Repetition Avoidance and the Exceptional Reduplication Patterns of Indo-European*

Sam Zukoff, szukoff@mit.edu
3/4/15

1. Introduction

- In the reduplicative systems of Ancient Greek, Gothic, and Sanskrit, we see differences in copying pattern dependent on the shape of the root-initial sequence.
  - Roots with an initial *consonant-vowel* (CV) sequence show C_1-copying.
    - $\sqrt{}C_1V$- → reduplicated $\underline{C_1V}$-C_1V-
  - Roots with initial *stop-sonorant* (TR) clusters tend to follow this default C_1-copying pattern.
    - $\sqrt{}T_1R_2V$- → reduplicated $\underline{T_1V}$-T_1R_2V-
  - However, roots with other initial clusters, notably *s-stop* (ST), display some other, distinct pattern:

(1) Non-default copying patterns in the Indo-European languages

	Copying Pattern	Base	Reduplicated form
Ancient Greek	Non-copying	$\sqrt{}S_1T_2V$-	$\underline{V}$-S_1T_2V-
Gothic	Cluster-copying	$\sqrt{}S_1T_2V$-	$\underline{S_1T_2V}$-S_1T_2V-
Sanskrit (cluster-initial roots)	C_2-copying	$\sqrt{}S_1T_2V$-	$\underline{T_2V}$-S_1T_2V-
Sanskrit (zero-grade bases)	"$C_1\bar{e}C_2$" pattern	S_1T_2-	$S_1\bar{e}T_2$-

- In addition to differing in the nature of the non-default pattern, the languages also vary in which types of clusters pattern with TR and which types pattern with ST.

❖ In this paper, I propose that these effects are all avoidance strategies for a single problem:

➢ **C_1-copying is blocked when it is *too difficult* to perceive the presence of root-C_1.**

- This will be formalized as the interaction between the (non-)availability of phonetic cues (cf. Wright 2004) and the principle of *repetition avoidance* (cf. Walter 2007).

Roadmap

* Special thanks to Donca Steriade, Adam Albright, Edward Flemming, Ryan Sandell, and Tony Yates. Thanks also to Andrew Byrd, Michael Kenstowicz, Jesse Lundquist, Juliet Stanton, Suyeon Yun, the audience at the 2015 LSA Annual Meeting, and audiences at MIT. All mistakes are, of course, my own.

2. (Im)perceptibility in $C_\alpha VC_\alpha C_\beta$ sequences

- It is well-known that there are biases against repetition in human language, and human cognition more generally (e.g., Walter 2007, and citations therein).
    - Walter (2007) demonstrates that, in phonology, there are both articulatory and perceptual biases against repetition, particularly against repetition of consonants in a local domain.
    - ➢ One specific bias in perception is "repetition blindness" (Kanwisher 1987), whereby subjects are unable to perceive repeated tokens as being separate entities (Walter 2007: chapter 5).
- It is also well-known that consonants are dispreferred in contexts where they are less perceptible ("Licensing by Cue"; Steriade 1997).

❖ A logical extension is that, when both of these conditions obtain in the same context, that context will be especially dispreferred:

(2) THE POORLY-CUED REPETITION PRINCIPLE (PCR):

> A CVC sequence containing identical consonants ($C_\alpha VC_\alpha$) is dispreferred, due to repetition blindness; it is especially dispreferred if one or both of the consonants do not bear phonetic cues which are important for the perception of its presence (in contrast to zero) in the speech signal.

- The intuition is the following:
    - Listeners are biased by *repetition blindness* to fail to identify the presence of a locally-repeated segment.
    - Listeners have difficulty recovering the presence of a consonant when it lacks robust phonetic cues to its presence.
    - When both of these conditions hold, accurate perception of the speech signal is especially difficult.

- I propose that this can project a constraint in the phonological grammar, such that these sequences may be actively avoided:

(3) POORLY-CUED REPETITION (PCR):

Assign a violation mark * to any $C_\alpha VC_\alpha$ sequence where the second consonant[1] does not bear the **requisite phonetic cues** to its presence.

➢ This constraint can begin deriving the differences between TR and ST clusters.

---

[1] Because we are dealing with a (#)$C_{\alpha1}VC_{\alpha2}C_\beta$ sequence, the first consonant will always be maximally-cued; the (#)_V context is the optimal context for perception of a consonant (see Wright 2004). Therefore, only perception of the second of the repeated consonants is at stake here.

- Among the phonetic cues which are most significant to perceiving the **presence** of a consonant[2] are *burst*, *intensity rise*, and *consonant-to-sonorant* (*CR*) *transitions*[3] (see Wright 2004).
  - All of these cues are present for a stop (T) before a sonorant (R).
  - None of them are present for a fricative (S), which inherently has no *burst*, before a stop, which can host neither an *intensity rise* nor *transitions*.
    - The *frication noise* of the fricative is normally a strong cue to its presence, particularly the high-intensity frication of sibilants.
    - It seems likely that *repetition blindness* may decrease the efficacy of this cue more than others; this will be discussed below (§4).

- Focusing first just on the TR vs. ST distinction, we can say that a consonant is *requisitely-cued* if it bears all of the following cues: *burst*, *intensity rise*, and *CR transitions*.

(4) PCR [*for TR vs. ST*]:

Assign a violation mark * to any $C_\alpha VC_\alpha$ sequence where the second consonant does not bear the requisite phonetic cues to its presence.

►**REQUISITE CUES**: *burst*, *intensity rise*, and *CR transitions*

3. Indo-European partial reduplication: TR vs. ST

- Ancient Greek (§3.1), Gothic (§3.2), and Sanskrit (§3.3) each display distinct behavior of TR-clusters vs. ST-clusters in reduplication.
- We will see that these can differences can be explained by the PCR.

3.1. Non-copying ST perfects in Ancient Greek

- Ancient Greek shows default C_1-copying when the root begins in a *stop-sonorant* (TR) cluster (shown in (5)), but “non-copying” in roots with initial *s-stop* (ST) (shown in (6)).

(5) <u>C_1-copying perfects to TR roots in Ancient Greek</u>

Root		Perfect Tense	
kri-	‘decide’	κέκριμαι [<u>k-e</u>-kri-mai]	not **[<u>e</u>-kri-mai]
tla-	‘suffer, dare’	τέτληκα [<u>t-e</u>-tlē̆-ka]	not **[<u>e</u>-tlē̆-ka]
pneu-	‘breathe’	πέπνυμαι [<u>p-e</u>-pnū-mai]	not **[<u>e</u>-pnū-mai]

(6) <u>Non-copying perfects to ST roots in Ancient Greek</u>

Root		Perfect Tense	
stel-	‘prepare’	ἔσταλκα [<u>e</u>-stal-ka]	not **[<u>s-e</u>-stal-ka]
strateu-	‘wage war’	ἐστράτευμαι [<u>e</u>-strateu-mai]	not **[<u>s-e</u>-strateu-mai]

[2] These are not necessarily the same cues which are most relevant for distinguishing the **place** of the consonant.

[3] I intend “CR transitions” to name the set of *consonant-to-vowel* (CV), *consonant-to-liquid* (CL), and *consonant-to-nasal* (CN) transitions, which stand in a stringency relationship: CV > CL > CN.

- With the current definition of the PCR from (4), we can motivate a difference between TR and ST roots. The actual repair is dependent on the ranking of a number of other constraints.

(7) Constraints modulating potential repairs for PCR

A. ONSET:
Assign a violation mark * for each onsetless syllable.
Potential PCR Repair: $\underline{V}$-C_1C_2V- (candidates (b))

B. C/V (≈ *COMPLEX):
Assign a violation mark * for every consonant which does not precede a vowel.[4]
Potential PCR Repair: $\underline{C_1C_2V}$-C_1C_2V- (candidates (c))

C. ANCHOR-L-BR:
Assign a violation mark * if the segment at the left edge of the reduplicant does not stand in correspondence with the segment at the left edge of the base.[5]
Potential PCR Repair: $\underline{C_2V}$-C_1C_2V- (candidates (d))

- To specifically generate the non-copying repair in Greek, ONSET must be the lowest ranked of these constraints, and it must also be dominated by PCR.[6]

(8) Non-copying in ST roots in Greek (PCR violation): √*stel-* → *e-stal-ka* 'I have made ready'

/RED, e, stal, a/	ANCHOR-L-BR	C/V	PCR	ONSET
a. s-e-stal-ka		*	*!	
b. ☞ e-stal-ka		*		*
c. st-e-stal-ka		**!		
d. t-e-stal-ka	*!	*		

- Candidate (a) is the default C_1-copy form – it is blocked from surfacing by PCR.
- The alternative candidates (b-d) each obviate PCR by avoiding the creation of the problematic repetition.
    - Candidate (d) does so by copying C_2, but fatally violates ANCHOR-L-BR.
    - Candidate (c) does so by copying the entirety of the root-initial cluster, interrupting the repetition with C_2, but this results in an extra C/V violation.
- The optimal candidate (b) copies nothing, at the expense only of low-ranked ONSET, whose violation is tolerable in service of PCR.

