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

WeCIEC 27, 10/24/15

Sam Zukoff, MIT
szukoff@mit.edu, web.mit.edu/szukoff/www

1. Introduction

- The reduplicative systems of Ancient Greek, Gothic, and Sanskrit each display 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.
    - e.g., √*pak-* → reduplicated *pe-pak-*
  - Roots with initial *stop-sonorant* (TR) clusters tend to also show C_1-copying.
    - e.g., √*prak-* → reduplicated *pe-prak-*
  - 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	√*stak-*	*e-stak-*
Gothic	Cluster-copying	√*stak-*	*ste-stak-*
Sanskrit (cluster-initial roots)	C_2-copying	√*stak-*	*ta-stak-*
Sanskrit (zero-grade bases)	"C_1ēC_2" pattern	(√*sat-* →) *st-*	*sēt-*

- 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).
  - I will refer to this interaction as the *Poorly-Cued Repetition Principle* (PCR).

- Based on the distribution of phonetic cues in different types of clusters, PCR can generate different behaviors for ST clusters than for TR clusters.
  - The repetition in a form *se-stak-* **violates** PCR, causing the alternative patterns in (1),
  - But the repetition in *pe-prak-* **satisfies** PCR, allowing default C_1-copying.

[*] Special thanks to Donca Steriade, Adam Albright, Andrew Byrd, András Cser, Edward Flemming, Jay Jasanoff, Michael Kenstowicz, Tyler Lau, Jesse Lundquist, Craig Melchert, Jeremey Rau, Ryan Sandell, Juliet Stanton, Tony Yates, and Suyeon Yun, the audience at the LSA Annual Meeting 2015, the audience at the Harvard GSAS Workshop on Historical and IE Linguistics, and audiences at MIT. All mistakes are, of course, my own. For an earlier version of this handout with additional data & analysis, see: http://web.mit.edu/szukoff/www/pdfs/Harvard_IE_Workshop15.pdf

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): *subjects are unable to perceive repeated tokens as being separate entities* (Walter 2007:§5).

- It is also generally accepted that consonants are dispreferred in contexts where they are less perceptible ("Licensing by Cue"; Steriade 1997).

❖ The central proposal in this paper is a logical extension of these two principles:
- Due to *repetition blindness*, listeners are biased to fail to identify the presence of a locally-repeated segment.
- Since *perception* is based on phonetic cues, listeners also have difficulty identifying the presence of a consonant when it lacks robust phonetic cues to its presence.
- ➢ Therefore, when a single stimulus meets both of these criteria (i.e. both biases obtain), listeners will be especially unlikely to accurately perceive that stimulus.

(2) THE POORLY-CUED REPETITION PRINCIPLE:

> 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.

➢ Since cluster types differ with respect to which phonetic cues they contain, this constraint can begin deriving the differences in copying behavior between different clusters.
- The phonetic cues that are most significant to perceiving the *presence* of a consonant[1] are listed in (3) below (see Wright 2004).

(3) Phonetic cues to the presence of consonant

a. **burst** stop release burst
→ *present in stop-initial and nasal-initial sequences*

b. **intensity rise** rise in spectral intensity between two segments
→ *present in rising sonority sequences*

c. **CR transitions** spectral transitions between a consonant and a following sonorant
→ *present in consonant-sonorant sequences*

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

- These cues are distributed very differently in different types of clusters:
  - *Stop-sonorant* (TR) clusters have **all** three cues
    - Stops have a burst
    - TR clusters rise in sonority, and thus have an intensity rise
    - Sonorants host CR transitions
  - *Fricative-stop* (ST) clusters have **none** of these cues[2]
    - Fricatives have no burst
    - ST clusters fall in sonority, and thus do not have an intensity rise
    - Stops are obstruents and thus do not host transitions

- I propose that the Poorly-Cued Repetition Principle is instantiated in the phonological grammar with a markedness constraint, provisionally defined in (4).

