Stress windows and Base Faithfulness in English suffixal derivatives

Juliet Stanton and Donca Steriade (MIT)

1 Introduction

1. This study proposes a new analysis of the phonological cycle, and illustrates its success with evidence from English stress.

2. **Our claim:** in the productive computation of the phonology of any morphologically complex expression, *all words lexically related to it are potential bases* ("base" = a derived input, a surface form relative to which the target form is expected to be faithful (Benua 1997)).

   We distinguish two types of bases:

   a. *Local bases* (B_Ls): exponents of derivatives' immediate syntactic subconstituents.
   b. *Remote bases* (B_Rs): forms lexically related to derivatives, but distinct from their local bases.

   Example: *átomicity* has B_L *atómic*, B_R *átom*.

   We propose that remote bases, in addition to local bases, are relevant in calculations of Base-Derivative faithfulness.

3. The English word *àtomícity* (B_L: *atómic*; B_R: *átom*) will introduce the main idea. In a multi-layered form like *atomicity*, faithfulness can be evaluated in at least three ways:

   a. non-cyclically, relative to underlying forms only:

   UR /ætɔm/ /ɪk/ /ɪti/
   D [ˌæɾəm –ˈɪs – əɾi]

   b. cyclically, i.e. relative to a derived representation of the *maximal* inner stem, the B_L:

   UR /ætɔm/ /ɪk/ /ɪti/
   B_L [əˈtɑmɪk]
   D [ˌæɾəmˈɪs – əɾi]

   c. relative to a derived representation of the inner stem, the B_R, and the underlying forms of affixes:

   UR /ætɔm/ /ɪk/ /ɪti/
   B_L [əˈtɑmɪk]
   B_R [ˈæɾəm]
   D [ˌæɾəm] –ˈɪs– əɾi]

4. We focus here on the less well-documented option, (3c). Elements of the analysis:

a. A preference exists for the local base to be the unique correspondent of the derivative's stem. We encode this preference as a violable constraint, CORRB_L

- CORRB_L: assign a * for each derivative whose stem does not correspond with its B_L.

b. This preference coexists however with a broader, undifferentiated form of faithfulness that can be satisfied by letting the inner stem of the derivative correspond to the stem of any lexically related form (General Correspondence, or GC). We assume that GC is inviolable, i.e. all derivatives must correspond with some base.

5. By expressing the preference for correspondence to local bases as a violable constraint, we reject the idea that locality is a defining property of cyclic derivations.[1]

a. **Prediction:** the preference for local bases, CORRB_L can lose, when M >> CORRB_L.

b. **Supporting evidence:** cases where faithfulness to the local base is abandoned, as it entails unacceptable departures from markedness, but faithfulness to remote bases is still active.

6. A simplified derivation of *atomicity* shows how we put to use the correspondence pattern to B_R (3c).

Relevant constraints:

- BD-IDENT(stress): assign a * for each pair of correspondent syllables differing in stress
- *CLASH: assign a * to each sequence of adjacent stresses.
- *EXTLAPSE: assign a * to each sequence of three stressless syllables (000)

If accentual faithfulness is evaluated relative to the remote base *atom*, all M constraints, as well as BD-IDENT(stress), are satisfied.

atom + ic + ity B_L atómic 010 B_R átom 10	BD-IDENT(stress)	*CLASH	*EXTLAPSE	CORRB_L
> a. àtom_R ícity 20100				*
b. atómic$_L$ ity 01000			*!	
c. atòmíc$_L$ ity 02100	*!	*		
d. àtomíc$_L$ ity 20100	*!**			

7. Are there alternatives?

a. Benua 1997: forms like *atomicity* are obtained by allowing accentual M constraints to outrank BD-IDENT(stress).

atom + ic + ity B: atómic 010	*CLASH	*EXTLAPSE	BD-IDENT(stress)
> a. àtomícity 20100			***
b. atómicity 01000		*!	*
c. atòmícity 02100	*!		

[1] For the idea that only local bases can be bases, see Benua 1997: 30-31, also Kager 1999: 281-282.

b. Prediction: clash will *always* be resolved in *–ity* words. This is incorrect: we derive the wrong result for words with accentually invariant lexical families, e.g. *àlcohòlícity* (203100), below.

alcohol + ic + ity B: àlcohólic 2010	*CLASH	BD-IDENT(stress)
💣 a. àlcoholícity 200100		*
☹ b. àlcohòlícity 203100	*!	

8. Under our analysis, the difference between *àtomícity* and *àlcohòlícity* follows from independent facts about their lexical families. A *CLASH violation is inevitable in *àlcohòlícity* because all members of its lexical paradigm (*álcohòl, àlcohólic, àlcohólìsm*, etc.) carry some stress on *–hol–*.

alcohol + ic + ity B_L: àlcohólic 2010 B_R: álcohòl 201 (...)	BD-IDENT(stress)	*CLASH	CORRB_L
> a. àlcohòl$_R$ ícity 203100		*	*
b. àlcohòlíc$_L$ ity 203100	*!	*	
d. àlcohol$_L$ ícity 200100	*!		

