech, Polish.
 - question: are distributional gaps in anything-goes languages just as systematic as in #TR-only languages?
 - regular answer: yes
 - my answer: no, they are accidental
 - why?
because of a prediction made by a particular theory, CVCV.

g. how can this be tested?

1. new words (loans, acronyms etc.) with a missing cluster can become regular words without any problem.

[Although #mc is not a native Russian cluster, the poem by Lermontov has become a regular Russian word (for those speakers who use it).]

2. I show that there is no rationale for dividing clusters in occurring and non-occurring ones: neither set forms a natural class by any possible criterion. Distribution is truly anarchic, i.e. lexical accident.

3. I show that ALL #RT clusters in ALL Slavic languages have been created through the loss of a yer: #C1-yer-C2 > #C1C2. Since of course there was no co-occurrence restriction between C1 and C2 in Common Slavic, the loss of the yer has promoted the CS lexical accident of the C1-C2 distribution to an initial cluster.

==> the resulting distributional pattern cannot be anything else than accidental and anarchic.

(3) typology of anything-goes languages

a. Moroccan Arabic

languages where any two consonants of the inventory are actually observed as the initial cluster of an existing word.

#TR	#RT	
brid	rbiT	cool down, bind
Drib	rDa	hit, accept
gliʕ	lga	remove, find
bka	kbir	cry, grow larger
nzil	zna	descend, commit adultery
dna	ndim	come near, regret
bqa	qbil	stay, accept

b. (ancient) Greek

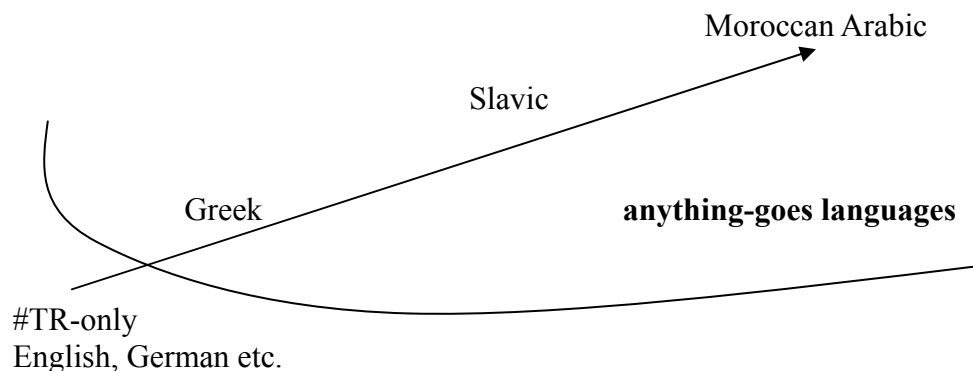
languages where just a few non-#TR clusters exist:

#pt, #kt and aspirated versions thereof, #mn

c. Russian, Czech, Polish, Ukrainian

languages where quite some, but not all logically possible non-#TR clusters exist.

d. hence a **scalar** classification of anything-goes languages between two poles?



e. where Greek is just a #TR-only language with some exceptions?

where the Indo-European unity also makes Slavic #TR-only with some more exceptions?

- (4) a particular phonological theory, CVCV, rebels against any scalar scenario:
- a. there are two and only two types of languages: 1) #TR-only, 2) anything-goes.
 - b. as soon as there is one single non-#TR cluster in a language, this language is anything-goes. It could not possibly be "#TR-only plus a few exceptions".
 - c. clusters that are absent from existing words are accidental gaps.

Background: CVCV

- (5) CVCV (Lowenstamm 1996, Szigetvári 1999, Scheer 2004a, Szigetvári & Scheer 2005) syllable structure boils down to a strict sequence of non-branching Onsets and non-branching Nuclei. The following representations for basic phonological objects ensue:

closed syllable	geminate	long vowel	[...C#]	branching Onset
O N O N	O N O N	O N O N	O N	O N O N
	/	/		
C V C ∅	C V	C V	C ∅ #	T ∅ R V