---

[4] I will only mark violations of C/V that arise from root-initial and reduplicant clusters.

[5] I assume that this constraint is not violated (i.e. vacuously satisfied) if no segments have been copied, as in the "non-copying" forms of Ancient Greek (e.g. candidate (8b)). In Sanskrit, ANCHOR-L-BR will not be vacuously satisfied by the non-copying candidates, because its patterns involve copying of a root-vowel (Steriade 1988). Sandell & Zukoff's (2014) synchronic analysis of the Gothic preterite system entails copying of the root vowel with consistent phonological reduction to [e] (i.e. phonological fixed segmentism).

[6] I omit MAX- and DEP-violating candidates for reasons of space. These constraints necessarily dominate C/V, as clusters are obviously permitted in all these languages.

- When the repetition caused by copying C_1 is well-enough cued to satisfy PCR, the ONSET violation incurred by the non-copying candidate becomes unnecessary, and so C_1-copying is permitted, as shown in (9):

(9) C_1-copying in TR roots in Greek (no PCR violation): √*kri-* → *k-e-kri-mai* 'I have been judged'

/RED, e, kri, mai/	PCR	ONSET
a. ☞ k-e-kri-mai		
b. e-kri-mai		*!

3.2. Cluster-copying reduplicated preterites in Gothic

- In Gothic there are not many relevant examples, but they again point to a distinction between TR and ST roots.
  - TR roots follow the default C_1-copying pattern (as seen in (10)).
  - ST roots display cluster-copying, i.e. a reduplicant in *STe-* (as seen in (11)).

(10) C_1-copying preterites to TR roots in Gothic (forms from Lambdin, 2006: 115)

	Infinitive	Preterite	
'to weep'	*gretan* [grēt-an]	*gaigrot* [gɛ-grōt]	not **grɛ-grōt

(11) Cluster-copying preterites to ST roots in Gothic

	Infinitive	Preterite	
'to possess'	*staldan* [stald-an]	*saistald* [stɛ-stald]	not **[sɛ-stald]
'to divide'	*skaidan* [skaið-an]	*skaiskaiþ* [skɛ-skaiθ]	not **[sɛ-skaiθ]

- This pattern falls out if we take the constraints and rankings proposed for Ancient Greek and simply swap ONSET and C/V:

(12) Cluster-copying in ST roots in Gothic (PCR violation): √*stald-* → *ste-stald* 'he possessed'

/RED, stald/	ANCHOR-L-BR	ONSET	PCR	C/V
a. se-stald			*!	*
b. e-stald	*!	*		*
c. ☞ ste-stald				**
d. te-stald	*!			*

- Here, the viable alternative to the PCR-violating C_1-copy candidate is the cluster-copying candidate (c).
  - When copying C_2 in addition to C_1 can avoid a poorly-cued repetition, a cluster in the reduplicant is tolerated.

- In all other cases, however, it is not:

(13) <u>Copying in TR roots in Gothic (no PCR violation)</u>: √*grēt-* → *<u>ge</u>-grōt* 'he wept'

/RED, grōt/	PCR	C/V
a. ☞ <u>ge</u>-grōt		*
b. <u>gre</u>-grōt		**!

❖ In §4, we will see that there are certain cases with C_1-copying which do not satisfy all the conditions of the current definition of the PCR. This will lead us to refine the definition.

3.3. TR-initial vs. ST-initial bases in Sanskrit

- The situation in Sanskrit is a bit more complicated.
  - There are two distinct non-default treatments, depending on the morpho-phonological origin of the base-initial cluster.
  - But the distribution of default vs. non-default treatment in both categories adheres to the principles of the PCR.

3.3.1. The behavior of cluster-initial roots in Sanskrit

- The division between TR and ST clusters for <u>cluster-initial roots</u> is illustrated in (14) & (15).
- We see again default C_1-copying to TR-initial roots:

(14) <u>C_1-copying perfects to TR-initial roots in Sanskrit (forms from Whitney 1885 [1988])</u>

Root	Perfect Tense	
bʰraj- 'shine'	<u>ba</u>-bʰrāj-a	not **<u>ra</u>-bʰrāj-a
drā- 'sleep'	<u>da</u>-drā-u	not **<u>ra</u>-drā-u
pracʰ- 'make'	<u>pa</u>-prācʰ-a	not **<u>ra</u>-prācʰ-a

- But in Sanskrit we see C_2-copying to ST-initial roots:

(15) <u>C_2-copying perfects to ST-initial roots in Sanskrit</u>

Root	Perfect Tense	
stʰā- 'stand'	<u>ta</u>-stʰā-u	not **<u>sa</u>-stʰā-u
stambʰ- 'prop'	<u>ta</u>-stambʰ-a	not **<u>sa</u>-stambʰ-a
sparç- 'touch'	<u>pa</u>-spr̥ç-ē	not **<u>sa</u>-spr̥ç-ē

- To derive the C_2-copying pattern for the ST-initial roots in Sanskrit, we again only need to permute the rankings proposed earlier for Greek and Gothic.
  - If ANCHOR-L-BR is the uniquely lowest-ranked relevant constraint, we predict C_2-copying as the repair for a PCR violation. This is shown in (16):

(16) C_2-copying in ST-initial roots in Sanskrit (PCR violation):

√*stamb*^{h}- → *ta-stamb*^{h}*-a* 'he has propped'

/RED, stamb^h, a/	ONSET	C/V	PCR	ANCHOR-L-BR
a. sa-stamb^h-a		*	*!	
b. a-stamb^h-a	*!	*		*
c. sta-stamb^h-a		**!		
d. ☞ ta-stamb^h-a		*		*

- TR-initial roots continue to copy C_1:

(17) C_1-copying in TR-initial roots in Sanskrit (no PCR violation):

√*prac*^{h}- → *pa-prāc*^{h}*-a* 'he has made'

/RED, prāc^h, a/	PCR	ANCHOR-L-BR
a. ☞ pa-prāc^h-a		
b. ra-prāc^h-a		*!

3.3.2. The behavior of cluster-initial zero-grade bases in Sanskrit

- The interaction between reduplication and zero-grade ablaut also induces PCR effects.
    - When C_1aC_2 roots are derived in the perfect active plural and the perfect middle, zero-grade ablaut would create a root allomorph of the shape $//C_1C_2//$.
- If the resulting C_1C_2-cluster is a TR cluster, C_1-copying is observed (18).

(18) C_1-copying perfects to -TR- zero-grade bases in Sanskrit[7]

Root		Perfect Tense	
b^h*ar-*	'bear'	ba-b^{h}r-ē	not **b^{h}ēr-ē
d^h*ar-*	'hold'	da-d^{h}r-ē	not **d^{h}ēr-ē
par-	'fill'	pa-pr-ur	not **pēr-ur

- If this new cluster would be an ST-cluster, as would be the case for the roots in (19), this allomorph would yield a PCR violation if accompanied by C_1-copying.
    - To avoid this, C_1-copying is blocked, just as in cluster-initial roots.
- But the non-default treatment is not C_2-copying; instead we see selection of a different allomorph, the "$C_1\bar{e}C_2$ pattern": $/C_1aC_2/ \rightarrow [C_1\bar{e}C_2$-].

[7] There are three *stop-liquid* roots which take the $C_1\bar{e}C_2$ pattern: *tērur* ← √*tar* 'pass', *p*^{h}*ēlire* ← √*p*^{h}*al* 'fruit', *p*^{h}*ēlur* ← √*p*^{h}*al* 'burst'. See Appendix B for discussion.

(19) $C_1\bar{e}C_2$ perfects to -ST- zero-grade bases in Sanskrit

Root		Perfect Tense	
sap-	'serve'	sēp-ur	not **sa-sp-ur, **pa-sp-ur
sad-	'sit'	sēd-ur	not **sa-sd-ur,[8] **da-sd-ur
çak-	'be able'	çēk-ur	not **ça-çk-ur, **ca-çk-ur
çap-	'curse'	çēp-ur	not **ça-çp-ur, **pa-çp-ur

- The unavailability to these roots of the C_2-copying pattern is explainable using Input-Reduplicant (IR) faithfulness (McCarthy & Prince 1995), specifically LINEARITY-IR.

(20) LINEARITY-IR:

For every pair of segments in the reduplicant *x'*, *y'*, such that *x'* precedes *y'*, assign a violation mark * if they have correspondents in the **underlying root** *x*, *y*, and *x* does not precede *y*.

- I assume that the reduplicant vowel corresponds to a segment in the underlying root,[9] such that LINEARITY violations are assigned as in (21):

(21) LINEARITY-IR violations: cluster-initial root vs. CaC root

	ZERO-GRADE CATEGORY (underlying vowel is deleted in output root)	LINEARITY-IR
Cluster-initial roots:	/RED, $s_1t^h{}_2\bar{a}_3$, ur/ → $\underline{t_2a_3}$-$s_1t^h{}_2$-ur	
CaC roots:	/RED, $s_1a_2p_3$, ur/ → **$\underline{p_3a_2}$-s_1p_3-ur	*

- LINEARITY-IR therefore blocks C_2-copying for these bases, and forces the use of a secondary repair strategy for the PCR, namely the $C_1\bar{e}C_2$ allomorph.
- I will treat this as morphological (following Sandell 2013), and allow the choice between reduplication and the $C_1\bar{e}C_2$ allomorph to be modulated by "USE X" constraints: USE REDUPLICATION » USE CēC.
  - These constraints are integrated into the phonological constraint ranking such that phonological constraints can force the use of the dispreferred morphological pattern.
- With these constraints in place, we can derive the four-part distribution shown in (22) with the tableaux in (23).