We can *predict* the difference between *àlcohòlícity* (with clash) and *àtomícity* (without). When a B_R is available whose stress pattern could lead to an accentually improved D (= an *accentually preferable* B_R), it is used.

9. Roadmap:

- §2: We document that BD-IDENT >> most M for all types of English suffixal derivatives.
- §3: We provide evidence that most M >> CORRB_L.
- §4: Discussion of suffix-specific effects: some M_{Suff} >> BD-IDENT; other M_{Suff} >> CORRB_L.
- §5: We verify our analysis against the English lexicon.

2 Evidence for BD-IDENT(stress) >> M in all English suffixal derivatives

10. Work on English morphology has typically divided affixes into Level 1 and 2. Arguments for Level 1 or 2 status usually appeal to an affix's interaction with stress (stress-shifting 1 vs. stress-neutral 2).

While there is widespread agreement among authors that some suffixes belong to Level 1 (e.g. *–ity, –al*, *–ic*), others, like *–able* and *–ize*, are more contentious: they display characteristics of both levels.

Examples	**SPE (1968)**	**Siegel (1974)**	**Aronoff (1976)**	**Spencer (1991)**
–ity,–al, –ic, –(at)ion	Level 1	Level 1	Level 1	Level 1
–ize, –able	Level 2	Level 1	Level 1 & 2	Level 1 & 2

Our claim: there is no binary or n-ary partition. The "split level suffixes" (*–able, –ize*) have predictable behavior, given the structure of the derivatives' lexical families and their sometimes distinct markedness needs.

11. "Level 1" and "split-level" derivatives are completely faithful to the stress of their bases, modulo certain window effects that we discuss here (and some other suffix-specific effects, discussed later).

Windows aside, for every[2] general M constraint on stress, we can document the F >> M ranking.

12. These faithfulness effects emerge when we distinguish two regions within the derived word:

    a. A right-edge window (Kager 2012) whose stress may be constrained, in suffix-specific ways.

Suffix examples	*Window size*	*Window constraint*
–ee, –esque, –ese	1σ	STRESSR$_{Suff}$
–ic	2σ	*LAPSER$_{Suff}$
–al, –ity	3σ	*EXTLAPSER$_{Suff}$

    b. The rest of the derived word (the *faithful domain*)

    The (b) region displays faithfulness to the stress of the base even when stress must shift in the (a) region, to satisfy the window constraints.

13. Example: *inàlienabílity* 02000100

Local base:	0	1	0	0	0	0			
	in	a	li	en	a	ble			
–ity D:	0	2	0	0	0		1	0	0
	in	a	li	en	a		bi	li	ty
	Faithful domain						*Window*		

Ranking that generates this: *EXTLAPSER$_{Suff}$ >> BD-IDENT(stress) >> *EXTLAPSE

inalienable + ity B$_L$: inálienable 010000	*EXTLAPSER$_{Suff}$	BD-IDENT(stress) *Window* *FD*	*EXTLAPSE
> a. inàlienabíl$_L$ ity 02000 100		*	*
b. inálienabil$_L$ ity 02000 000	*!		****
c. inàlièabíl$_L$ ity 02020 100		* *!	

14. Others like *inalienability*, where WINDOW >> BD-IDENT(stress) >> *EXTLAPSE:

Suffix	*# of forms*	*Example from each suffix*		B_L		*Window size*
–ation	27	càннibalizátion	200010	cánnibalìze	1002	3σ
–ic	1	hèndecasyllábic[3]	200010	hèndecasýllable	200100	2σ
–ite	1	méteorolìte	20001	méteor	100	3σ
–ity	4	dìsciplinabílity	2000100	dísciplinable	10000	3σ
–ize	3	málleablìze	20001	málleable	10002	3σ

Upshot: With respect to BD-IDENT(stress) >> *EXTLAPSE, *–ity, –ic* (etc.) are just like *–ish* (etc.)

[2] With small caveats: a preference for #231 (vs. #021) (initial trochaic inversion) can sometimes, but not always, lead to BD-IDENT(stress) violations (see Kager 1989: 171, Pater 2000). *LAPSEL, in English, is also inviolable.

[3] There is some variability in the pronunciation of 'hendecasyllabic,' due to the variability of the pronunciation of its base, 'hendecasyllable'. 'Hendecasyllable' can be either 200100 or 020100; the *–ic* D either 200010 or 020010.