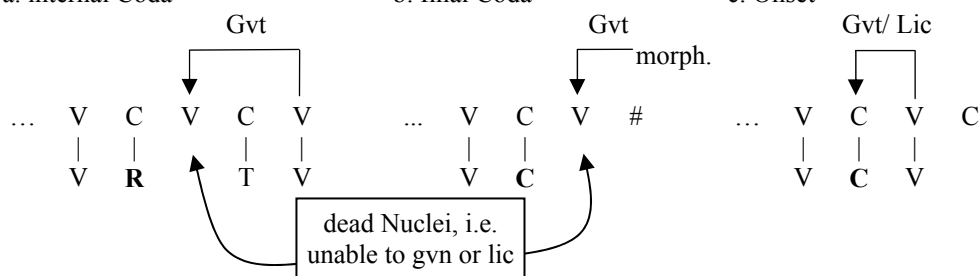
- (6) syllable structure
 - a. traditional: arboreal structure expresses co-occurrence restrictions and varying affinity among segments.
 - b. CVCV: this function is shifted onto **lateral relations** that are assumed to hold between constituents, **Government** and **Licensing**.
Effects that are usually attributed to the fact that a given segment belongs to this or that syllabic constituent are claimed to stem from the configuration regarding Government and Licensing that it is involved in.
 - c. the result is supposed not to be a null-sum game: The **lateralisation of structure and causality** buys you more than arboreal syllable structure.
- (7) illustration: the Coda
 - a. Coda phenomena are effects that are triggered by Codas and either appear on the Coda itself (lenition, devoicing etc.) or on the preceding vowel, in which case they are usually called closed syllables effects (vowel shortening, nasalisation etc.).
 - b. classical: the Coda disjunction is reduced to a non-disjunctive statement by saying that consonants in word-final and pre-consonantal position belong to a specific constituent, the Coda.
 - c. in CVCV, a coda consonant is a consonant that occurs before a governed empty Nucleus: **Coda** = __ $\emptyset$
==> this is **lateralisation of structure**: if you want to know whether a given consonant is an onset or a Coda, you don't **look up** (to the node they belong to), but **right** (whether the following Nucleus is empty or filled).

the Coda in CVCV: a consonant that occurs before a governed empty Nucleus

a. internal Coda

b. final Coda

c. Onset



(8) lateralisation of causality I: lenition, the challenge

a. illustration: **lenition**

[the Coda Mirror, Ségéral & Scheer 2001,2005, Scheer 2004a: §§110,556]

b. the traditional literature in Romance and Germanic recognizes the so-called **Strong Position**, i.e. where consonants are shielded against damage.

Strong Position = "word-initially and after a consonant" = {#,C}__

Coda = "word-finally and before a consonant" = __{#,C}

c. hence the challenge:

1. reduce the Strong Position disjunction to a non-disjunctive statement

2. explain why the Coda and the Strong Position are the exact mirror image of each other:

	structural description		segmental effect		syllabic analysis
Coda	__ {#,C}	=	weakness	=	before empty Nuclei
	vs.		vs.		vs.
Coda Mirror	{#,C}__	=	strength	=	after empty Nuclei

d. arboreal syllable structure fails on both challenges:

1. consonants in the Strong Position are Onsets, but intervocalic consonants are as well.

2. there is no reason why the Coda should be weak, rather than strong, and its Mirror strong, rather than weak.

(9) lateralisation of causality II: lenition in CVCV

[the Coda Mirror, cf. above]

a. Government and Licensing

two antagonistic forces: Government inhibits, Licensing enhances the segmental expression of its target.

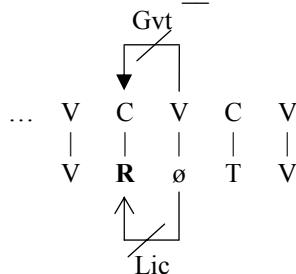
b. Nuclei need to be either phonetically expressed (contentful) or governed. Otherwise the structure is ill-formed.

[first approximation, more on that soon]

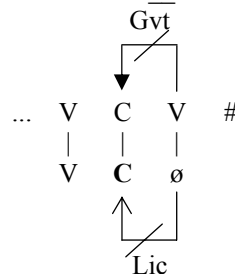
c. status of the 5 relevant positions:

consonants in Codas: ungoverned and unlicensed

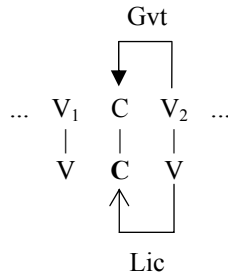
a. internal Coda __.C



b. final Coda __#

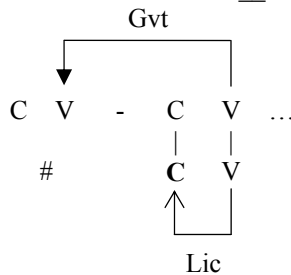


intervocalic consonants: governed and licensed

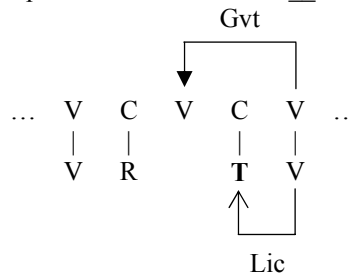


consonants in the Coda Mirror: ungoverned but licensed

a. initial consonant #__



b. post-Coda consonant C. __



d. the Strong Position

"after a heterosyllabic consonant" means "after a governing empty Nucleus" in CVCV.

==> **there must be an empty Nucleus preceding word-initial consonants**

e. this is the initial CV

[Lowenstamm 1999, Scheer 1999, 2004a:§83]

= CV

f. the top of the iceberg for another conception of the representation of extra-phonological information in phonology:

NOT through diacritics such as #, the Prosodic Hierarchy or the like,

but through truly phonological objects, i.e. that exist in the phonology independently of any issue related to the interface.