(4) THE POORLY-CUED REPETITION CONSTRAINT (PCR):
Assign a violation mark * to any $C_\alpha VC_\alpha$ sequence where the second consonant[3] does not bear the **requisite phonetic cues** to its presence.
►**REQUISITE CUES**: *burst*, *intensity rise*, and *CR transitions*

- When this constraint is active in the phonological grammar, forms with poorly-cued repetitions may be actively avoided.
  - This is what induces the alternative patterns which we observed in (1).

TR and ST clusters represent the two poles of "cued-ness" of their first member: TR clusters have the best cued first member, ST clusters have the worst cued first member. This predicts that, if a language is to make any default/non-default distinction based on cluster type, TR clusters will show default behavior and ST clusters will show non-default behavior. *This is indeed what we find across the Indo-European languages.* Other clusters that fall in the middle show less consistent behavior, as would be expected. See Appendix for discussion.

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.
- In light of the distribution of cues laid out above, this distinction follows from the *Poorly-Cued Repetition Principle*, and the PCR constraint defined in (4).

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)a), but "non-copying" in roots with an initial *s-stop* (ST) cluster (shown in (5)b).

[2] The *frication noise* present in sibilants, which is normally a strong cue to the presence of a fricative, does not appear to be a strong cue in the repetition context, perhaps because *repetition blindness* decreases the efficacy of cues like frication noise which extend over a significant duration.
[3] 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).

(5) Ancient Greek perfects by cluster type

a. TR roots → C_1-copying perfects

Root		Perfect Tense	
kri-	'decide'	κέκριμαι [k-e-kri-mai]	not **[e-kri-mai]
tla-	'suffer, dare'	τέτληκα [t-e-tlē-ka]	not **[e-tlē-ka]
pneu-	'breathe'	πέπνυμαι [p-e-pnū-mai]	not **[e-pnū-mai]

b. ST roots → non-copying perfects

Root		Perfect Tense	
stel-	'prepare'	ἔσταλκα [e-stal-ka]	not **[s-e-stal-ka]
strateu-	'wage war'	ἐστράτευμαι [e-strateu-mai]	not **[s-e-strateu-mai]

- The PCR constraint provides a way of penalizing C_1-copying to ST roots; but the specific strategy which will be employed in order to avoid C_1-copying is dependent on the ranking of other constraints.

(6) Constraints modulating potential repairs for PCR

a. ONSET:
Assign a violation mark * for each onsetless syllable.
Potential PCR Repair: V-C_1C_2V- (candidates (b))

b. *CC (≈ *COMPLEX):
Assign a violation mark * for every consonant cluster.[4]
Potential PCR Repair: 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: C_2V-C_1C_2V- (candidates (d))

- To generate the non-copying repair for Ancient Greek, ONSET must be the lowest ranked of these constraints, and it must also be dominated by PCR. This ranking derives the proper form, as shown in (7) below.

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

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

[4] I will only mark violations of *CC 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. Unlike Gothic and Sanskrit, the "reduplicative" vowel in Greek is not copied, but rather underlying (cf. Zukoff under review).

- Candidate (a) is the default C_1-copying form.
    - It violates PCR since it has a consonant repetition which is not cued by *burst*, *intensity rise*, and *CR transitions*.
    - PCR thus eliminates this candidate.
- 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 this 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 *CC violation.
- The optimal candidate (b) copies nothing, at the expense only of low-ranked ONSET, whose violation is tolerable in service of PCR.

- For TR roots, the repetition created by copying C_1 is well-enough cued to satisfy PCR, since it has *burst*, *intensity rise*, and *CR transitions*.
    - This makes the ONSET violation incurred by the non-copying candidate unnecessary.
    - Therefore, C_1-copying is permitted to TR roots, as demonstrated in (8):

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

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

3.2. Cluster-copying reduplicated preterites in Gothic

- In Gothic there are few relevant examples, but they again point to a distinction between TR and ST roots.
    - TR roots follow the default C_1-copying pattern ((9)a).
    - ST roots display cluster-copying, i.e. a reduplicant in *<u>STe</u>-* ((9)b).