15. In addition to *EXTLAPSE, BD-IDENT(stress) dominates many other general high-ranked stress phonotactics. We sketch here a few of the arguments; see the appendix for further justification.

a. BD-IDENT(stress) >> *CLASH, *LIGHTCLASH[4]: *colòssálity* ≻ **còlossálity* (cf. *colóssal*)

colossal + ity B_L: colóssal 010	BD-IDENT(stress)	*CLASH *LIGHTCLASH
> a. colòssál$_L$ ity 02100		*(C), *(LC)
b. còlossál$_L$ ity 20100	*!*	

Other examples:
- *–ize*: *hòtélize* 213, *102 cf. B_L *hòtél* 21
- *–ic*: *dèltáic* 210, *010 cf. B_L *délta* 10
- *–ee*: *expèllée* 021, *201 cf. B_L *expél* 01

b. BD-IDENT(stress) >> WSP: *házardous* ≻ **hazárdous* (cf. *házard*)

hazard + ous B_L: házard 10	BD-IDENT(stress)	WSP
> a. házard$_L$ ous 100		*
b. hazárd$_L$ ous 010	*!*	

Other examples:
- *–al$_{Adj}$*: *súffixal* 100, *010 cf. B_L *súffix* 10 (13/18 speakers consulted)
- *–able*: *challengable* 1000, *0100 cf. B_L *challenge* 10
- *–ary: légendàry* 1020, *0100 cf. B_L *légend* 10

c. BD-IDENT(stress) >> EXTNONFIN: *forbíddal* ≻ **fórbiddal* (cf. *forbíd*)

forbid + al B_L: forbíd 01	BD-IDENT(stress)	EXTNONFIN
> a. forbídd$_L$ al 010		*
b. fórbidd$_L$ al 100	*!*	

Other examples:
- *al$_{Adj}$*: *commíttal* 010, *100 cf. B_L *commít* 01
- *al$_{Adj}$*: *dismíssal* 010, *100 cf. B_L *dismíss* 01
- *al$_{Adj}$*: *referral* 010, *100 cf. B_L *refér* 01

16. **Comparison with alternatives:** We reconstruct the predictions of the currently standard analysis of English Level 1 stress (Liberman and Prince 1977, Kiparsky 1982, Halle and Vergnaud 1987; Pater 2000) for the data just examined. These predictions are based on the following ideas:

- Level 1 forms exhaustively parsed into feet (= *EXTLAPSE, LAPSE undominated beyond R edge).
- Feet are q-sensitive on right edge (= WSP satisfied); can be q-insensitive elsewhere.
- Monomoraic feet are avoided (FTBIN ≈ *LIGHTCLASH generally satisfied)

[4] Constraint definitions provided in the appendix.

In translating rule-based LPM to OT analyses, and foot-based to foot-free constraints some of these predictions are harder to reconstruct for a clear comparison. We cite specific authors whose statements or rankings speak directly to the predictions listed below:

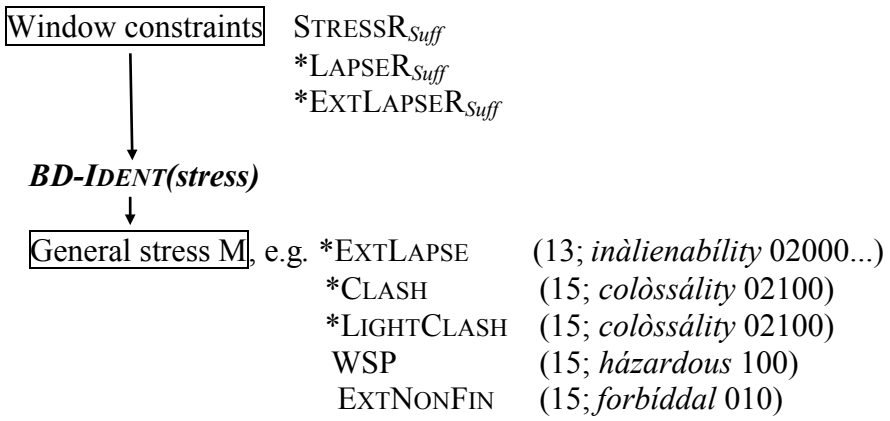
This analysis	*Predicts*	*Standard alternative*	*Predicts*
BDIDENT>>EXTLAPSE	inàlienabílity 02000100	exhaustive bin. parse (Kiparsky 1982): PARSEσ >> BDIDSTRESS	*inàlièn abílity 0(20)(20)(10)
BDIDENT>>*LIGHTCLASH	colòssálity 02100	FTBIN >>BDIDSTRESS (Pater 2000)	*còlossálity (20)(10)0
BDIDENT>>WSP BDIDENT >> EXTNONFIN	házardous 100 forbíddal 010	FTBIN, NONFIN >> ALIGNR >> BDIDSTRESS (Benua 1997)	*hazárdous 0(1)0 *fórbiddal (10)0

What we find significant is that, in our analysis, BD-IDENT(stress) systematically outranks all general M constraints (caveat in fn. 2 aside), even at "Level 1".