==> **Direct Interface**

[Scheer 2005a,b 2006, forth a,b]

(10) summary

a. properties of the 5 positions

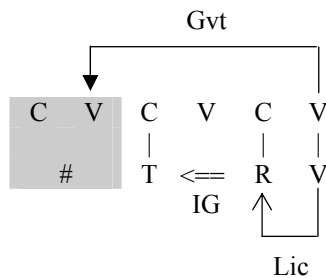
position	usual name	phonological identification	lateral situation	segmental health
1. #_V	word-initial	Coda Mirror = ø__	licensed but ungoverned	splendid
2. VC__V	post-Coda			
3. V__.CV	internal Coda	Coda = __ø	unlicensed and ungoverned	unfavourable
4. V__#	final Coda			
5. V__V	intervocalic	= elsewhere	licensed and governed	unfavourable

b. comparison

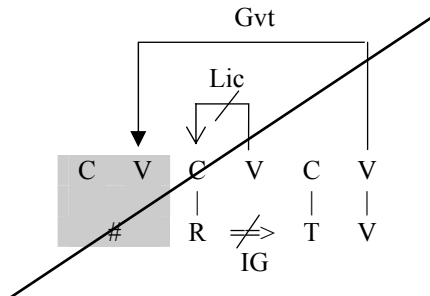
	arboreal	CVCV
identity of the Coda	the constituent "Coda"	<u> </u> ∅
identity of the Strong Position	no answer	∅ <u> </u>
why is the Coda weak?	no answer	-gvt, -lic
why is the Coda Mirror strong?	no answer	-gvt, +lic

(11) branching Onsets in CVCV

a. well-formed structure: #TRV...



b. ill-formed structure: #RTV...



b. [Scheer 2004a:§102]

1. IG = Infrasegmental Government
2. sonorants govern obstruents, but need to be licensed to do so by a full vowel.
3. a Nucleus sandwiched within a domain of IG may remain empty.

Back to initial clusters

(12) predictions:

- a. the initial CV makes any non-#TR cluster ill-formed.
- b. the initial CV makes word-initial consonants strong.
- c. so what about languages where
 1. non-#TR clusters exist
 2. word-initial consonants are weak?
- d. answer:
 1. these languages are the same
 2. they lack the initial CV
- e. Seigneur-Froli (2003,forth) shows that Greek combines both "typological" properties:
 1. it has non-#TR clusters
 2. word-initial consonants are weak (intervocalic)
- f. in terms of interface theory:

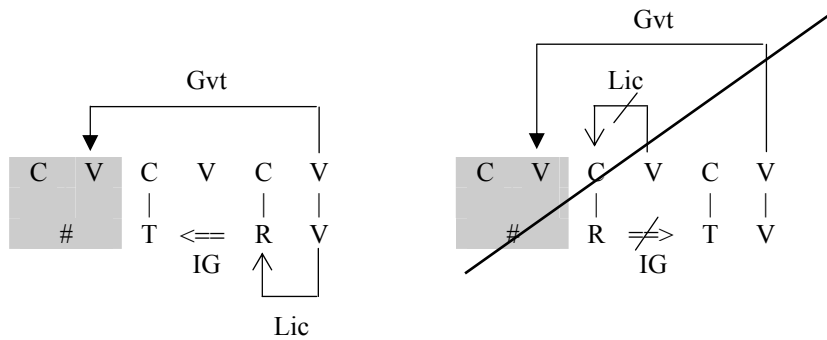
the distribution of the initial CV is a **parametric choice** made by the Translator's Office, i.e. the instance that transforms morpho-syntactic structure into phonological information (Prosodic Phonology).

==> the initial CV is phonological material of non-phonological (and non-lexical) origin.
- g. restrictions on initial clusters are only binary: #TR-only or **ANYTHING** goes
 - the initial CV is either present or absent
 - there is no third possibility
 - hence there are only two grammars regarding initial properties

1. languages that **have** the initial CV

#TR: well-formed

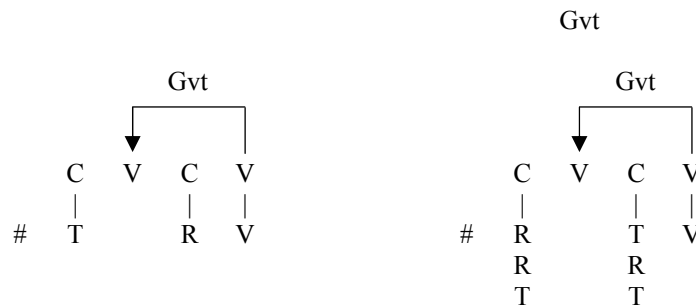
#RT: ill-formed



2. languages that **lack** the initial CV

#TR: well-formed

#RT: well-formed



- h. the presence of non-#TR clusters in a language guarantees the absence of the initial CV.
- i. hence in a language where some #RT clusters exist but others do not, the latter are accidental gaps: grammar is unable to distinguish any other restriction than "initial CV present vs. absent".