(9) <u>Class VII preterites in Gothic (forms from Lambdin 2006:115)</u>

a. TR roots → C_1-copying preterites

	Infinitive	Preterite	
'to weep'	*gretan* [grēt-an]	*gaigrot* [<u>ge</u>-grōt]	not **<u>gre</u>-grōt

b. ST roots → cluster-copying preterites

	Infinitive	Preterite	
'to possess'	*staldan* [stald-an]	*staistald* [<u>ste</u>-stald]	not **[<u>se</u>-stald]
'to divide'	*skaidan* [skaið-an]	*skaiskaiþ* [<u>ske</u>-skaiθ]	not **[<u>se</u>-skaiθ]

- This pattern falls out if we take the constraints and rankings proposed for Ancient Greek and simply swap ONSET and *CC:[6]

[6] The mapping to a reduplicated preterite in synchronic Gothic is likely more complicated than presented in this tableau; see Sandell & Zukoff (2015).

(10) Cluster-copying in ST roots in Gothic (PCR violation): √*stald-* → *<u>ste</u>-stald* 'he possessed'

/RED, stald/	ANCHOR-L-BR	ONSET	PCR	*CC
a. <u>se</u>-stald			*!	*
b. <u>e</u>-stald	*!	*!		*
c. ☞ <u>ste</u>-stald				**
d. <u>te</u>-stald	*!			*

- With the change in relative rankings, the viable alternative in Gothic to the PCR-violating C_1-copying 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.

- When a poorly-cued repetition is not at stake, C_1-copying is again preferred:

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

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

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

- Sanskrit also displays distinct behavior for different cluster types in reduplication, but with an added twist:
  - There are two distinct non-default treatments, depending on the morphophonological origin of the base-initial cluster.
- Nevertheless, the distribution of default vs. non-default treatment in both categories adheres to the *Poorly-Cued Repetition Principle*.

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 (12).
  - TR roots again show default C_1-copying ((12)a).
  - ST roots here show C_2-copying ((12)b).

(12) <u>Perfects to cluster-initial roots in Sanskrit (forms from Whitney 1885 [1988])</u>

a. TR roots → C_1-copying perfects

Root	Perfect Tense	
b^{h}raj- 'shine'	<u>ba</u>-b^{h}rāj-a	not **<u>ra</u>-b^{h}rāj-a
drā- 'sleep'	<u>da</u>-drā-u	not **<u>ra</u>-drā-u
prac^h-'make'	<u>pa</u>-prāc^h-a	not **<u>ra</u>-prāc^h-a

b. ST roots → C_2-copying perfects

Root		Perfect Tense	
st^{h}ā-	'stand'	ta-st^{h}ā-u	not **sa-st^{h}ā-u
stamb^{h}-	'prop'	ta-stamb^{h}-a	not **sa-stamb^{h}-a
sparç-	'touch'	pa-spr̥ç-ē	not **sa-spr̥ç-ē

- To derive the C_2-copying pattern for the ST roots in Sanskrit, we again need only permute the rankings proposed earlier for Greek and Gothic.
  - If ANCHOR-L-BR is the uniquely lowest-ranked relevant constraint, we generate C_2-copying as the repair for a PCR violation. This is shown in (13):

(13) 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	*CC	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 roots continue to copy C_1:

(14) 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 make perfect active plurals and perfect middles, 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, as in (15).

(15) 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, **ra-b^{h}r-ē
d^{h}ar-	'hold'	da-d^{h}r-ē	not **d^{h}ēr-ē, **ra-d^{h}r-ē
par-	'fill'	pa-pr-ur	not **pēr-ur, **ra-pr-ur

- If this new cluster would be an ST cluster, as for the roots in (16) below, 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.

---

[7] There are two *stop-liquid* roots that take the C_1ēC_2 pattern: *tērur* ← √*tar* 'pass' and *p^{h}ēlur* ← √*p^{h}al* 'burst; fruit'.

- But the non-default treatment is not C_2-copying; instead we see selection of a different allomorph, the "$C_1ēC_2$ pattern": /C_1aC_2/ → [$C_1ēC_2$-].