(22) Distribution of stem-formation patterns in the Sanskrit perfect

	ST cluster	TR cluster
Zero-grade base	$C_1\bar{e}C_2$	C_1-copying
Cluster-initial root	C_2-copying	C_1-copying

[8] This form, as well all three examples with /ç/, can be ruled out independently on phonotactic grounds. Nonetheless, there are many other cluster types which undergo the $C_1\bar{e}C_2$ pattern despite being phonotactically licit.

[9] See Steriade 1988 for arguments in favor of this approach.

(23) Reduplication in Sanskrit: TR vs. ST

			LINEARITY-IR	PCR	USE RED	C/V	ANCHOR-L-BR	USE CēC
Zero-grade	ST	/RED, s<a>p, ur/						
		a. sa-sp-ur		*!		*		*
		b. pa-sp-ur	*!			*	*	*
		c. ☞ sēp-ur			*			
	TR	/RED, p<a>r, ur/						
		a. ☞ pa-pr-ur				*		*
		b. ra-pr-ur	*!			*	*	*
		c. pēr-ur			*!			
Cluster-initial	ST	/RED, stamb^h, a/						
		a. sa-stamb^h-a		*!		*		*
		b. ☞ ta-stamb^h-a				*	*	*
		c. stēmb^h-a			*!	*		
	TR	/RED, prāc^h, a/						
		a. ☞ pa-prāc^h-a				*		*
		b. ra-prāc^h-a				*	*!	*
		c. prēc^h-a			*!	*		

3.3.3. The Sanskrit intensives

- While PCR explains the behavior of reduplication in the present and the perfect, we may run into problems when we consider the intensive.
- The intensive is formed by creating a heavy syllable reduplicative prefix (sometimes followed by a linking -*i*-). Onset clusters are simplified. (See Steriade 1988).

(24) Naturally-occurring intensives to ST roots

Root		Intensive	
sku	'tear'	cō-ṣkūyatē	(not **sō-ṣkūyatē)
skand	'leap'	kan-i-ṣkan, cani-ṣkadat	(not **san-i-ṣkand-)
spand	'quiver'	pan-i-ṣpad-	(not **san-i-ṣpad)
stan	'thunder'	taṅ-stanīhi	(not **saṅ-stanīhi)

  - There are also many grammarian-cited forms of the same sort.

- These intensives show the same behavior as the present and perfect: ST-clusters copy the T.
  - However, it is difficult to directly attribute this behavior to PCR.
- PCR as defined so far operates only over repetitions of the sort $C_\alpha VC_\alpha$.
  - When choosing between the candidates **S_1aR-i-S_1T_2(a)R- vs. T_2aR-i-S_1T_2(a)R-, no $C_\alpha VC_\alpha$ sequence is created, and thus PCR cannot penalize the C_1-copying form.

- ❖ If PCR did not apply to the intensives, yet they show the same behavior, we have a duplication problem in our grammar.
    - o But maybe we can save PCR.
- We could say that the *-i-* forms are secondarily derived from non-*i-* forms.
    - o The comparison is then between **$\underline{S_1aR\text{-}}S_1T_2(a)R$- vs. $\underline{T_2aR\text{-}}S_1T_2(a)R$-.
        - ▪ When R = /w,y/, it will contract with the reduplicant's /a/ to yield a long monophthong, as in *<u>cō</u>-ṣkūyatē* (= //<u>cau</u>-//).
        - ▪ When R = /n/, it surfaces as anusvara < ṅ >, as in *<u>taṅ</u>-stanīhi*, which we might interpret as nasalization of the vowel rather than a fully articulated nasal stop.
    - o In both cases, therefore, it is possible to think of the reduplicated consonant and the root-initial consonant as being separated only by a vowel, and thus within the domain of PCR.

4. Refining the PCR: the behavior of other cluster types

- We have now seen that the PCR can be used to explain the division between TR and ST clusters in reduplication in Ancient Greek, Gothic, and Sanskrit.
- But, each of these languages allows other types of root-initial clusters beside just TR and ST.
- In this section, we will evaluate which of these clusters pattern with TR and which pattern with ST in the respective languages, and we'll consider how the PCR can be defined to capture these distinctions.

4.1. Greek vs. Gothic

- While Ancient Greek and Gothic differ significantly in their cluster inventory, they differ minimally in their distribution of default vs. non-default treatment for various cluster types.
- The tables in (25) illustrate these distributions.
    - ❖ Attested initial clusters are marked with ✓; non-occurring initial clusters are marked with ✗ and dark grey.
    - ❖ White cells are those which display default C_1-copying; light grey cells are those occurring clusters with non-default treatment (non-copying in Greek, cluster-copying in Gothic).

(25) <u>Initial clusters and reduplicative behavior</u> (see Appendix A for complete data)

<u>Greek</u>

C_1 \ C_2	Stop (T)	Fricative (S)	Nasal (N)	Liquid (L)
Stop	✓	✓	✓	✓
Fricative	✓	(✓)[10]	✓	✓
Nasal	✗	✗	✓	✗

<u>Gothic</u>

C_1 \ C_2	Stop (T)	Fricative (S)	Nasal (N)	Liquid (L)
Stop	✗	✗	✗	✓
Fricative	✓	✗	✗	✓
Nasal	✗	✗	✗	✗

[10] Greek permits root-initial geminates in *-ss-* and (more frequently) *-rr-*. These roots show non-copying in the perfect: e.g. √*sseu* 'chase' → perfect *e-ssu-mai*. Whether this should be taken as a PCR effect remains a question.

- In Gothic, the only attested cluster with non-default treatment is indeed ST.
  - There are two other cluster types attested with reduplication,[11] both of which show default C_1-copying: *stop-liquid* (TL) and *fricative-liquid* (SL, S = {f,s})
- Greek has a more robust cluster inventory, but a very restricted distribution of C_1-copying.
  - TR (i.e. TL & TN) clusters copy C_1, whereas all other types show non-copying.

(26) Attested clusters and their behavior in Greek and Gothic

C_1 \ C_2	Stop	Nasal	Liquid
Stop	Greek: ✗	Greek: ✓	Greek: ✓
	Gothic: Ø	Gothic: Ø	Gothic: ✓
Fricative	Greek: ✗	Greek: ✗	Greek: ✗
	Gothic: ✗	Gothic: Ø	Gothic: ✓

✓ = C_1-copying
✗ = non-default treatment
Ø = unattested cluster type

❖ The most notable difference between Greek and Gothic in this respect is that Greek *does not* show default behavior for SL clusters (nor SN clusters) (S = /s/), whereas Gothic *does*.

- To see how this relates to the PCR, we must consider what cues are available in each cluster:

(27) Availability of cues to presence of C_1 in CC-clusters

C_1 \ C_2	**Stop** (T)	**Nasal** (N)	**Liquid** (L)
Stop (T)	TT · burst	TN · burst · intensity rise · CR transitions	TL · burst · intensity rise · CR transitions
Fricative (S)	ST · frication noise	SN · frication noise · intensity rise · CR transitions	SL · frication noise · intensity rise · CR transitions

❖ **Question**: Can we refine the PCR's "requisite cues" clause to capture the distinct distributions in the two languages?

- For Gothic, the "requisite cues" clause must pick out TL and SL to the exclusion of ST.
  - Two cues independently satisfy this condition: *intensity rise* and *CR transitions*.

(28) PCR [*for Gothic*]:

Assign a violation mark * to any $C_\alpha VC_\alpha$ sequence where the second consonant does not bear the requisite phonetic cues to its presence.

►**REQUISITE CUES**: *intensity rise* and/or *CR transitions*

[11] A few additional root shapes (*sm*-, *sn*-, *sw*-, etc.) exist in Gothic, but are not attested among reduplicating roots.

- For Greek, on the other hand, we must distinguish TR, which shows C_1-copy, from SR, which does not.
  - o We must also pick out TR to the exclusion of ST, TT, and NN (*nasal-nasal*), all of which are attested with non-copying in Greek.
- The major difference between TR and SR sequences with respect to cues is that TR sequences have *burst* but SR sequences have *frication noise*.
  - o While both of these are robust cues to the presence of a consonant (see Wright 2004), *frication noise* might be more apt to suffer from *repetition blindness*.
    - ❖ *Repetition blindness* might bias listeners towards disregarding the *frication*-less gap in a *fricative-vowel-fricative* sequence, and instead lead them to interpret the entirety of the frication noise as belonging to a single articulation.
    - ❖ Since the *burst* cue does not extend over a duration, it would be impossible for speakers to hear a second burst yet attribute it to the first.
    - ➢ Therefore, *burst* would seem to be a better cue in the repetition context than *frication noise*. This seems to be necessary for Greek.
- *Intensity rise* must also be a necessary cue, since *burst* alone would include TT and NN, neither of which permits C_1-copying.
  - ➢ Therefore, *burst* <u>and</u> *intensity rise* must both be requisite cues.[12]

(29) PCR [*for Greek*]:

Assign a violation mark * to any $C_\alpha VC_\alpha$ sequence where the second consonant does not bear the requisite phonetic cues to its presence.