(13) immediate benefits I

the binary parameterisation of the initial CV gets the typology right

- a. #CV-only trivial: no clusters at all
- b. #TR-only presence of the initial CV
- c. #TR and #RT absence of the initial CV
- d. #RT-only cannot exist because the existence of #RT implies the absence of the initial CV, which in turn allows for any possible cluster.

(14) immediate benefits II

a better solution for extrasyllabicity

[Scheer 2004a:§339,2004b]

- a. the regular extrasyllabic analysis (e.g. Rubach & Booij 1990) predicts that there can be a random number of extrasyllabic consonants.
- b. this is because any number of unsyllabifiable initial consonants will be left unparsed by the syllabification algorithm (or equivalent constraints). They are then reintegrated into the Prosodic Hierarchy (adjunction) at a later derivational stage.

Depending on the analysis, they

- 1. either simply stand astray (Hall 1992, Wiese 1996)
- 2. or are adjoined to the Onset, and Onsets are then said to be able to violate Sonority Sequencing at the surface (but not when core syllabification takes place) [e.g. Hall 1992:122ss, 2000:248]

3. or are directly adjoined to some constituent of the Prosodic Hierarchy, e.g. the prosodic word, the phonological word (e.g. Rubach & Booij 1990, Rubach 1997).
- c. in any case there is no restriction defined regarding the number of extrasyllabic consonants that can be adjoined: how many consonants and of which sonority slope can Prosodic Word contain? On the grounds of which co-occurrence restrictions?
- d. by contrast, CVCV predicts that there can be one "extrasyllabic" consonant at most: any additional consonant implies an additional empty Nucleus, and two empty Nuclei in a row are ill-formed.

This appears to be a true statement, even for a "wild" language like Polish, cf. the detailed demonstration in Scheer (2004a:§373).

Two arguments

(15) argument 1

- a. if gaps are accidental, new words (loans, acronyms, nonce-words) with non-occurring #RT clusters can enter the language without problem.
If gaps are systematic, they cannot enter the language.
- b. borrowings of Georgian words with non-Russian #RT clusters into Russian

#RT		
#mts	Mcyri	poem by Lermontov, and the corresponding character
#mt	Mtacminda	mountain in Tbilisi
#mz	Mziuri	Georgian dance band
#mts	Mckheta	town in Georgia
#rk	rkaciteli	popular brand of wine
#rz	Rza	personal name
- c. acronyms with non-Czech #RT clusters in Czech
[must be socio-linguistically controlled, difference between speakers who use them frequently and those who do not. The latter vocalize every letter]

ČVUT	České vysoké učení technické
LFUK	Lekařská Fakulta University Karlova
JČU	Jihočeská Universita
JSA	Jazyk symbolických adres
LFOP	Lidová Fronta pro Osvobození Palestiny
LSU	Liberální Sociální Unie
LŠU	Lidová Škola Umnění
- d. other Slavic languages?
other sources?
- e. contrast with #TR-only languages into which the same items would not be able to sneak.
- f. the same must be true for Greek
 1. ongoing work on North-Eastern dialects (Lesbos) where pre-tonic syncope has "blindly" created #RT clusters: Seigneur (forth).
 2. ongoing work on the acquisition of Greek. Hypothesis: Greek infants can learn any non-#TR cluster and will accept items bearing them as regular words, while an English control group will not: Sanoudaki (forth).

(16) argument 2

- a. if in an anything-goes language gaps in #RT clusters are not accidental, they must be systematic.
Hence there must be some organizational principle that
 - 1) identifies the existing clusters as a natural class
 - 2) identifies the non-existing clusters as a natural class
- b. I show that for all Slavic languages with #RT clusters neither can be achieved.
Rather, the set of existing #RT clusters is lexical accident.
This lexical accident is the direct consequence of yer-loss.
Yer-loss is the only organizing principle for Slavic #RT clusters.

(17) empirical situation of Slavic #RT clusters

- a. corpus based on 13 Slavic languages
- b. on the grounds of the following dictionaries
 - Havlová, Eva, Adolf Erhart et alii 1989-1999. Etymologický slovník jazyka staroslověnského. 1ère - 9e livraison A-obrěsti. Praha: Nakladatelství Československé Akademie Věd/ Akademie Věd České Republiky.
 - Holub, Josef & František Kopečný 1952. Etymologický slovník jazyka českého. Praha: Státní Pedagogické Nakladatelství.
 - Holub, Josef & Stanislav Lyr 1978. Stručný etymologický slovník jazyka českého. Praha: Státní Pedagogické Nakladatelství.
 - Machek, Václav 1957. Etymologický Slovník Jazyka Českého a Slovenského. Praha: Nakladatelství Československé Akademie Věd.
 - Miklosich, Franz von 1886. Etymologisches Wörterbuch der Slavischen Sprachen. Reprint Amsterdam 1970: Philo Press.
- c. all data have been checked by native speakers (all of which in addition were linguists).
- d. feed-back from a posting on Linguist List (volume 12-358, published on February 12th, 2001)
- e. quite some words are archaic, uncommon, belong to the passive vocabulary etc. They typically occur in dictionaries, but natives are not very enthusiastic. The corpus has been cleaned in this respect, but a permissive policy has been adopted: if at least some natives have a passive competence for a given word, it has been retained.
- f. entries are ordered according to Common Slavic etymons.
- g. for each item, one line indicates whether it is represented with an #RT cluster in a particular language, and another line whether it is represented with a non-#RT cluster.
- h. result: 47 CS etymons, which are numbered from 1 to 47.