(16) $C_1ēC_2$ 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

- We can explain why C_2-copying is not available for these roots by using Input-Reduplicant (IR) faithfulness (McCarthy & Prince 1995), specifically the constraint LINEARITY-IR.
  - This constraint will be *violated* by copying a consonant that underlyingly follows the root vowel (as in these cases), but not by copying a root consonant that underlyingly precedes the root vowel (as is the case for C_2-copying from root-initial clusters).

- LINEARITY-IR therefore blocks C_2-copying for these bases, and forces the use of a secondary repair strategy for PCR, namely the $C_1ēC_2$ allomorph.
  - This allows us to generate the four-part distribution shown in (17):[9]

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

	PCR-violating	**PCR-satisfying**
	ST cluster	TR cluster
Zero-grade base	$C_1ēC_2$	C_1-copying
Cluster-initial root	C_2-copying	C_1-copying

3.4. Other PCR effects in IE reduplication

- There are several more reduplicative patterns in Indo-European which are motivated by PCR.

(18) Latin infixing perfect reduplication to ST roots (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-i*

- ❖ The reduplicant moves inside the root to avoid creating an SVST sequence, which would violate PCR.

[8] This form, and possibly also the examples with /ç/, can be ruled out independently on phonotactic grounds. Nonetheless, there are many other cluster types which undergo the $C_1ēC_2$ pattern despite being phonotactically licit.
[9] We also need "Use X"-type allomorphy constraints to properly select the $C_1ēC_2$ allomorph if we ultimately treat this as morphological, following Sandell (2015:§8). The necessary ranking is: USE REDUPLICATION » USE CēC.

(19) (Classical) Sanskrit infixing desiderative reduplication to VCC roots

	Root		Desiderative	
	√*arc* 'praise'	→	*ar-ci-c-iṣ-*,	not ***a-ri-rc-iṣ-*
	√*ard* 'stir'	→	*ar-di-d-iṣ-*,	not ***a-ri-rd-iṣ-*
	√*ubj* 'force'	→	*ub-ji-j-iṣ-*,	not ***u-bi-bj-iṣ-*
Cf.	√*akṣ* 'attain'	→	*ā-ci-kṣ-iṣ-*,	not ***āk-ṣi-ṣ-iṣ-*

- ❖ Minimal infixation would place the reduplicant between V and C_1.[10]
  - However, given the types of VCC roots which are attested, this almost always would result in a PCR violation.
  - Therefore, the infixation moves further inward to between C_1 and C_2.
  - For the root *akṣ*, infixation between V and C_1 does not result in a PCR violation due to velar palatalization, and therefore we do see minimal infixation.

(20) Attic Reduplication perfects in Pre-Greek (cf. Zukoff 2014, under review)

Ancient Greek	<	Pre-Greek		
perf. *edēd-*	<	√*$h_1$ed-* 'eat'	→ perf. *h_1əd-e-h_1d-*,	not ***h_1-e-h_1d-*
perf. *agēger-*	<	√*$h_2$ger-* 'gather'	→ perf. *h_2əg-e-$h_2$ger-*,	not ***h_2-e-$h_2$ger-*
perf. *olōl-*	<	√*$h_3$el-* 'destroy'	→ perf. *h_3əl-e-h_3l-*,	not ***h_3-e-h_3l-*

- ❖ Cluster-copying + reduplicant internal epenthesis repairs a PCR-type violation which is specific to laryngeals.

4. The C_1ēC_2 pattern in Sanskrit, Germanic, and elsewhere

- We have seen that Sanskrit selects a C_1ēC_2 allomorph in zero-grade categories of the reduplicated perfect.
  - o This same pattern can be identified in several other Indo-European languages.

- ❖ It is seen most robustly in the preterite plurals of CeC strong verbs in Germanic (Class IV-V), which display an otherwise rather unexpected long vowel [ē] in the root:

(21) Gothic Class IV & V preterite plurals (forms from Lambdin 2006:51)

	Infinitive	Preterite Plural (1PL.)	
'to bear'	*bairan* [ber-an]	*berum* [bērum]	(as if from *be-br-um)
'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)

[10] The motivation for infixation in the first place appears to be to prevent the initial vowel from being preceded by a consonant in the reduplicated form when it is absolute initial in the isolation form.