►**REQUISITE CUES**: *burst* and *intensity rise*

➢ Greek and Gothic thus have different sets of "requisite cues" for the purpose of the PCR.
  - o We will see that positing different sets of requisite cues for different languages / language stages will derive the different scope of various effects.

4.2. Sanskrit

- The distribution of C_1- vs. C_2-copying in Sanskrit cluster-initial roots is shown in (30).

(30) <u>Attested clusters and reduplicative behavior in Sanskrit cluster-initial roots</u>

(see Appendix A for complete data)

C_2 / C_1	Stop	Affricate	Fricative	Nasal	Liquid	Glide
Stop	×	×	✓(?)	✓	✓	✓
Affricate	×	×	×	✓	✓	✓
Fricative	✓	✓	×	✓	✓	✓
Nasal	×	×	×	✓	✓	✓

[12] Yun (2014a) has identified these two cues (together termed "acoustic disjuncture") as being significant for the typology of epenthesis site. However, more recently Yun (p.c.) thinks that intensity rise on its own may be sufficient to explain the epenthesis typology, possibly with intensity rise defined such that release bursts create intensity rises.

- Sanskrit permits C_1-copying to all *consonant-sonorant* (CR) root-initial clusters.
  - This implies that *CR transitions* are sufficient for PCR satisfaction in Sanskrit.[13]

- Additionally, it also seems to permit C_1-copying to *stop-fricative* (TS) clusters (*ps*, *ts*, *kṣ*).
  - This is surprising, given that, before a fricative, a stop will bear none of the cues available in TR clusters, i.e. *burst*, *intensity rise*, or *CR transitions*.
- But the data here is minimal. There are only two relevant examples:
  - (i) √*tsar*- 'approach stealthily' → perfect *tatsāra*
    - This form is attested only in the Rig-Veda (Whitney 1885 [1988]: 68).
    - This period may contain certain remnants of an earlier stage which was more permissive than later stages wrt PCR (see discussion in Appendix B).
  - (ii) √*psā*- 'devour'[14] → perfect *papsāu*
    - This form is only cited by grammarians, rather than occurring in actual texts (Ibid.: 104).
  - Roots in #*kṣ* copy C_1, but they are freed from PCR effects by the independent process of reduplicant velar palatalization; √*kṣad*- 'divide' → perfect *cakṣadē*.
    - Since /k/ copies as [c], repetition is obviated and PCR is satisfied, whether or not it would normally be violated in TS clusters.
- Therefore, if we disregard *papsāu* as artificial and identify *tatsāra* as an unproductive archaism, we do not have to say that synchronic Sanskrit (at any historical stage) displayed C_1-copying to TS-initial roots.

- The division between C_1-copying and C_2-copying among cluster-initial roots can thus be characterized as the presence or absence of *CR transitions*:

(31) PCR [*for Sanskrit*]:

Assign a violation mark * to any $C_\alpha VC_\alpha$ sequence where the second consonant does not bear the requisite phonetic cues to its presence.

►**REQUISITE CUES**: *CR transitions*

❖ The facts regarding zero-grade bases are harder to pin down, but follow the same general shape:
  - CR clusters tend to show C_1-copying,
  - Other clusters tend to show non-default treatment (i.e. the $C_1\bar{e}C_2$ pattern).
- There are though a number of cases on the borderline which contradict this statement.
  - Phonotactics and diachrony interfere significantly, such that it is difficult to tease apart what is directly applicable to the PCR at any given stage.

❖ See Appendix B for a tentative analysis of the developments.

---

[13] This holds as long as we assume that (non-homorganic) NN sequences have *CR transitions*. There is a single data point for NN roots: √*mnā* 'note' → *mamnāu*; this form is not found in naturally-occurring texts, but rather only cited by grammarians.

[14] This is a secondary root built from √*b^has* 'devour' + extension -*ā*- (Whitney 1885 [1988]: 104).

5. Other PCR effects in IE Reduplication

5.1. The $C_1\bar{e}C_2$ pattern in Germanic (and elsewhere…)

- The $C_1\bar{e}C_2$ pattern of Sanskrit seems to have a direct counterpart in Germanic.
- The preterite plurals of CeT / CeS roots (Strong Class V) unexpectedly show a long vowel [ē] in the root.

❖ From a historical/derivational perspective, we might able to view this as reduplication followed by deletion of root-C_1 + compensatory lengthening.

(32) Gothic Class V preterite plurals (forms from Lambdin 2006:51)

	Infinitive	Preterite Plural (1PL.)	
'to give'	*giban* [gib-an]	*gebum* [gēb-um]	(as if from *ge-gb-um)
'to say'	*qiþan* [k^wiθ-an]	*qeþum* [k^wēθ-um]	(as if from *k^we-k^wθ-um)

(33) Deriving $C_1\bar{e}C_2$ in *Pre-Germanic*[15]

	/RED, C_1eC_2, um/
Reduplicate: copy CV	C_1e-C_1eC_2-um
Zero-grade: delete root vowel[16]	C_1e-C_1C_2-um
Deletion + compensatory lengthening: $eC_\alpha \rightarrow \bar{e}$ / C_α_C	[$C_1\bar{e}C_2$-um]

- The deletion + CL rule can be viewed as a repair for a PCR violation.
    - In the $Ce^T/_S$ roots of Class V, zero-grade ablaut places a consonant-repetition before another consonant (specifically an obstruent), leaving the repetition poorly-cued.
    - The [ē] reflex is not seen in roots of the shape CeRC (Strong Class I-III), because they had a sonorant which could vocalize and provide a well-cued repetition:
        - √CeRC → preterite plural Ce-CR̥C-

❖ It is unclear to me whether the Class IV (CReC roots) preterite plurals participated in this pattern at this stage.
- In Gothic, Class IV works just like Class V:
    - *bair-an* [bɛr-an] 'to bear' → pret. pl. *ber-um* [bērum]
- Was this the case for Pre-Germanic, or did these roots show C_1-copying:
    - Pre-Germanic: *bērum* or *bebrum* ?
- It is possible that Class IV was secondarily attracted to this pattern after a larger change in the morphological system (cf. Sandell & Zukoff 2014).

➢ Due to this uncertainty, I will not speculate further on the details of the PCR at this stage.

❖ For a preliminary OT analysis of this pattern, see Appendix C.

[15] This is not how the pattern is encoded in the synchronic grammar of Gothic (cf. Sandell & Zukoff 2014), so it must be attributed to an earlier stage. This is necessary anyway, since the reflexes of this pattern are seen across the Germanic languages, even those which lack reduplication in the reflexes of Class VII verbs.
[16] This deletion was likely originally conditioned by accent, which surfaced on the suffix in the plural.

- A similar grammar could be used to generate the $C_1\bar{e}C_2$ pattern in early Sanskrit (i.e. generate it phonologically rather than through allomorph selection).
- Equivalent forms also exist in Old Irish:

(34) Old Irish ā-preterites (Thurneysen 1966: 429; cf. Niepokuj 1997: 151-152)

Present stem	Preterite stem	
tech- [tʲex-] 'flees'	*tách-* [tāx-]	(as if from *ta-tx-)
reth- [rʲeθ-] 'serve'	*ráth-* [rāθ-]	(as if from *ra-rθ-)
fig- [fʲigʲ-] 'weaves'	*fáig-* [fāgʲ-]	(as if from *fa-fgʲ-)

- The similarity of repair between these languages suggests that the pattern could be projected even farther back to Proto-Indo-European.
  - Given the qualities of the vowels in the different languages, it is unlikely that we can reconstruct the *forms* themselves to PIE.
  - But, if the process remained productive, independent changes in the phonologies of the individual languages could account for the differences in vowel quality.
- ❖ Sandell (2014) argues that some of the "Narten" roots/formations in PIE, i.e. present stems with unexpected long-vowels, could be examples of exactly this pattern; but see Jasanoff (2012) for arguments against deriving Narten forms from reduplication.

5.2. Attic Reduplication

- In Ancient Greek, certain vowel-initial roots show VC-copying rather than the more productive vowel-lengthening pattern.
  - √*ager-* 'gather' → perf ἀγήγερμαι [agēgermai] (vs. √*ag-* 'lead' → perf ἦχα [ēkʰa])
  - ❖ This VC-copying pattern is referred to as Attic Reduplication.
- Attic Reduplication can be reconstructed as a non-default reduplication strategy imposed on laryngeal-initial roots in Pre-Greek (Zukoff 2014, in prep).
  - *√*h₂ger-* → perfect *<u>*h₂əg-e*</u>-*h₂ger-*
- The motivation for the pattern is a desire to avoid repeated identical laryngeals in pre-consonantal position: *√*h₂ger-* → perfect ˣ<u>*h₂-e*</u>-*h₂ger-*.
  - ➢ This is a PCR effect.