Example: #lateral-obstruent

[full corpus at www.unice.fr/dsl/rt/slavicRT.htm]

1

	Root	#CC	Common Slavic (Old Church Slavonic)	IE and comparatistic evidence	gloss CS	West					South				East				
						Czech	Slovak	Upper Sorbian	Lower Sorbian	Polish	Kashu-bian	Bulgarian	Mace-donian	Bosno-Serbo-Croatian	Slovenian	Belorussian	Ukrainian	Russian	
26	l̥b-	lb	l̥bъ	IE leubh-	skull GENsg NOMsg	lbi, rare lbu leb, lebi, rare lebu				lba leb								lbá (GENsg of lób)	
27	l̥g-ati	lg	l̥gati, l̥g-jo	NHG lügen	lie inf, 1sg	lhát, lžu				lgać, lže	lgac						lhaty	lgát', lgún	
						lež	luhat'	fać	dgaś			l̥'ž'a	laže	lagati	lagati	ilhać			
		lž	l̥ž-a			lie GENsg	lži												lžá (arch)
						lie NOMsg	lež			bža, bžě	dža, džy			l̥žá	laže	laž	laž		lož, olža
28	l̥g-	lg	l̥g-ъкъ, l̥go-stajъ	IE legwh-u-, skr laghú-, gr elakhys, lat levis, NHG leicht	light	lhostejný				lgi (arch)									
						lehký	l'ahký, l'ahostaj	lochki	lažki	lekki	letk'i	lék	lek	lak, lagan, laknuti	lahek, lahak	l'ochki	lehkyi	l'óhkij	
						respice, deadline	lhůta, topo Lhota										l'hota	l'góta, l'gá	
							lehký	lehota, topo Huta							odlog	il'hota			
		lz	l̥dza			it is suitable to	lze				lza, lza (arch)								l'žá
							lehký	nel'za							lahko	il'ha	nel'ha	nel'zjá	
29	l̥k	lk	l̥k	onom (s)luug-, NHG schlucken	mourn	lkát				lkać									
						po-lykat		lunk		połykać			l'oka						
30	l̥p-	lp	l̥p-	NHG bleiben, leben	cling, stick	lpět, lpít, lnout				lgnąć	lnanc						l'nuty	l'nút'	
						lepit		lěpić	lipaś	lepić		lepílo	lepak, lepi	lepiti	lepiti	il'nuć, lipnuć			
31	l̥sk-, l̥šč-	ls	l̥šč-ati (se) l̥sk-ati, l̥sk-ati	IE leuk-, gr leukhos, lat lux, OHG liht (> NHG Licht), skr ročate	shine, twinkle	lsknouti se (arch), lštíti se				lsknąć się, lsnać się, lsnić							l'sce		
						lesk, GENsg lesku		ščany (arch)	ščas se, šciś se	šlnić (arch)		l̥'skav, l̥štjá	leskot	laštiti se	lesk, leščati (arch), lesketati	il'snicca	il'sce	losnít' sa	

	Root	#CC	Common Slavic (Old Church Slavonic)	IE and comparatistic evidence	gloss CS	West					South				East			
						Czech	Slovak	Upper Sorbian	Lower Sorbian	Polish	Kashu- bian	Bulgarian	Mace- donian	Bosno- Serbo- Croatian	Slovenian	Bielorussian	Ukrainian	Russian
32	лѣтъ		лѣтъ < OHG listiz (> NHG List)		cunning, ruse	lsti, listivost, listivý, lstný				łści (arch), łściwy							л'стіт'	
						lest	lest'	lesć	lasć	leść (arch)		л'st (arch)		last (arch), lastan	lest (arch)	lestъ (arch), lislivić	l'est'	l'ést'
33	лѣвъ	lv	лѣвъ	< CGerm *liuwaz (> NHG Löwe), cf. lat leo, gr leon	lion GENsg	lva, lví, lvíče, lvíce, lvoun				lwa							l'va	lvá
					lion NOMsg	lev	lev	law	law	lew	lev	л'v	lav	lav	lev, GENsg leva	leu, GENsg il'va	lev	lév
34	слѣз-	lz	слѣза, слѣза	*lugjō, NHG schlucken	tear					łza, łzawy								
						slza		sylza	dza	łez (GENpl)					solza			
35	лѣз-	lž	лѣзica, лѣзѣка	lat ligula	spoon	lžice		lžica										
						žlíce (dial)	lyžica		žyca	łyżka		лѣзica	lažica	Cr žlica	žlica	lyžka	ložka, lyžka	lóżka