❖ Similar forms also exist in Old Irish:

(22) Old Irish ā-preterites (Thurneysen 1966:429; cf. Niepokuj 1997:151-2)

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-)

➢ From a historical/derivational perspective, this pattern could be viewed as reduplication followed by deletion of root-C_1 + compensatory lengthening.
  - Table (23) below illustrates this for the Pre-Germanic stage which produces the Class IV/V preterite plurals.

(23) Deriving $C_1\bar{e}C_2$ in Pre-Germanic

	/RED, C_1eC_2, um/
Reduplicate: copy CV	C_1e-C_1eC_2-um
Zero-grade: delete root vowel	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.
  - For CeC roots, zero-grade ablaut places a consonant-repetition before another consonant (specifically an obstruent in Class V), leaving the repetition poorly-cued.[11]
- 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-CR̩C-

❖ Though it has nothing to do with ablaut, Melchert (2015) also suggests that Hittite *šipand-* 'to libate' (beside normal *išpand-*) should be viewed as reduplication + deletion:

(24) Hittite: /RED, špand-/ → reduplication → *ši-špand-* → $C_\alpha VC_\alpha C$ deletion → [šipand-]

  - Since this is occurring with an ST cluster in reduplication, it seems quite likely that this is also a PCR effect.

➢ The $C_1\bar{e}C_2$ pattern can thus be generated phonologically (at least in a derivational theory of phonology).[12]
  - Therefore, at least at some stage, we do not need to appeal to allomorph selection to generate the pattern (as was implied above for Sanskrit).

[11] It is unclear to me whether the Class IV (CeR roots) preterite plurals participated in this pattern at this stage
[12] It is actually fairly difficult to derive this pattern in a parallel, non-derivational version of OT.

- Furthermore, a similar phonological grammar could be used to generate the patterns in each of these languages.
    - This suggests that the pattern, and the grammar that generates it, could potentially be projected even farther back to Proto-Indo-European.
    - However, given the differing 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. PCR effects outside of reduplication

5.1. An exception to Bartholomae's Law

- Sanskrit has a set of voiced aspirated stops [bʱ, dʱ, ḍʱ, gʱ] (= Dʱ), but this aspiration is only licensed in pre-sonorant position (i.e. cannot surface pre-obstruent or word-final).
- When an underlying Dʱ would surface in a position where aspiration is not licensed, the aspiration can migrate to a nearby stop, in one of two ways:[13]

(25) Patterns of aspiration mobility

a. **Aspiration Throw Back (ATB)**
If the preceding consonant is a *plain voiced stop*, the aspiration can surface on that stop.
❖ √budʱ 'know' : root noun LOC.PL. /**budʱ**-su/ → [**bʱ**u**t**-su]

b. **Bartholomae's Law (BL)**
If the immediately following consonant is a *plain stop*, the aspiration can surface on that stop. (That stop also becomes voiced.)
❖ √rudʱ 'obstruct' : nasal-infix present 3SG. /ru-na-**dʱ**-**t**i/ → [ru-na-**d**-**dʱ**i]

- 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:

(26) Reduplicated present of √dʱā 'place' (white cells display ATB)

	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

[13] I adopt the position that the IE "diaspirate" roots are represented synchronically in Sanskrit as /DVDʱ/ not /DʱVDʱ/.

- The ACTIVE 3DL and MIDDLE 3SG (shown in bold) have the conditions to support either BL or ATB, but they show ATB rather than BL, *contrary to normal preference* (cf. [bu**d-dʱ**a-]).

- If BL had applied to these forms, it would have created a poorly-cued repetition:
    - MID.3SG //da-dʱ-te// → BL → **[**da-d**-dʱe]

➢ Therefore, the choice of the normally dispreferred process (ATB) over the default process (BL) can be viewed as a PCR effect.