- The evidence suggests that all non-laryngeal cluster-initial roots showed C_1-copying in Pre-Greek.
  - The non-productive reduplicated presents (and the perfects associated with them) show C_1-copying even if they have a non-TR root allomorph in the present:

(35) Present reduplication to non-TR clusters

Root	Present		Perfect	
mnē- 'remind'	μιμνήσκω	[<u>m</u>-i-mnē-sk-ō]	μέμνημαι	[<u>m</u>-e-mnē-mai]
stē- 'stand'	ἵστημι	[<u>h</u>-i-stē-mi] (< **sistēmi*)	ἕστηκα	[<u>h</u>-e-stē-ka] (< **sestēka*)
pet- 'fall'	πίπτω	[<u>p</u>-i-pt-ō]	πέπτωκα	[<u>p</u>-e-pt-ōka]

- Also, there is variation in the perfect of one particular core lexical item:
  √*kta*- 'acquire' →
    - Expected: non-copying pattern ἔκτημαι [e-ktε̄-mai]
    - Unexpected: C_1-copying pattern κέκτημαι [k-e-ktε̄-mai]
    - ➢ The latter should be taken as an archaism.

- Taking these facts to reflect the earlier stage of the language, we have evidence for a dichotomy between laryngeal-initial clusters, on the one hand, and (all) other clusters, on the other hand.
  - $\#H_1C_2V\text{-} \rightarrow \underline{H_1əC_2\text{-}e}\text{-}H_1C_2V\text{-}$
  - $\#C_1C_2V\text{-} \rightarrow \underline{C_1}\text{-}\underline{e}\text{-}C_1C_2V\text{-}$, when C_1 = *stop, sibilant, or nasal* [17]
- Under the assumption that the laryngeals were non-strident fricatives – and therefore had *low-intensity* frication noise, this distribution can lead us to a cogent definition of the PCR for this stage:

(36) PCR [*for Pre-Greek*]:

Assign a violation mark * to any $C_\alpha VC_\alpha$ sequence where the second consonant does not bear the requisite phonetic cues to its presence.

►**REQUISITE CUES**: (i) *burst* or (ii) *high-intensity frication noise*

- *Burst* licenses C_1-copying for stops and nasals.
- *High-intensity frication noise* licenses C_1-copying for *s*-initial clusters, to the exclusion of the low-intensity frication of *laryngeal*-initial clusters.

5.3. Latin *ST*- infixing reduplication

- In §3, we saw that ST-initial roots undergo non-default treatment in Greek, Gothic, and Sanskrit. The same can be said for Latin.[18]
- The behavior of the ST roots can be characterized as *infixing* reduplication:

(37) Latin infixing perfects (forms from Weiss, 2009: 410)

Root		Perfect	
√*st* 'stand'	→	*s-te-t-ī*,	not ***se-st-ī* (but present *si-st-ō*)
√*spond* 'promise'	→	*s-po-pond-ī*,	not ***so-spond-ī*
√*scid* 'cut'	→	*s-ci-cid-ī*,	not ***si-scid-ī*

- This can be described by saying that *left-alignment* of the reduplicant and *contiguity* of the root are less important than avoiding a poorly-cued repetition.
  - The alignment approach also explains why this infixation is minimally displaced from the left edge.

[17] The behavior of *approximant-consonant* clusters which may have arisen through zero-grade ablaut needs to be considered further.

[18] We can only contrast ST roots to CV roots, as TR roots are not attested with reduplication. This may well be a non-accidental gap, but I will not try to account for it here.

(38) Infixing reduplication in Latin (PCR violation) (bolded string = the “base” of reduplication)

/ RED, scid, ī /	ANCHOR-L-BR	PCR	C/V	CONTIG-ROOT	ALIGN-RED-L
a. si-**scid-ī**		*!	*		
b. sci-**scid-ī**			**!		
c. ci-**scid-ī**	*!		*		
d. ☞ s-ci-**cid-ī**			*	*	*
e. sc-id-**id-ī**			*	*	**!

- Since we do not have direct evidence from reduplication of other cluster types, we cannot determine the exact nature of the Latin PCR.
- Regardless, we will see that PCR effects involving *s* in Latin are also attested outside of reduplication (§6.2).

6. PCR effects outside of reduplication[19]

6.1. An exception to Bartholomae’s Law

- In Sanskrit, voiced aspiration (probably murmur or breathy voice) shows a complex distribution.
    - Breathy voice appears on stops [b^{ɦ}, d^{ɦ}, ḍ^{ɦ}, g^{ɦ}] (= D^{ɦ}), but is only licensed in pre-sonorant position.
        - All laryngeal contrasts – voicing, aspiration, breathy voice (≈ voicing + aspiration) – are neutralized in pre-obstruent and word-final position.
    - When an underlying D^{ɦ} would surface in a position where the laryngeal contrast is not licensed, the breathy voicing can migrate to a nearby stop (MAX[^{ɦ}]), subject to certain restrictions.
        - If there is no stop on which it can land, the breathy voice is lost.

- There are two places the breathy voice can migrate to:[20]

    (i) If the preceding consonant (or a member of the preceding consonant cluster) is a *plain voiced stop*, the breathy voice can surface on that stop.
    - ❖ √bud^{ɦ} ‘know’ : root noun LOC.PL. /**bud^{ɦ}**-su/ → [**b^{ɦ}ut**-su]
    - ➢ This is often referred to as Aspiration Throw Back (ATB).

    (ii) If the immediately following consonant is a *plain voiceless (or plain voiced) stop*, the breathy voice can surface on that stop. (That stop also becomes voiced.)
    - ❖ √rud^{ɦ} ‘obstruct’ : nasal-infix present 3.SG. /ru-na-**d^{ɦ}-t**i/ → [ru-na-**d-d^{ɦ}**i]
    - ➢ This process is known as Bartholomae’s Law (BL).

[19] Thank you to Donca Steriade for bringing these examples to my attention.
[20] Here, I adopt the position that the IE “diaspirate” roots are represented synchronically in Sanskrit as /DVD^{ɦ}/ not /D^{ɦ}VD^{ɦ}/.

- When both ATB and BL are in principle available, BL is preferred:
  - √budʱ 'know' : past participle /**budʱ-t**a-/ → [bu**d-dʱ**a-] (BL), not **[**bʱ**u**t**-ta-] (ATB)

- There is at least one such example where the usually dispreferred ATB option surfaces:

(39) Reduplicated present of √dʱā 'place'

	ACTIVE			MIDDLE		
	SING	DUAL	PL	SING	DUAL	PL
1ST	dá-dʱā-mi	da-dʱ-vás	da-dʱ-más	da-dʱ-é	dá-dʱ-vahe	dá-dʱ-mahe
2ND	dá-dʱā-si	dʱa-t-tʰás	dʱa-t-tʰá	dʱa-t-sé	da-dʱ-ā́te	dʱá-d-dʱve
3RD	dá-dʱā-ti	**dʱa-t-tás**	dá-dʱ-ati	**dʱa-t-té**	da-dʱ-ā́te	dá-dʱ-ate

❖ White cells are those in which ATB applies. The bolded cells – ACT.3DL and MID.3SG – have the conditions to support either BL or ATB.
  - ➢ They show ATB rather than BL!
- This is a PCR effect. Consider what the BL form would have been:
  - ❖ MID.3SG //da-dʱ-te// → BL → **[**da-d**-dʱe]
- BL would create a sequence of two identical stops in a position where the second of the repeated stops does not have CR transitions.
  - This would be a violation of the Sanskrit PCR.

➢ Avoidance of the PCR violation thus results in choosing ATB instead of BL.