(18) summary

distribution of #RT clusters among Slavic languages

		West					South				East			
#RT	cluster	Cz	Sk	Po	USo	LSo	Ka	Bu	Mac	SC	Sn	Ru	Uk	Bru
j+T	jd	+			+									
	jh	+												
	jm	+	+										+	
	js	+												
r+T	rb									+				
	rt̩s	+		+										
	rt̩ʃ	+								+				
	rk, řk	+												
	rd, rd̩z, rd̩z̥	+		+						+		+		
	rz	+								+				
	r̩3	+		+	+					+		+	+	
	rf													
	rs													
	rt	+		+						+		+	+	
	rv, řv	+		+						+		+	+	
	l+T	lb	+		+								+	
lg, lh		+		+			+					+	+	
l̩3		+										+	+	
lz		+		+	+									
lk		+		+										
lp		+												
ls, l̩ç		+		+								+		
m+T	l̩ʃ													+
	lv	+		+								+	+	
	md	+	+	+										
	mg, mh	+	+	+			+					+	+	
	m̩3	+		+								+	+	
	mz													
	mx			+								+		
	m̩ʃ	+	+	+	+		+					+	+	
	mk	+		+			+					+		
	mt̩ʃ													+
	ms, m̩ç	+		+								+	+	
	mz	+	+	+	+							+	+	
n+T	mt	+												
	absent													

(19) generalisations

- lack #RT altogether:
Lower Sorbian, Bulgarian, Macedonian, Slovenian, Bielorussian
- only #rT occurs, and r in this case is always syllabic
Serbo-Croatian
- poverty of sources
Upper Sorbian, Kašubian
- 5 languages with #RT clusters
Czech, Slovak, Polish, Russian, Ukrainian
- the Slovak #RT words may turn out to be presently vanishing from the language:
almost all words belong to the passive vocabulary.

(20) zoom on Polish and Czech, the languages where #RT clusters are most widespread

a. choice among possible #RT made by Czech and Polish

b. Polish:

126 possible #RT sequences (6 sonorants, 21 obstruents), 22 occurring #RT clusters (18%).

c. Czech:

108 possible clusters (6 sonorants, 18 obstruents), 27 occurring #RT clusters (25%).

		Polish									Czech						
		C ₁	j	l	r	ɲ	n	m			C ₁	j	l	r	ɲ	n	m
C ₂	p								p								p
	t				+				t							+	t
	k			+				+	k				+	+		+	k
	b			+					b								b
	d				+			+	d		+		+			+	d
	ɡ			+				+	ɡ								ɡ
	<u>ts</u>				+				<u>ts</u>								<u>ts</u>
	<u>tʃ</u>								<u>tʃ</u>					+			<u>tʃ</u>
	<u>tɕ</u>								<u>tɕ</u>								c
	<u>dz</u>				+				<u>dz</u>								ʃ
	<u>dʒ</u>								<u>dʒ</u>								f
	<u>dʒ̥</u>								<u>dʒ̥</u>								v
	f								f				+	+		+	s
	v			+	+				v					+	+	+	z
	s			+				+	s							+	ʃ
	z			+				+	z							+	ʒ
	ʃ							+	ʃ								x
	ʒ				+			+	ʒ								h
	ɕ			+				+	ɕ								
	ʒ̥								ʒ̥								
	x							+	x								
			j	l	r	ɲ	n	m				j	l	r	ɲ	n	m

(21) Polish, a well studied language I

a. initial clusters have been extensively studied in Polish:

Kuryłowicz (1952), Rubach & Booij (1990), Gussmann (1991), Cyran & Gussmann (1998,1999), Sawicka (1974), Rowicka (1999:309ss)

b. exhaustive list of two-membered initial clusters

[following Sawicka 1974, Rowicka 1999, see also Scheer 2004a: §§375,622]

c. statistics

1. 616 logically possible clusters: 22 possibilities for C₁ times 28 possibilities for C₂ (s-sounds = [s,z,ʃ,ʒ,ɕ,ʒ̥] are counted out for C₁)

2. 130 clusters attested (21%)

of which 56 respect sonority sequencing, against 74 violating it

d. attempts at discovering an organizing principle have been relatively more or less successful, but none

- could eliminate the number of clusters that should exist but do not

- could eliminate those that exist but should not

==> no analysis can characterize the set of existing and non-existing clusters as a natural class.

Gaps are randomly distributed.