5.2. The Latin –*is*… endings

- Cser (2015:13) documents allomorphy involving the perfect endings -*(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.
        - This is easily interpretable as hiatus avoidance.
    - For stems ending in /s/, we see free variation according to the pattern in (27):

(27) Variability in *s*-final stems

a. *access-istis* ~ *acces-tis*	c. *admis-isse* ~ *admis-se*,
b. *divis-isse* ~ *divis-se*	d. *dire*[*ks*]-*isti* ~ *dire*[*ks*]-*ti*[14]

- Use of the full form of these suffixes creates a PCR violation just in the case of /s/-final roots.
    - The PCR constraint can thus militate against the use of the full forms.

6. Conclusion and discussion

- In the reduplicative systems of the Indo-European languages, there are a number of cases in which default C_1-copying is blocked, and an alternative pattern is found instead.
- In this paper, I have argued that these cases can be unified as repair/avoidance strategies for the *Poorly-Cued Repetition Principle*:

(28) 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.

- This proposal identifies repetition blindness (Kanwisher 1987), coupled with, and framed in terms of, cue-based perception (see, e.g., Wright 2004), as (one of) the motivation(s) for repetition avoidance phenomena (cf. Walter 2007).

- The PCR constraint can induce avoidance of C_1-copying in reduplication, contrary to the normal pattern of the Indo-European languages. These patterns are summarized in (29):

[14] Cited by Cser as "*derexisti* ~ *derexti* 'arrange'."

(29) Reduplicative PCR effects and details

Language	Non-default treatment induced by PCR	Context
Ancient Greek	Non-copying	ST roots
Pre-Greek	Attic Reduplication	HC roots and bases
Gothic	Cluster-copying	ST roots
Pre-Germanic	$C_1\bar{e}C_2$ preterite plurals	CeC roots
Sanskrit	C_2-copying	ST roots
	$C_1\bar{e}C_2$ allomorphy	Zero-grade bases
	Non-minimal infixation	VCC desideratives
Latin	Infixing reduplication	ST roots
Old Irish	$C_1\bar{a}C_2$ preterites	CaC roots
Hittite	*ši-špand-* → *šipand-*	ST roots

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

❖ Various other "haplology" effects around IE may well be amenable to a PCR explanation.

6.1. Discussion

➢ The PCR proposal has a number of advantages over previous accounts.

Fleischhacker (2005) develops a general theory of cluster simplification driven by facts about perceptual similarity.

- To explain cluster simplification in reduplication, she uses employs a fixed ranking of similarity-based constraints on the Base-Reduplicant relationship.
- The distributions in Greek, Gothic, and Sanskrit (see Appendix), and the broader typology, are derived by interpolating other phonotactic constraints at different points in the fixed ranking.
- Based on the evidence I have presented, there are at least two shortcomings of Fleischhacker's approach:
  - The $C_1\bar{e}C_2$ pattern finds no logical interpretation within her system, as it results in completely dissimilar consonants between Base and Reduplicant.[15]
  - Because it is tied to the Base-Reduplicant relationship, her account cannot explain the non-reduplicative cases discussed in §5.
- Fleischhacker's account though does have benefits not shared by this proposal, namely that it ties cluster simplification in partial reduplication to other domains where similarity drives cluster-dependent processes (epenthesis, rhyme, alliteration, etc.).
  - Though we may wonder if the results attributed by Fleischhacker to "perceptual similarity" might not alternatively be expressible in terms of the distribution of phonetic cues, of the sort used in this paper.

[15] But note again that, even in the PCR system, generating this pattern in a parallel, non-derivational version of OT is far from trivial, so this critique must be tempered.

Keydana (2012) derives the distinct behavior of ST clusters, and the variation in repairs across the languages, through assumptions about the phonological representation of ST clusters.

- He asserts that ST clusters should be analyzed as a type of complex segment (with slightly different properties in the different languages) and that this can be used to derive their exceptional behavior.
- However, as demonstrated in the Appendix, the division is not simply between ST clusters and all other clusters; it is rather the case that clusters which fall between the two poles of ST and TR can have differing affinities for those two poles.
- Keydana would thus be forced to say that many other types of clusters, e.g. TT and SN in Ancient Greek, function as complex segments. This is not an appealing solution.