- Using cover constraints for what is determining the choice between ATB and BL (NO ATB » NO BL), we can see how PCR causes misapplication:

(40) PCR blocks BL:

/da-dʱ-te/	Cʱ / _R	MAX[ʱ]	PCR	NO ATB	NO BL
a. dadʱté	*!				
b. datté		*!			
c. daddʱé			*!		*
d. ☞ dʱatté				*	

6.2. The Latin –*is*… suffixes

➢ Cser (to appear (2015): §2.3.2.2) documents allomorphy involving the perfective suffixes -*(i)sse*, -*(i)stī*, and -*(i)stis*.
  - In the normal case, when attached to consonant-final stems, they surface with the -*i*- variant: *nōv-isse*, *tetig-isse*, etc.
  - When attached to vowel-final stems (other than *u*-final stems), the vowel-less variant surfaces: *complē-sse*, *abī-sse*, etc.
  - For stems ending in /s/, we see free variation according to the pattern in (41):

(41) <u>Variability in *s*-final stems</u>

a. *access-istis ~ acces-tis* c. *admis-isse ~ admis-se,*
b. *divis-isse ~ divis-se* d. *dere*[*ks*]*-isti ~ dere*[ks]*-ti*

- ➢ PCR can motivate phonological deletion of the entire suffix-initial /is/ sequence, without resorting to choosing the wrong allomorph.
- A variable ranking between PCR and MAX-AFX (*don't delete an affix segment*) will generate the variable realization of these suffixes following *s*-final stems.
  - When MAX-AFX outranks PCR (41A), the suffix will be realized faithfully.
  - When PCR outranks MAX-AFX (41B), there is deletion.
    - Two other requirements make deletion of just a single segment suboptimal:

(i) Geminates must be inter-sonorant *GEM//OBS
(ii) Contiguity between affix segments is maintained CONTIGUITY-AFX

(42) <u>Variable ranking between PCR and MAX(AFX) generates variable realization</u>

A. When MAX-AFX » PCR: faithful realization of suffix

/dereks, -isti/	*GEM//OBS	CONTIGUITY-AFX	MAX-AFX	PCR
a. ☞ dereks-isti				* (sist)
b. dereks-sti	*!		* (i)	
c. dereks-iti		*! (i↔t)	* (s)	
d. dereks-ti			*!* (is)	

B. When PCR » MAX-AFX: deletion of suffix-initial /is/

/dereks, -isti/	*GEM//OBS	CONTIGUITY-AFX	PCR	MAX-AFX
a. dereks-isti			*! (sist)	
b. dereks-sti	*!			* (i)
c. dereks-iti		*! (i↔t)		* (s)
d. ☞ dereks-ti				** (is)

7. Conclusion

- In the reduplicative systems of the Indo-European languages, there are a number of cases in which default C_1-copying is blocked.
- In this paper, I have argued that these cases can be unified as repair strategies for the Poorly-Cued Repetition Principle (PCR):

(43) THE POORLY-CUED REPETITION PRINCIPLE (PCR):

> A CVC sequence containing identical consonants ($C_\alpha VC_\alpha$) is dispreferred, due to repetition blindness; it is especially dispreferred if one or both of the consonants does not bear phonetic cues which are important for the perception of its presence (in contrast to zero) in the speech signal.

- The proposal centers around the logical union of *repetition avoidance* (cf. Walter 2007) and the availability and robustness of phonetic cues (cf. Wright 2004).

- The PCR can induce avoidance of C_1-copying in reduplication, contrary to the normal pattern of the Indo-European languages.

❖ In the systems in (44), the details of the PCR constraint can be identified with some degree of confidence.

(44) Reduplicative PCR effects and details

Language	Non-default treatment induced by PCR	Requisite Cues for PCR[21]
Ancient Greek	Non-copying	*burst + intensity rise*
Gothic	Cluster-copying	*intensity rise* and/or *CR transitions*
Sanskrit	C_2-copying	*CR transitions*
Pre-Greek	Attic Reduplication	(i) *burst* or (ii) *high-intensity frication noise*

❖ The systems in (45) also display PCR effects in reduplication.
  - But lack of relevant comparisons and/or diachronic complexity prevents conclusive definition of the PCR constraint.

(45) Additional reduplicative PCR effects

Language	Non-default treatment induced by PCR
Sanskrit (multiple stages?)	$C_1\bar{e}C_2$ allomorphy to zero-grade bases
Pre-Germanic	$C_1\bar{e}C_2$ preterite plurals to Class V (and maybe Class IV?) roots
Latin	Infixing reduplication to ST roots

- The PCR is also responsible for minor irregularities outside of reduplication:
  - An exception to Bartholomae's Law in Sanskrit
  - Phonological allomorphy in Latin suffixation

❖ Further questions:

- Are there other effects in the Indo-European languages, either relating to reduplication or more generally, which can be attributed to PCR?
- Can PCR help explain cluster-dependent reduplication asymmetries outside of IE?
  - Likely yes: Klamath works exactly like Gothic (Steriade 1988, Fleischhacker 2005).
- Can we make strong predictions about the types of cluster asymmetries based on the relative strength of different cues?
  - Repetition with non-sibilant fricatives should imply repetition with sibilant fricatives.
    - *High intensity* frication noise is a more robust cue than *low intensity*.
    - ➢ In Pre-Greek, /s/ supports C_1-copy but /h_x/ does not.
  - If CR transitions are not all created equal, i.e. CV > CL > CN, we predict languages which permit repetition to CL but not CN.
    - ➢ (Classical?) Sanskrit $C_1\bar{e}C_2$ allomorphy may point in this direction…

[21] In each case, *CV transitions* license C_1-copying as well.

8. References

Cser, András. to appear (2015). The Nature of Phonological Conditioning in Latin Inflectional Allomorphy. *Acta Linguistica Hungarica* 62.

Fleischhacker, Heidi Anne. 2005. Similarity in Phonology: Evidence from Reduplication and Loan Adaptation. PhD Dissertation, University of California, Los Angeles.

Jasanoff, Jay H. 2012. Long-vowel Preterites in Indo-European. In *The Indo-European verb: proceedings of the Conference of the Society for Indo-European Studies, Los Angeles, 13-15 September 2010*, edited by H. Craig Melchert, 127–35. Wiesbaden: Reichert Verlag.

Kanwisher, Nancy. 1987. Repetition Blindness: Type Recognition without Token Individuation. *Cognition* 27: 117–43.

van de Laar, Henri M. F. M. 2000. *Description of the Greek Individual Verbal Systems*. Amsterdam: Rodopi.

Lambdin, Thomas O. 2006. *An Introduction to the Gothic Language*. Eugene, Oregon: Wipf & Stock Publishers.

McCarthy, John J., & Alan Prince. 1995. Faithfulness and Reduplicative Identity. In *University of Massachusetts Occasional Papers in Linguistics 18: Papers in Optimality Theory*, edited by Jill Beckman, Suzanne Urbanczyk, and Laura Walsh Dickey, 249–384. Amherst, MA: GLSA, University of Massachusetts. http://works.bepress.com/john_j_mccarthy/44.

Niepokuj, Mary. 1997. *The Development of Verbal Reduplication in Indo-European*. Washington, D.C.: Institute for the Study of Man.

Sandell, Ryan. 2013. Vedic Perfect Weak Stems of the Form C_1eC_2. Paper presented at the *25th Annual UCLA Indo-European Conference*, Los Angeles, October 26. Available at: https://ucla.academia.edu/RyanSandell.

Sandell, Ryan. 2014. The Phonological Origins of Indo-European Long-Vowel (“Narten”) Presents. Paper presented at the *26th Annual UCLA Indo-European Conference*, Los Angeles, October 25. Available at: https://ucla.academia.edu/RyanSandell.

Sandell, Ryan, & Sam Zukoff. 2014. A New Approach to the Origin of Germanic Strong Preterites. Poster presented at *NELS 45*, November 1, MIT. Available at https://mit.academia.edu/SamZukoff.

Steriade, Donca. 1988. Reduplication and Syllable Transfer in Sanskrit and Elsewhere. *Phonology Yearbook* 5: 73–155.

Steriade, Donca. 1997. Phonetics in Phonology: The Case of Laryngeal Neutralization. Ms., UCLA.

Struijke, Carolina Maria. 2000. Reduplication, Feature Displacement, and Existential Faithfulness. PhD Dissertation, University of Maryland.

Thurneysen, Rudolph. 1966. *A Grammar of Old Irish*. Translated by D. Binchy and O. Bergin. Dublin: Dublin Institue for Advanced Studies.

Walter, Mary Ann. 2007. Repetition Avoidance in Human Language. PhD Dissertation, MIT.

Weiss, Michael. 2009. *Outline of the Historical and Comparative Grammar of Latin*. Ann Arbor: Beech Stave Press.

Whitney, William Dwight. 1885 [1988]. *The Roots, Verb-Forms, and Primary Derivatives of the Sanskrit Language: A Supplement to his Sanskrit Grammar*. Reprinted 1945 as American Oriental Series 30. New Haven: American Oriental Society.

Wright, Richard. 2004. A Review of Perceptual Cues and Cue Robustness. In *Phonetically Based Phonology*, edited by Bruce Hayes, Robert Kirchner, and Donca Steriade, 34–57. Cambridge: Cambridge University Press.

Yun, Suyeon. 2014a. The Role of Acoustic Cues in Nonnative Cluster Repairs. In *Proceedings of the 31st West Coast Conference on Formal Linguistics*, edited by Robert E. Santana-LaBarge, 514–23. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #3055.

Yun, Suyeon. 2014b. The Typology of Compensatory Lengthening: A Phonetically-Based Optimality-Theoretic Approach. In *Proceedings from the Annual Meeting of the Chicago Linguistic Society*, 46:341–55. Chicago Linguistic Society.

Zukoff, Sam. 2014. On the Origins of Attic Reduplication. In *Proceedings of the 25th UCLA Indo-European Conference*. Bremen: Hempen.

Zukoff, Sam. in prep. The Reduplicative System of Ancient Greek and a New Analysis of Attic Reduplication. Ms.