17. The picture, so far:

    a. Window constraints >> BD-IDENT(stress). This ensures that, for certain classes of suffixal derivatives, stress falls close to the right edge.

    b. For all suffixal families, BD-IDENT(stress) >> all general stress M discussed here.



3 *Remote bases*

18. Having established the general BD-IDENT >> M ranking for stress in derived words, we show that M-improvements are possible by appealing to accentually preferable B_Rs.

19. A preliminary question: what criteria must a word satisfy to count as a remote base? Two things:

    a. The B_R is **semantically related** to its potential D: this excludes pairs like *ínvalid, inválidate*.

    b. The B_R is **more frequent** than the D (preliminary indications: large differences in Google hits between the B_R and the D). This condition is based on the assumption that productive Ds have low token frequency: any B by reference to which the D is formed *must be known to the speaker* to exert an influence, hence is likely to have higher frequency than its D.

20. Example: the lexical family of *atomicity*, with B_L *atomic*:

a. *atomicity* is semantically related to many other words, other than its B_L: *atom, atomechanics, atomical, atomician, atomicism, atomism, atomist, atomistic, atomization, atomize, atomless....*

b. Of these, however, only a subset are more frequent than *atomicity* (655k Google hits[5]):

   i. More frequent: *atom* (425m), *atomician* (2.7m), *atomization* (686k), *atomize* (967k)

   ii. Less frequent: *atomechanics* (68k), *atomical* (181k), *atomicism* (87k), *atomism* (477k), *atomist* (157k), *atomistic* (574k), *atomization* (686k), *atomize* (967k), *atomless* (51k).

c. What this means: of the many semantically related forms to *atomicity*, only four of them (*atom*, *atomize, atomician, atomization)* satisfy both criteria in (2), and qualify as B_Rs for *atomicity*. All have accentual and segmental properties similar to those of *átom* and different from *atóm–*.

21. Illustration of *atomicity*'s lexical family (arrows indicate an asymmetrical BD relation):

D: *atomicity*	B_Rs:	Other forms (not potential B_Rs):
	atom, atomician	*atomechanics, atomical, atomicism, atomist,*
B_L: *atomic*	*atomize, atomization*	*atomistic, atomization, atomize, atomless...*

Comments

- The B_L, *atomic*, enjoys a privileged status: *atomicity* prefers to be faithful to it.
- B_Rs (*atom, atomician, atomize, atomization*), like the B_L, are **independently and productively derived**. They are more frequent than their co-derivative *atomicity*, and can influence its form.

B_R: *átom*

atom	NONFIN
> a. átom	
b. atóm	*!

B_L: *atómic*

atom + ic B_L: átom	*LAPSER$_{Suff}$
> a. atómic	
b. átomic	*!

B_R: *átomìze*

atom + ize B_L: átom	WSP
> a. átomìze	
b. átomize	*!

D: *àtomícity*

atom + ic + ity B_L: atómic B_R: átom, àtomíze (...)	BD-IDENT	CORRB_L
> a. àtom$_R$ icity		*
b. àtomíc$_L$ ity	*!*	

- The others (*atomechanics*, etc.) are not B_Rs of *atomicity*, because they are less frequent than it.

[5] These are preliminary numbers – the number of Google hits is not the best way to assess lexical frequency, as the number of hits is not directly tied to the number of times the word is used (for example: very uncommon words may receive inflated counts from dictionaries, translation websites, etc.).

22. We illustrate the use of B_Rs with *–ity* Ds. Relevant M constraints:

- *CLASH: a * for each sequence of two adjacent stresses.
- *LIGHTCLASH: a * for each sequence of two adjacent stresses, where the first syllable is light.

23. We consider first *–ity* derivatives that have the following properties:

- their local bases end in 10, causing *EXTLAPSER$_{Suff}$, *CLASH, and BD-ID(stress) to conflict.
- there are no accentually preferable B_Rs available.

The result: violations of all varieties of *CLASH[6].

**(LIGHT)CLASH violations 24 total, including:*		B_L	
àlcoh[ɑ̀]lícity	203100	àlcohólic	2010
c[ɛ̀]llárity	2100	céllar	10
col[ɑ̀]ssálity	02100	colóssal	010
sph[ì]rícity	2100	sphéric	10
hòriz[ɑ̀]ntálity	203100	hòrizóntal	2010

24. Now: a different set of *–ity* derivatives with the following properties:

- their local bases end in 10, causing *EXTLAPSER$_{Suff}$, *CLASH, and BD-ID(stress) to conflict.
- there are accentually preferable B_Rs available.