(22) Slavic #RT are a lexical accident

a. diachronic origin of modern Slavic #RT clusters

#RT	nb roots < #RvT	uncertain origin
	< #RyerT	< #RvT
#jC	4	1 (5 j-es-mъ)
#rC	15	4 (14 rufijanъ 15 rusъ 21 rjuti 25 rez)
#lC	10	0
#mC	12	0
	41	5
		1
		Total 47

b.

all Slavic #RT < #RvT

c. this cannot be accidental

there is a causal relation between the loss of yers and modern #RT clusters.

d. the modern anarchic distribution of #RT clusters is obvious when knowing that all #RTs have come into being through the loss of an intervening yer:

Since of course there was no co-occurrence restriction between C1 and C2 in a Common Slavic #C1-yer-C2 sequence, the loss of the yer has promoted the CS lexical accident of the C1-C2 distribution to a modern initial cluster.

e. this also explains the strange situation of [m] and [n]:

- if #NT clusters are allowed, why not #nT sequences?

- if any discrimination obtains between both, #nT is expected to exist.

==> however, #mT exists (and is actually quite frequent), while there is not a single word with an #nT cluster in any Slavic language.

f. prediction if the diachronic hypothesis is correct:

there were #m-yer-T roots in CS, but no #n-yer-T roots (at least they did not survive).

Etymological dictionaries (e.g. Havlová & Erhart (eds) 1989-2002:557s, Miklosich 1886:218, Holub & Kopečný 1952:241, Machek 1957:321) know only one candidate root, CS *nъštvi "trough" (< IE *nigw "washing", e.g. gr víζειν "to wash"), but which has experienced (irregular) vocalisation of the yer in all modern reflexes: s-cr načve, cz necky, pol niecka, old rus načvy.

Hence there are no #nT clusters simply because there was no etymological basis.

(23) the numeric extrapolation of the 47 roots is sound

a. why are #RT roots so rare in Slavic?

Because the only possible source of modern #RT is #R-yer-T.

b. let us see whether 47 roots is a reasonable number if only 2 vowels have fallen out.

c. assumptions

1. linear distribution of consonants and vowels in Common Slavic #C₁V₁C₂V₂ sequences

2. disregard of diachronic loss of roots.

d. calculus

1. eleven vowels in Common Slavic: <i,y,e,ě,a,o,u,ę,ǫ,ɤ,ɤ>

2. two elevenths of the Common Slavic #C₁V₁C₂V₂ sequences can be assumed to have produced #C₁C₂V₂ items.

3. obstruent inventory of Common Slavic: 20 items
14 obstruents: <p,b,v,d,t,s,z,š,ž,c,č,k,g,ch>
6 sonorants: <m,n,ɲ,r,l,j>
4. probabilities:
 - #TV₁RV₂ represents 14/20 for T in C₁ and 6/20 for R in C₂: $7/10 \times 3/10 = 21\%$
 - #RV₁TV₂ has the same probability: 21 %
 - #TV₁TV₂ represents $14/20 \times 14/20 = 49\%$
 - #RV₁RV₂ has a probability of $6/20 \times 6/20 = 9\%$
5. only two elevenths of the #C₁V₁C₂V₂ sequences, namely #C₁yerC₂V₂, produce #C₁C₂V₂ items.
6. hence 47 #RT roots = 21% of the Common Slavic #C₁V₁C₂V₂ stock that is 3,81%
7. ==> Common Slavic had 1231 #C₁V₁C₂V₂ roots - a reasonable number.

(24) conclusion

- a. languages may have no, some, quite some, a lot or all possible #RT clusters.
- b. this surface gradation grows on the grounds of just two possible grammatical situations:
 - either grammar imposes #TR-only
 - or grammar imposes no restriction at all
- c. if a language has #RT clusters and no matter how complete they are, it may have any initial cluster.
- d. in addition of the arguments made, acquisition data are most promising: infants of #RT-languages will be able to acquire any initial cluster, while infants of #TR-only languages will not.
- e. this approach explains why languages cannot have five, seven or 15 extrasyllabic consonants in a row.

References

WEB: references followed by this mention are available at www.unice.fr/dsl/tobias.htm.

- Clements, George 1990. The role of the sonority cycle in core syllabification. Papers in Laboratory Phonology I, edited by John Kingston & Mary Beckmann, 283-333. Cambridge: Cambridge University Press.
- Cyran, Eugeniusz & Edmund Gussmann 1998. Polish consonantal sequences: a phonological testing ground. Structure and Interpretation, edited by Eugeniusz Cyran, 127-138. Lublin: Pase. WEB.
- Cyran, Eugeniusz & Edmund Gussmann 1999. Consonant clusters and governing relations: Polish initial consonant sequences. The syllable, Views and Facts, edited by Harry van der Hulst & Nancy Ritter, 219-248. Berlin, New York: de Gruyter.
- Gussmann, Edmund 1991. Polish syllable structure: a hypothesis and its problems. Words are Physicians for an Ailing Mind, edited by Maciej Grochowski & Daniel Weiss, 207-213. München: Sagner.
- Hall, Tracy 1992. Syllable Structure and Syllable-Related Processes in German. Tübingen: Niemeyer.
- Hall, Tracy 2000. Phonologie. Eine Einführung. Berlin & New York: de Gruyter.
- Havlová, Eva & Adolf Erhart (eds) 1989-2002. Etymologický slovník jazyka staroslověnského. 1st - 11th booklet A-poditi. Praha: Nakladatelství Československé Akademie Věd/ Akademie Věd České Republiky.