9. Appendix A: The data

9.1. Ancient Greek (see, e.g., Van de Laar 2000)

(46) C$_1$-copying root-shapes in Ancient Greek

❖ (a) Roots with initial singleton consonants

	Root		Present Tense	Perfect Tense
(i)	pemp-	'send'	pemp-ō	p-e-pemp-tai
	dō-	'give'	d-i-dō-mi	d-e-dō-ka
(ii)	sōs-	'save'	sōzd-ō	s-e-sōs-mai
	lu-	'loosen'	lu-ō	l-e-lu-ka

❖ (b) Roots with initial *stop* + *sonorant* (TR) clusters

Root		Present Tense	Perfect Tense
kri-	'decide'	krī-n-ō	k-e-kri-mai
tla-	'suffer, dare'	tla-ō	t-e-tlē-mai
pneu-	'breathe'	pne-ō	p-e-pnū-mai

(47) Non-copying root-shapes in Ancient Greek

❖ (a) Roots with initial *stop* + *obstruent* clusters

	Root		Perfect Tense	
(i)	kten-	'kill'	e-kton-a	not **k-e-kton-a
	p^{h}t^{h}i-	'decay'	e-p^{h}t^{h}i-ka	not **p-e-p^{h}t^{h}i-ka
(ii)	pseud-	'lie'	e-pseus-mai	not **p-e-pseus-mai

❖ (b) Roots with initial *s* + *consonant* clusters

	Root		Perfect Tense	
(i)	stel-	'prepare'	e-stal-ka	not **s-e-stal-ka
	strateu-	'wage war'	e-strateu-mai	not **s-e-strateu-mai
(ii)	smɛ̄k^{h}-	'wipe'	e-smɛ̄g-menos	not **s-e-smɛ̄g-menos

❖ (c) Roots with initial geminates

Root		Perfect Tense	
rreu-	'flow'	e-rru-ɛ̄ka	not **r-e-rru-ɛ̄ka
sseu-	'hasten'	e-ssu-mai	not **s-e-ssu-mai

9.2. Gothic (see Lambdin 2006: 115)

(48) C_1-copying root-shapes in Gothic

❖ (a) Roots with initial singleton consonants

	Infinitive	Preterite (1/3SG.)
'to fold'	falθ-an	fɛ-falθ
'to tend'	hald-an	hɛ-hald
'to boast'	h^{w}ōp-an	h^{w}ɛ-h^{w}ōp
'to touch'	tēk-an	tɛ-tōk
'to play'	laik-an	lɛ-laik

❖ (b) Roots with initial *stop* + *liquid* clusters

	Infinitive	Preterite (1/3SG.)
'to weep'	grēt-an	gɛ-grōt

❖ (c) Roots with initial *fricative* + *liquid* clusters

	Infinitive	Preterite (1/3SG.)
'to sleep'	slēp-an	sɛ-slēp (also sɛ-zlēp)
'to bewail'	flōk-an	fɛ-flōk
'to tempt'	frais-an	fɛ-frais

(49) Cluster-copying root-shapes in Gothic

❖ Roots with *fricative* + *stop* clusters

	Infinitive	Preterite (1/3SG.)	
'to possess'	stald-an	stɛ-stald	not **sɛ-stald
'to divide'	skaið-an	skɛ-skaiθ	not **sɛ-skaiθ

9.3. Sanskrit cluster-initial roots (see Whitney 1885 [1988])

- White cells are those which show C_1-copying; light grey cells are those with C_2-copying; dark grey cells are unattested clusters.
- Forms marked with brackets [ ] are those Whitney reports as being cited only in grammatical texts rather than in naturally-occurring texts.

- *Italicized* forms are those in which the reduplicated consonant is a palatal affricate corresponding to a root-initial velar stop or *h*.
    - These would be expected to escape PCR violations by virtue of their (significant) non-identity.

C_2 / C_1	Stop	Affricate	Sibilant	Nasal	Liquid	*w*	*y*
Stop			tatsāra [papsāu] *cakʂamē* *cakʂadē*	dad^{h}māu[22]	bab^{h}rāja dadrāu paprāc^{h}a tatrē pupluvē *jagrab^{h}a* *cakranda*	didvēʂa tatvarē [*cakvāt^{h}a*]	dad^{h}yāu cak^{h}yāu tatyāja didyota
h				[*juhnuvē*]	*jihrāya*	[*jahvāla*]	
Affricate				jajɲāu	[jijrāya]	jajvāla [jajvāra]	cucyuvē jijyāu
s	tast^{h}āu tastamb^{h}a tastāra *caskanda* *cask^{h}āla* paspr̥çē paspaçē			sasmāra sismāya sasnur	susrāva sasransur	sasvadē sasvajē sasvanur	sasyandē
ʂ, ç	tiʂʈ^{h}ēva tuʂʈ^{h}āva	cuççota		[çaçnāt^{h}a]	çaçrāma çaçrat^{h}ē çaçlāg^{h}irē	çaçvāsa [çiçvāya]	siʂyanda suʂvāpa [çaçyē]
Nasal				[mamnāu]	mumloca mamlāu		mimyakʂa
w					vavrāja		vivyāca vivyād^{h}a

10. Appendix B: Sanskrit zero-grade bases

- Because there are few (relevant) co-occurrence restrictions between root-C_1 and root-C_2 in CaC roots, zero-grade ablaut can in theory bring just about any two consonants into contact.
    - ➢ Many of these sequences are not phonotactically licit.
    - Copying either consonant would not change the illegality of the zero-grade sequence.

[22] It is unclear whether this should be treated as a cluster-initial form or a zero-grade form.

  - The $C_1\bar{e}C_2$ pattern circumvents this problem altogether by replacing the would-be zero-grade allomorph with one that has a vowel between the consonants.
  - Therefore, any would-be cluster type which is not phonotactically legal that shows the $C_1\bar{e}C_2$ pattern cannot be used as direct evidence for the PCR.
- There may also be diachronic interference, as the $C_1\bar{e}C_2$ pattern seems to expand its scope of applicability over time in a non-strictly-phonological way (cf. Sandell 2013).
  - Among roots which attest both a C_1-copying form and a $C_1\bar{e}C_2$ form, the C_1-copying form is almost always older, and usually ceases to be attested in the later periods.
  - This implies that there has been grammatical change between the earliest period and the later periods.
  - Furthermore, the $C_1\bar{e}C_2$ pattern even eventually spreads beyond CaC roots to some CRaC / CaRC roots.[23]
  - ➢ Therefore, it is in principle possible that, in the later language, some factor other than the PCR could be conditioning the selection of the $C_1\bar{e}C_2$ allomorph.
- Although much is left to interpretation, the data which will be presented below suggests that Sanskrit has undergone a (gradual) change in the strictness of the PCR and the scope of the repair (i.e. the $C_1\bar{e}C_2$ pattern) for zero-grade bases.
  - The earliest (possibly pre-historic) stage is one in which the PCR was very permissive (or inactive) and repair was dispreferred.
  - The latest stage, on the other hand, is one in which the PCR was more strict and the repair was not dispreferred at all (rather it may have come to be preferable to actual zero-grade ablaut).

10.1. The data

- ❖ The following chart shows all the attested perfect forms to CaC roots which have either (i) C_1-copying with zero-grade of the root[24] (white cells) or (ii) $C_1\bar{e}C_2$ allomorphy (grey cells).[25] (Dark grey cells are unattested clusters.)
- ❖ *Italicized* forms are those in which the reduplicated consonant is non-identical to root-C_1, either due to velar palatalization in the reduplicant or place assimilation of root-C_1 to root-C_2.
- ❖ Forms marked with brackets [ ] are those Whitney reports as being cited only in grammatical texts rather than in naturally-occurring texts; he does not report which grammarian(s) cites such forms, and thus I do not know their chronology. They should be taken with a grain of salt anyway.
- ❖ Forms marked in parentheses ( ) are presents or other derivatives which appear to have (or clearly do have) reduplication.

[23] For example: √*tras* 'be terrified' → perf active singular *tatrāsa*, but perf middle plural *trēsur* (not ***tatr̥sur*); √*b^h^ram* 'wander' → perf. active singular *bab^h^rāma*, but perf. middle plural *b^h^rēmatur* (not ***bab^h^r̥matur*)

[24] There are other CaC roots with C_1-copying in zero-grade categories but with an unexpected full-grade of the root. This seems to be another avoidance strategy for bad clusters (whether phonotactic or PCR), but I have not yet examined these systematically.

[25] There are additional examples of *stop-liquid* roots; all other cells are virtually exhaustive to the best of my knowledge.

- ❖ Each $C_1\bar{e}C_2$ form is accompanied by a ✓ (indicating that the cluster resulting from zero-grade would be phonotactically **legal**) or a ✗ (indicating that the cluster would be phonotactically **illegal**, assuming no assimilations took place). Those in parentheses are those which I am unsure of in this regard.
- ❖ The solid vertical line separates clusters where C_2 is an obstruent (left) from those where C_2 is a sonorant (right).