In these cases, the derivative systematically resembles the stress pattern of its remote base.

Derivative matches B_R 23 total, including:		B_R		B_L	
àtomícity	20100	átom	10	atómic	010
apòstolícity	020100	apóstle	010	àpostólic	2010
càlorícity	20100	cálorie	100	calóric	010
mùtagenícity	200100	mútagen	100	mùtagénic	2010
hìstorícity	20100	hístory	100	histórico	100

25. The appeal to B_Rs, in these cases, is a *clash-avoidance strategy*: *CLASH >> CORRB_L.

Full tableau, to demonstrate, for *apostolicity*:

apostle + ic + ity B_L: àpostólic 2010 B_R: apóstle 010	*EXTLAPSER$_{Suff}$	BD-IDENT(stress)	*CLASH	CORRB_L
> a. apòstol$_R$ ícity 020100				*
b. àpostòlíc$_L$ ity 203100		*!	*	
c. àpostolíc$_L$ ity 200100		*!*		
d. àpostólic$_L$ ity 201000	*!			

[6] What about *atómic*, *oríginal*, (etc.) with no clash? We assume that these and other frequent clash-less forms are lexicalized.

Discussion of candidates:

- (25a): *apòstol$_R$ ícity* corresponds with *apostle*$_R$, and incurs only a violation of CORRB$_L$.
- (25b): *àpostòlíc$_L$ ity* corresponds with *apostolic*$_L$. It incurs violations of BD-IDENT(stress) and *CLASH, as a stress is added to *–lí–* to satisfy the window constraint (*EXTLAPSER$_{Suff}$).
- (25c): *àpostolíc$_L$ ity* corresponds with *apostolic*$_L$. It incurs two violations of BD-IDENT(stress), as a stress is added to *–lí–* (to satisfy *EXTLAPSER$_{Suff}$) and deleted from *–sto–* (to satisfy *CLASH).
- (25d): *àpostólic$_L$ ity* corresponds with *apostolic*$_L$. It incurs a violation of *EXTLAPSER$_{Suff}$, as none of the final three syllables are stressed.

26. Other M constraints >> CORRB$_L$. Two arguments sketched below; more in the appendix.

a. *EXTLAPSE >> CORRB$_L$: *demónstrable* (B$_R$: *demónstrative*) ≻ **démonstrable* (B$_L$: *démonstràte*)

demonstrate + able B$_L$: démonstràte 102 B$_R$: demónstrative 0100	*EXTLAPSE	CORRB$_L$
> a. demónstr$_R$ able 0100		*
b. démonstr$_L$ able 1000	*!	

Cf. *challengable*, from *challenge* (with a violation of *EXTLAPSE).

b. WSP >> CORRB$_L$: *trìúmphal* (B$_R$: *trìúmphant*) ≻ **tríumphal* (B$_L$: *tríumph*)

triumph + al B$_L$: tríumph 10 B$_R$: trìúmphant 210	WSP	CORRB$_L$
> a. trìúmph$_R$ al 210		*
b. tríumph$_L$ al 100	*!	

Cf. *súffixal*, from *súffix* (with a violation of WSP; 13/18 speakers consulted).

27. Preliminary analysis, window constraints aside:

BD-IDENT(stress)
↓
General stress M, e.g. *CLASH (5: *apòstol$_R$ ícity*)
*LIGHTCLASH (see above)
*EXTLAPSE (6a: *demónstr$_R$ able)*
WSP (6b: *trìúmph$_R$ al*)
↓
CORRB$_L$

28. The behavior of some suffixes (the Germanic suffixes) suggests that the ability to correspond with remote bases varies parametrically across suffixes: *–ish* and *–ing*, for example, do not allow any M-improvements by appealing to B$_R$s.

See the appendix for details and a minor revision to the analysis, to account for this.

4 Further evidence for suffix-specific constraints

29. The proposed analysis claims that the difference between "Level 1" and "Level 2" derivatives can be attributed entirely to suffix-specific constraints. Here we provide more evidence that these constraints are independently necessary (see also Plag 1999, Raffelsiefen 2004).

30. We identify two types of suffix-specific constraints:

- (31): those that dominate BD-IDENT(stress).
- (32): those that dominate CORRB$_L$, providing more evidence for remote bases.

31. **M$_{Suff}$ >> BD-IDENT(stress)**: *LIGHTCLASH in *–ician* derivatives

While heavy clashes are allowed in *–ician* derivatives, light clashes are not.

Heavy clashes allowed *8 total, including:*	*Light clashes banned* *16 total, including:*
sỳntàctícian 2310 (cf. sýntàx 12, syntáctic 010)	còsmetícian 2010, *2310 (cf. còsmétic 210)
semàntícian 0210 (cf. semántic 010)	logícian 010, *210 (cf. lógic 10)
pàradòxícian 20310 (cf. páradòx 102)	magícian 010, *210 (cf. mágic 10)

How to account for this: *LIGHTCLASH$_{–ician}$ >> BD-IDENT(stress) >> *LIGHTCLASH.