Zukoff (2014) used a syllable-based OCP constraint (cf. McCarthy 1986) to account for the facts in Greek.

- It required positing that clusters that show non-default behavior were heterosyllabified while those that show C_1-copying were tautosyllabified.
- These syllabifications do not comport with evidence from syllable weight in weight-sensitive phonological processes and metrics, both of which point strongly to all types of clusters being identical for the purposes of weight computation (at least word-internally) (see Steriade 2015 for recent discussion).
- Therefore, we require a non-syllable-based analysis, such as PCR.

References

Cser, András. 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.

Keydana, Götz. 2012. Evidence for Non-Linear Phonological Structure in Indo-European: The Case of Fricative Clusters. In *The Sound of Indo-European: Phonetics, Phonemics, and Morphophonemics*, edited by Benedicte Nielsen Whitehead, Thomas Olander, and Birgit Anette Olsen, 99–117. Denmark: Museum Tusculanum Press.

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. 1986. OCP Effects: Gemination and Antigemination. *Linguistic Inquiry* 17: 207–63.

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.

Melchert, H. Craig. 2015. Initial *sp- in Hittite and *šipand*- 'to libate'. Paper presented at the 34th East Coast Indo-European Conference, Vienna, June 6.
Niepokuj, Mary. 1997. *The Development of Verbal Reduplication in Indo-European*. Washington, D.C.: Institute for the Study of Man.
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. 2015. Productivity in Historical Linguistics: Computational Perspectives on Word-Formation in Ancient Greek and Sanskrit. PhD Dissertation, UCLA.
Sandell, Ryan, and Sam Zukoff. 2015. A New Approach to the Origin of Germanic Strong Preterites. In *the Proceedings of NELS 45*.
Steriade, Donca. 1997. Phonetics in Phonology: The Case of Laryngeal Neutralization. Ms., UCLA.
Steriade, Donca. 2015. Changes in the Weight of Cluster: Evidence from Ancient Greek. Paper presented at the LSA Annual Meeting, Portland, Oregon, January 8-11.
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.
Zukoff, Sam. 2014. On the Origins of Attic Reduplication. In *Proceedings of the 25th UCLA Indo-European Conference*, edited by Stephanie W. Jamison, H. Craig Melchert, and Brent Vine, 257–78. Bremen: Hempen.
Zukoff, Sam. under review. The Reduplicative System of Ancient Greek and a New Analysis of Attic Reduplication. Ms, MIT.

Appendix: Refining PCR – the behavior of other cluster types

- §3 demonstrated that a PCR constraint based on *burst*, *intensity rise*, and *CR transitions* can 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.

➢ Which of these clusters pattern with TR and which pattern with ST in the respective languages?
➢ Does PCR need to be adjusted to capture these distinctions; and, if so, how?

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 (30) below illustrate these distributions.

Grey cells = initial clusters that are not attested in the perfect/preterite
✓ = clusters with default C_1-copying
× = clusters with non-default treatment (Greek: non-copying; Gothic: cluster-copying)

(30) Initial clusters and reduplicative behavior[16]

Greek

C_1 \ C_2	Stop (T)	Fricative (S)	Nasal (N)	Liquid (L)
Stop	× e-kton-	× e-pseus-	✓ p-e-pnū-	✓ k-e-kri-
Fricative	× e-stal-	(×)[17] (e-ssu-)	× e-smε̄g-	(grey)

Gothic

C_1 \ C_2	Stop (T)	Fricative (S)	Nasal (N)	Liquid (L)
Stop	(grey)	(grey)	(grey)	✓ gε-grōt
Fricative	× stε-stald	(grey)	(grey)	✓ fε-flōk

- Both languages attest two cluster types with default treatment, *stop-liquid* (TL) and one other:
  - Greek permits C_1-copying to *stop-nasal* (TN), but does not attest *fricative-liquid* (SL).
  - Gothic permits C_1-copying to *fricative-liquid* (SL, S = {f,s}), but does not attest TN.
- Both languages show non-default treatment for ST.
  - Gothic has no further clusters with reduplication.[18]
  - Greek has a more robust cluster inventory, but all remaining types show non-default treatment (non-copying).