(50) Treatment of zero-grade clusters in reduplication (data from Whitney 1885 [1988])

C_1 \ C_2	Stop	Affricate	*h*	Sibilant	Nasal	Liquid	Glide
Stop	paptur pētur ✓ pēdur ✗ dēb^{h}ur ✓ tēpē ✓	pēcur ✗	dēhē ✗	*jakşur* (*bapsati*)	dad^{h}mirē tatnē papnē *cak^{h}nur* *jagmur* tēnē ✓ [pēṇē] ✗ [p^{h}ēṇur] ✗	bab^{h}rē dad^{h}rē dadrē paprur *cakrē* tērur ✓ p^{h}ēlirē ✗ [p^{h}ēlur] ✗	bib^{h}yur *cikyur* *jigyē* *jig^{h}yur*
Sibilant	sēdur ✗ sēpur ✓ çēkur ✗ çēpur ✗	*saçcur* sēcirē ✗ sējur ✗	sēhur ✗		(sasni) (siṣṇu) [sēnē] ✓ çēmur (✓)	sasrē çaçrē [çēlē] (✗)	*suṣvāṇa* çiçyē
Nasal	? [26] mēthur ✗ nēdur ✓ [nēbhē] ✗		[nēhē] ✗	nēçur (✓)	mamnāt^{h}ē mēnē ✓ nēmē ✓	mamrur	ninyē mimyur
Liquid	lēb^{h}ē ✗ lēpur (✓) rēb^{h}ē ✓ [rēdur] ✓ [rēṭur] ✓	rējur ✓	rēhur (✗)	lēṣur (✓) rēsur ✓	rēmē ✓		lilyē
Glide	yētē ✗	yējē(?) ✗			vavnē(?) vēmur (✗) yēmur ✗	vavrē	vivyē

[26] There may be some members of this class, but each possible example is ambiguous at best. To the root which Whitney lists as "√mat^h, mant^h- 'shake' " (p. 117), we could imagine the forms with medial [n] as being originally reduplicated: √mat^h → //ma-mt^h-// → mant^h-. The same situation obtains for "√mad, mand- 'be exhilarated, exhilarate' " (p. 118). Another root given by Whitney is "√nand 'rejoice' (pp. 87-88), with a present *nandati*. He implies that this root is to be connected with √nad 'sound'. It might alternatively be possible to connect it with √mad, with reduplicative copying of the place-assimilated nasal. While these could logically be viewed as reduplicated in origin, it is difficult to rule out other explanations (e.g., the medial [n] being etymological or being originally the nasal infix). In all cases, the forms with the doubled nasals are attested already in Vedic.

10.2. Interpretation

- In the cells to the left of the dividing line, the vast majority of forms show $C_1\bar{e}C_2$.
    - This follows the pattern seen for cluster-initial roots.
- The exceptions all date to the earliest attested period of Sanskrit (the Rig-Veda).
    - For the roots with doublets, the C_1-copying form is older:
    - *paptur* > *pētur*, *saçcur* > *sēcire*
- Three of the exceptions show non-identity between reduplicated consonant and surface root-C_1 (they are italicized in the table):
    - *√sac → saçcur*, *√g^has → jakşur*, *√b^has → bapsati*
    - Velar palatalization certainly escapes a PCR violation (there are no velar-initial roots that take the $C_1\bar{e}C_2$ pattern).
    - Whether the other types of non-identity are significant enough to escape a PCR violation is unclear, but this may point to that being the case.
- To the right of the line, most of the examples show C_1-copying, as expected relative to the cluster-initial roots.
    - Some of the counterexamples can be ruled out independently by phonotactics.
    - e.g., *√p^hal → p^hēlirē* because ***pa-p^hl-ire* would have an illegal **-*p^hl*- sequence.
    - Here again, for roots with doublets, the C_1-copying form is older:
        - *tatnē* > *tēnē*, *sasni* / *sişņu* > [*sēnē*], *mamnāt^hē* > *mēnē*
    - Even for those $C_1\bar{e}C_2$ forms without doublets, they are mostly not attested until the later language (an exception being *tēnē*, already attested in the Atharva-Veda).
- While this leaves much up to interpretation, it seems likely that this picture represents one of transition, with approximately three major stages:
    (i) In the (prehistoric) stage that precedes Vedic Sanskrit, it might have been the case that all clusters which were phonotactically licit (or made phonotactically licit through assimilation) copied C_1 (i.e. default behavior).
        - This explains archaic forms like *paptur* and *saçcur*.
    (ii) In the Vedic (or at least post-Rig-Vedic) period, any cluster which did not have *CR transitions* took on the $C_1\bar{e}C_2$ allomorph.
    (iii) In Classical/Epic Sanskrit, there is some variation in *consonant-nasal* clusters, with the trend seeming to be moving towards $C_1\bar{e}C_2$, except when C_1 is a stop, in which case C_1-copying still predominates.
- If this characterization of Classical/Epic Sanskrit is correct, and the distribution at that point is still governed by some version of the PCR, then it seems that it is possible for there to be distinctions made between *CN transitions* and *CL transitions*.
    - The retention of C_1-copying for *stop-nasal* sequences indicates that *burst* and/or *intensity rise* could have become significant factors in the PCR of this stage.
- But since there does not seem to be any equivalent change in cluster-initial roots, it might be preferable to not attribute this to PCR, but rather to the general morphological / morpho-phonological change that is expanding the scope of the $C_1\bar{e}C_2$ pattern to include even non-CaC roots (see Sandell 2013).

10.3. A few more data points

- The following shows the behavior of *h*- and *affricate*-initial zero-grade bases.

(51) <u>Treatment of *h*- and *affricate*-initial zero-grade clusters in reduplication</u>

C_1 \ C_2	Stop	Nasal	Liquid	Glide
h			*jahrur*	*juhve* [27]
Affricate	jepur ✗ [cete] ✗	jajɲur cemur ✗	celur ✗ cerur ✗ jerur (✗)	

- The *h*-initial roots will always be exempted from PCR effects, because *h* reduplicates as *j* (for historical reasons).
- The only potential zero-grade cluster among these roots which is phonotactically licit is the *-jɲ-* of *jajɲur* (← √*jan*).
    - *-jr-* seems to be marginally permitted in the later language.
- Therefore, it is difficult to tell determine the "expected" behavior of affricate-initial clusters.

11. Appendix C: Preliminary analysis of the $C_1\bar{e}C_2$ pattern in Germanic

- To model this particular repair in parallel OT, we will need to make use *existential faithfulness* (Struijke 2000).

(52) ∃-MAX-C-IO:

Assign one violation mark * for every consonant in the input which does not have *at least one* correspondent in the output.

- In Struijke's theory, the reduplicant is directly subject to Input-Output correspondence; therefore, ∃-MAX-C-IO will be satisfied under any one of three circumstances:
    - (i) There is a single output correspondent of the consonant, and it is in the root.
    - (ii) There is a single output correspondent of the consonant, and it is in the reduplicant.
    - (iii) There are multiple output correspondents of the consonant.

- We will also need:

(53) ∃-ANCHOR-L-IO:

Assign one violation mark * if the segment at the left edge of the underlying root does not have some surface correspondent at the left edge of the output.[28]

[27] This may belong with the cluster-initial root forms.
[28] This conflates certain issues relating to underlying order. It is more straightforwardly captured in an Input-Reduplicant faithfulness model; but it may not be desirable to simultaneously employ IR-faithfulness and existential faithfulness, as the problems which they are invoked in order to solve at least partially overlap.

(54) ∃-$\text{M}\textsc{ax}_{\text{RT}}$-$\text{X}_{\text{VC}}$-IO (based on Yun 2014b's system for compensatory lengthening)

Assign one violation mark * if there is a timing slot in the input, which is associated with the root and belongs to an underlying VC sequence, which does not have some surface correspondent in the output.

- When ranked as in (55), we can generate the desired form.

(55) <u>Pre-Germanic Class V preterite plurals in $C_1\bar{e}C_2$</u>

/RED, e, $\mathbf{g_1b_2}$, um/	DEP-IO	∃-MAX-C-IO	∃-ANCH-L-IO	PCR (Pre-Gmc)	C/V	∃-ANCH-L-BR	∃-MAX_{RT}-X_{VC}-IO
a. $\underline{g_1\text{-e}}\text{-}\mathbf{g_1b_2}\text{-um}$				*!	*		
b. $\underline{g_1\text{-e}}\text{-}\mathbf{b_2}\text{-um}$						*	*!
c. ☞ $\underline{g_1\text{-}\bar{e}}\text{-}\mathbf{b_2}\text{-um}$						*	
d. $\underline{\text{e}}\text{-}\mathbf{g_1b_2}\text{-um}$			*!		*		
e. $\underline{g_1\text{-e}}\text{-}\mathbf{g_1}\text{-um}$		*!					
f. $\underline{b_2\text{-e}}\text{-}\mathbf{b_2}\text{-um}$		*!	*!				*
g. $\underline{b_2\text{-e}}\text{-}\mathbf{g_1}\text{-um}$			*!			*	
h. $\underline{b_2\text{-e}}\text{-}\mathbf{g_1b_2}\text{-um}$			*!		*	*	
i. $\underline{g_1b_2\text{-e}}\text{-}\mathbf{g_1b_2}\text{-um}$					*!*		
j. $\underline{g_1\text{e}b_2\text{-e}}\text{-}\mathbf{g_1b_2}\text{-um}$	*!				*		