32. **BD-IDENT(stress) >> M$_{Suff}$ >> CORRB$_L$**: *LAPSE in Latinate derivatives

a. Some Latinate suffixes (e.g. *–ite*) carry with them an additional restriction on lapsed sequences. This preference becomes clear when we examine forms without a morphological base.

–ite baseless forms
àmphíbolìte 2103, *1002
pèntácrinìte 2103, *1002
pỳrolúsìte 2013, *1002

Why *àmphìbolíte* (with a violation of *CLASH) and not *àmphibolíte* (cf. *Kàlamazóo*)? We claim that *all lapsed sequences are penalized in Latinate derivatives.*

b. BD-IDENT(stress) >> *LAPSE$_{Lat}$: violations of *LAPSE$_{Lat}$ allowed so that Ds can resemble Bs.

*–ite form violating *LAPSE$_{Lat}$* *18 total, including:*	B_L
crócodilìte[7] 1002	crócodìle 102
méteorìte 1002	méteor 100
méteorolìte 10002	méteor 100
gálleryìte 1002	gállery 100

[7] From the OED: H. Mason (1624) *New Art of Lying* ii. 35 This muddy Nylus so fertile of Crocodiles, I mean of this sophistique Crocodilites, whereby vnware men are ouer-reached and caught.

Tableau for *méteorite:*

meteor + ite B_L: méteor 100	BD-IDENT(stress)	*LAPSE$_{Lat}$
> a. méteorìte 1002		*
b. mèteórìte 2013	*	

c. *LAPSE$_{Lat}$ >> CORRB$_L$: correspondence with remote bases permitted to satisfy *LAPSE$_{Lat}$.

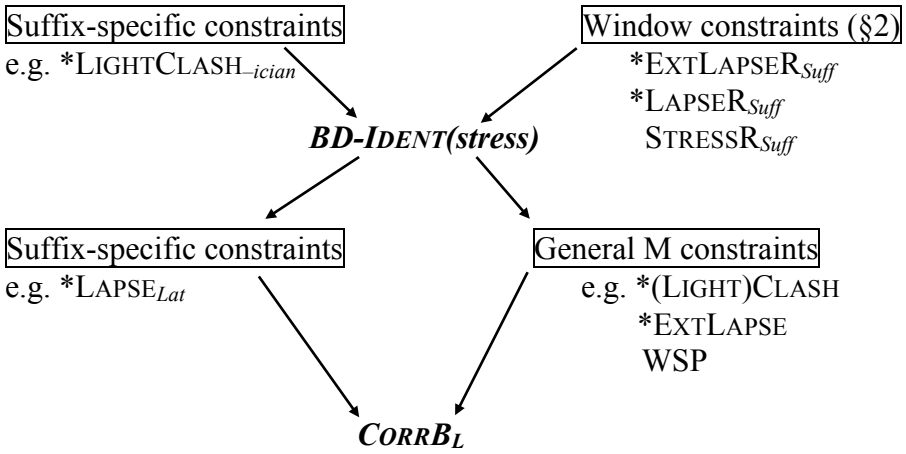
–ite derivative	B_R	B_L
metábolìte 0102	metábolism 01020	mètabólic 2010
hỳdròġeníte 2103	hỳdrógenàte 2103	hỳdrogen 100

Tableau for *metábolìte* (B_L identified as *metabolic* by the OED[8]):

metabolic + ite B_L: mètabólic 2010 B_R: metábolìsm 01020	*CLASH	*LAPSE$_{Lat}$	CORRB$_L$
> a. metábol$_R$ ìte 0102			*
b. métabol$_L$ ìte 1002		*!	
c. mètaból$_L$ ìte 2013	*!		

d. Examples of other suffixes governed by *LAPSE$_{Lat}$: *–ate, –ify, –ary*

33. Schema of the analysis (see the appendix for a more complete picture):



5 *Global support*

34. Question**:** does a global look at the English lexicon support the analysis proposed here?

35. Testing this: suffixal derivatives of each type listed in Marchand (1969). Derivative lists are from from Lehnert (1971) and the OED; judgments verified by the first author[9].

[8] We assume productive truncation here: metabolic + ite = metabol~~ic~~ite. Cf. Aronoff 1976.
[9] Derivative lists available upon request.

36. Two findings:

a. The number of derivatives that have *potential access* to an accentually preferable remote base ("B_R better") is small: 6% on average.