- Table (31) compares these distributions more directly.
  - We see that, unfortunately, there are no cluster types beyond TL and ST which are attested in both languages, and thus we cannot perfectly compare the two.
  - The closest point of comparison is the behavior of *fricative-sonorant* clusters, represented by SN (S = {s}) in Greek and SL (S = {f,s}) in Gothic.

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

C_1 \ C_2	**Stop (T)**	**Nasal (N)**	**Liquid (L)**
Stop (T)	Greek: ×	Greek: ✓	Greek: ✓
	Gothic: Ø	Gothic: Ø	Gothic: ✓
Fricative (S)	Greek: ×	Greek: ×	Greek: Ø
	Gothic: ×	Gothic: Ø	Gothic: ✓

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

- Table (32) below shows the distribution of phonetic cues in these different cluster types.

[16] Consult van de Laar (2000) for the range of forms in Greek; consult Lambdin (2006) for the forms in Gothic.
[17] 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.
[18] A few additional root shapes (*sm-*, *sn-*, *sw-*, etc.) exist in Gothic, but are not attested among reduplicating roots.

(32) 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

- SN and SL have the same set of available cues.[19]

➢ Since Greek and Gothic differ in how they treat these clusters, this implies that their respective versions of PCR are different.
  - o Such a difference can be captured if we posit that the "requisite cues" clause of the constraint differs between the two languages.

- For Gothic, the "requisite cues" clause must pick out TL and SL to the exclusion of ST.
  - o Two cues independently satisfy this condition: *intensity rise* and *CR transitions*.
  - o *Burst* redundantly allows TL, but cannot on its own allow SL.

(33) 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*

- Given Greek's increased cluster inventory, the calculation of requisite cues takes a bit more thought: it must permit TR (TN & TL), but exclude ST, TT, and SN.
  - o Both TN & TL have *burst*, but so does TT.
  - o Both TN & TL have *intensity rise* and *CR transitions*, but so does SN.
  - ➢ Therefore, Greek must require **both** *burst* + *intensity rise* or *burst* + *CR transitions* (or possibly all three).

(34) 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**: (i) *burst*, and
(ii) *intensity rise* and/or *CR transitions*

- We therefore require two slightly different PCR constraints for Greek and Gothic, but each has a well-defined set of "requisite cues".

[19] They are identical only insofar as if there is no significant difference between CR transitions involving nasals vs. those involving liquids. This may not be the case.

Sanskrit

- The distribution of C_1- vs. C_2-copying in Sanskrit cluster-initial roots is shown in (35).

(35) Attested clusters and reduplicative behavior in Sanskrit cluster-initial roots

C_1 \ C_2	Stop	Affricate	Fricative	Nasal	Liquid	Glide
Stop			✓(?)	✓	✓	✓
Affricate				✓	✓	✓
Fricative	✗	✗		✓	✓	✓
Nasal				✓(?)	✓	✓

- 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.[20]
- Additionally, it also seems to permit C_1-copying to *stop-fricative* (TS) clusters (*ps*, *ts*, *kṣ*), though the data is minimal.[21]
  - Since the stop in these clusters bears none of the cues available in TR clusters, if we do want to permit these clusters via PCR, we need an alternative cue.
  - Fricatives, particularly sibilant fricatives, likely bear some sort of (relatively weak) transitional cue.
  - Therefore, it might be that *any sort of transitions* is sufficient for the Sanskrit PCR.
- 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* (or possibly any transitions):

(36) 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* (and possibly *TS 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.

---

[20] 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.

[21] There are only two relevant examples: (i) √*tsar*- 'approach stealthily' → perfect *tatsāra*, which is only attested in the Rig-Veda (Whitney 1885 [1988]:68), and may represent a sort of archaism; and (ii) √*psā*- 'devour' (built to the basic root √*b^h^as* 'devour' with an extension in -*ā*-; Whitney 1885 [1988]:104) → perfect *papsāu*, which 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*.