- Holub, Josef & František Kopečný 1952. Etymologický slovník jazyka Českého. Praha: Státní nakladatelství učebnic.
- Kuryłowicz, Jerzy 1952. Uwagi o polskich grupach spółgłoskowych [Remarks on Polish consonantal groups]. *Biuletyn Polskiego Towarzystwa Językoznawczego* **11**, 54-69.
- Lowenstamm, Jean 1996. CV as the only syllable type. *Current trends in Phonology. Models and Methods*, edited by Jacques Durand & Bernard Laks, 419-441. Salford, Manchester: ESRI. WEB.
- Lowenstamm, Jean 1999. The beginning of the word. *Phonologica 1996*, edited by John Rennison & Klaus Kühnhammer, 153-166. La Hague: Holland Academic Graphics. WEB.
- Machek, Václav 1957. Etymologický slovník jazyka českého a slovenského. Praha: Nakladatelství Československé Akademie Věd.
- Miklosich, Franz von 1886. Etymologisches Wörterbuch der Slavischen Sprachen. Reprint Amsterdam 1970: Philo Press.
- Rowicka, Grażyna 1999. On Ghost vowels. A Strict CV Approach. Ph.D. dissertation, Leiden University.
- Rubach, Jerzy 1997. Extrasyllabic Consonants in Polish: Derivational Optimality Theory. *Derivations and Constraints in Phonology*, edited by Iggy Roca, 551-581. Oxford: Clarendon.
- Rubach, Jerzy & Geert Booij 1990. Edge of constituent effects in Polish. *Natural Language and Linguistic Theory* **8**, 427-463.
- Sanoudaki, Eirini forth. The acquisition of non-occurring initial consonant clusters by Greek infants. Ph.D dissertation, University College London.
- Sawicka, Irena 1974. Struktura grup spółgłoskowych w językach słowiańskich. Wrocław, Warszawa, Kraków, Gdańsk: Wydawnictwo Polskiej Akademii Nauk.
- Scheer, Tobias 1999. A theory of consonantal interaction. *Folia Linguistica* **32**, 201-237. WEB.
- Scheer, Tobias 2004a. A Lateral Theory of Phonology. Vol.1: What is CVCV, and why should it be? Berlin: Mouton de Gruyter.
- Scheer, Tobias 2004b. A better solution for extrasyllabicity than extrasyllabicity. Paper presented at GLOW 27, Thessaloniki 19-21 April. WEB.
- Scheer, Tobias 2005a. We need a translator's office, but the buffer has to go: Direct Interface. Paper presented at the 36th Poznań Linguistic Meeting, Poznań 22-24 April. WEB.
- Scheer, Tobias 2005b. When higher modules talk to phonology, they talk to empty Nuclei. Paper presented at the conference Sounds of Silence, Tilburg 19-22 October. WEB.
- Scheer, Tobias 2006. Interface Dualism. Paper presented at the 37th Poznan Linguistic Meeting, Poznan 20-23 April. WEB.
- Scheer, Tobias forth a. A Lateral Theory of Phonology. Vol.2: On Locality, Morphology and Phonology in Phonology. Berlin: Mouton de Gruyter.
- Scheer, Tobias forth b. Why the Prosodic Hierarchy is a diacritic and why the Interface must be Direct. To appear in the proceedings of the Tilburg Sounds of Silence conference. WEB.
- Ségéral, Philippe & Tobias Scheer 2001. La Coda-Miroir. *Bulletin de la Société de Linguistique de Paris* **96**, 107-152. WEB.
- Ségéral, Philippe & Tobias Scheer 2005. What lenition and fortition tells us about Gallo-Romance *Muta cum Liquida*. *Romance Languages and Linguistic Theory 2003*, edited by Twan Geerts, Ivo van Ginneken & Haike Jacobs, 235-267. Amsterdam: Benjamins.
- Seigneur-Froli, Delphine 2003. Diachronic consonant lenition & exotic word-initial clusters in Greek: a unified account. *Studies in Greek Linguistics. Proceedings of the 23rd annual meeting of the department of linguistics*, edited by M. Stavrou-Sifaki & A.

- Fliatouras, 345-357. Thessaloniki: University of Thessaloniki. WEB.
- Seigneur-Froli, Delphine forth. Cudas initiales et lénition en grec et ailleurs. Ph.D dissertation, University of Nice.
- Szigetvári, Péter 1999. VC Phonology: a theory of consonant lenition and phonotactics. Ph.D dissertation. Eötvös Loránd University, Budapest.
- Szigetvári, Péter & Tobias Scheer 2005. Unified representations for the syllable and stress. *Phonology* **22**, 37-75.
- Wiese, Richard 1996. *The Phonology of German*. Oxford: Oxford University Press.