Suffix	*# of forms*	*B_R better*	*%*	*Suffix*	*# of forms*	*B_R better*	*%*
–a/ence	250	0	0%	*–ese*	108	1	1%
–a/ency	102	16	16%	*–esque*	16	1	6%
–a/ent	283	2	1%	*–ette*	39	0	0%
–acy	60	2	3%	*–ful*	87	0	0%
–age	108	0	0%	*–hood*	52	0	0%
–al	701	22	3%	*–ian*	126	7	6%
–an	30	0	0%	*–ic*	78	2	3%
–arian	23	3	13%	*–ician*	50	9	3%
–ary	190	10	5%	*–ify*	21	11	52%
–ate	166	8	5%	*–ite*	50	2	4%
–ation	158	32	20%	*–ity*	76	26	34%
–cy	22	0	0%	*–ive*	163	17	10%
–dom	52	0	0%	*–ize*	145	15	11%
–ee	90	22	24%	*–oid*	45	5	4%
–eer	42	0	0%	*–olatry*	15	0	0%
–ery	95	0	0%	*–ous*	55	2	4%

b. Of this class, an overwhelming majority resemble the accentual profile of their B_R ("B_R used").

Suffix	*B_R best*	*B_R used*	*%*	*Suffix*	*B_R best*	*B_R used*	*%*
–al	22	22	100%	*–ite*	2	2	100%
–arian	3	3	100%	*–oid*	5	5	100%
–ary	10	10 (1 var.)	100%	*–ize*	15	15	100%
–ate	8	8	100%	*–ous*	2	2	100%
–ation	32	32	100%	*–ive*	17	16 (2 var.)	94%
–ese	1	1	100%	*–ity*	26	23 (1 var.)	88%
–esque	1	1	100%	*–ee*	20	17 (5 var.)	85%
–ian	7	7	100%	*–a/ent*	2	1	50%
–ic	2	2	100%	*–acy*	2	1	50%
–ician	9	9 (2 var.)	100%	*–a/ency*	16	0	0%[10]
–ify	11	11	100%				

37. Certain remote bases are predictably inaccessible:

–ee: in 1/3 of the cases where an *–ee* D fails to resemble its B_R, the last stem consonant in the local B differs from the corresponding consonant in the accentually preferable remote base.

resi[n]ee 021, from resign ([n]) (*resi[g]nee 102, from resi[gn]ation 20-10)

[10] All 16 cases where an *–a/ency* derivative fails to resemble the stress of an accentually preferable remote base are similar: the *–a/ency* form (e.g. *excellency* 1000) resembles the an *–a/ence* form (e.g. *excellence* 100), even though an accentually preferable remote base (e.g. excel 01) exists. These deviations were explained by Chomsky & Halle (1968: p. 130 ff.) and Liberman & Prince (1977: 293-4) by claiming that <y> and other final sonorants are non-syllabic when stress is assigned.

Similarly on *–able* derivatives (Steriade 1999[11]):

contríbu[t]able 01000 from contríbu[t]e 010 (*contribú[ʃ]able 20100 from còntribú[ʃ]on 2010)

38. Other remote bases are variably accessible, a fact not always reflected in the OED:

–ity: in 2/3 of the cases where an *–ity* D fails to resemble its B_R, the judgments are variable and highly speaker-dependent, but not reflected as such in the OED.

ar[ə]maticity: 200100 (OED) ~ ar[oʊ]maticity 020100 (6/10 speakers consulted)
isot[ə]nicity 200100 (OED) ~ isot[oʊ]nicity 203100 (6/14 speakers consulted)

7 Brief discussion and summary

39. Brief discussion:

a. We've seen that whether or not a given derivative will correspond with a local base, or another related word, is a predicted by M >> CORRB_L.

b. **Question**: when a derivative corresponds with the stress of a B_R, can it correspond with its segmentals, too? Preliminary evidence suggests yes. See the appendix.

40. Summary:

a. **The English-specific point:** the F >> M ranking is consistent for all English suffixes, "Level 1" or "Level 2". Differences among them can be attributed to the activity of suffix-specific constraints, which are independently necessary.

b. **The broader point:**

- The phonological cycle is the formal mechanism ensuring that a derivative phonologically resembles members of its lexical family.
- The preference that derivatives resemble their local bases is *violable*.

c. Further evidence for this view comes from phenomena in other languages:

- Italian agentive nouns (Burzio 1998)
- French adjectival liaison (Steriade 1999)
- Romanian inflected nouns (Steriade 2008)
- Ukrainian and East Slavic nominal derivatives (Steriade & Yanovich, to appear)

[11] These cases have been analyzed as involving multiple correspondence between the D, the B_L and the B_R. The proposed generalization is that the D must correspond to the B_L in certain segmental properties: if the B_R mismatches these features, its stress becomes inaccessible as well. This proposal is consistent with the main point developed here: that B_Rs are used in computing the stress of most English derivatives. Its formalization is presented in Steriade (1999), but reconciling that with the evidence developed here is left to future work.

References

Albright, Adam (2006). Why eatees are not E.T's: Blocking of aspiration by output-output constraints. Paper presented at the 14th Manchester Phonology Meeting.

Aronoff, Mark (1976). *Word Formation in Generative Grammar.* Cambridge, MA: MIT Press.

Benua, Laura (1997). *Transderivational identity: phonological relations between words*. Ph.D. dissertation, UMass Amherst.

Bermúdez-Otero, Ricardo (2011). Cyclicity. In M. van Oostendorp, C.J. Ewen, E. Hume & K. Rice (eds.), *The Blackwell companion to phonology* **4.** 2019-2048.

Burzio, Luigi (1998). Multiple correspondence. *Lingua* **104.** 79-109.

Chomsky, Noam & Morris Halle (1968). *The sound pattern of English.* Cambridge, MA: MIT Press.

Collie, Sarah (2007). *English stress preservation and Stratal Optimality Theory.* Ph.D. dissertation, The University of Edinburgh.

Collie, Sarah (2008). English stress preservation: the case for 'fake cyclicity'. *English Language and Linguistics* **12.** 505-532.

Embick, David (2010*). Localism versus Globalism in Morphology and Phonology*. Cambridge: MIT Press.

Halle, Morris & Jean-Roger Vergnaud (1987). *An essay on stress.* Cambridge, MA: MIT Press.

Kager, René (1989). A metrical theory of stress and destressing in English and Dutch. Berlin: Walter de Gruyter.

Kager, René (1999). *Optimality theory*. Cambridge: Cambridge University Press.

Kager, René (2012). Stress in windows: language typology and factorial typology. *Lingua* **122.** 1454-1493.

Kiparsky, Paul (1982). Lexical Morphology and Phonology. In *Linguistics in the morning calm: selected papers from SICOL-1981*, pp. 3-92. Seoul: Hanshin Publishing Company.

Kiparsky, Paul (2001). Opacity and cyclicity. *Linguistic Review* **17.** 351-366.

Lehnert, Martin (1971). *Reverse dictionary of present-day English.* VEB Verlag Enzyklopädie.

Liberman, Mark & Alan Prince (1977). On Stress and Linguistic Rhythm. *LI* **8.** 249-336.

Marchand, Hans (1969). The Categories and Types of Present-Day English Word-Formation. CH Beck'sche Verlagsbuchhandlung.

Pater, Joe (2000). Non-uniformity in English secondary stress: the role of ranked and lexically specific constraints. *Phonology* **17.** 237-274.

Plag, Ingo (1999). *Morphological Productivity: Structural Constraints in English Derivation.* Berlin: Mouton de Gruyter.

Raffelsiefen, Renate (2004). Absolute ill-formedness and other morphophonological effects. *Phonology* **21.** 91-142.

Siegel, Dorothy Carla (1974). *Topics in English morphology.* Ph.D. dissertation, MIT.

Spencer, Andrew (1991). *Morphological Theory: An Introduction to Word Structure in Generative Grammar*. Cambridge: Cambridge University Press.

Steriade, Donca (1999). Lexical conservatism in French adjectival liaison. In M. Authier, B. Bullock & L. Reed (eds.), *Proceedings of the 25th Linguistic Colloquium on Romance Languages*. Amsterdam: John Benjamins.

Steriade, Donca (2008). A pseudo-cyclic effect in Romanian morphophonology. In Asaf Bachrach & Andrew Nevins (eds.), *Inflectional Identity*, pp. 313-359. Oxford: Oxford University Press.

Steriade, Donca & Igor Yanovich (to appear). Accentual allomorphs in East Slavic: an argument for inflection dependence. In Eulàlia Bonet, Maria-Rosa Lloret & Joan Mascaró (eds.), *Understanding Allomorphy*. Sheffield, UK: Equinox Press.

Appendix 1

Constraint definitions

General M constraints, and their suffix-specific versions:

- *CLASH: assign one * for each sequence of adjacent stressed syllables.
  - *CLASH$_{\text{–}ation}$: assign one * for each sequence of adjacent stressed syllables for *–ation* forms.
- *LIGHTCLASH: assign one * for each sequence of adjacent stressed syllables, where the first syllable in the sequence is light (e.g. [bɪ], [bɛ]).
- *EXTLAPSE: assign one * for each sequence of stressless syllables 00 such that neither is adjacent to a stress or a word boundary.
  - *LAPSE$_{Lat}$: assign one * for each sequence of two unstressed syllables in Latinate derivatives.
- WSP: assign one * if a heavy syllable is not stressed.
- EXTNONFIN: assign one * if stress falls on one of the final two syllables.

Window constraints:

- *EXTLAPSER$_{Suff}$: assign one * if none of the final three syllables are stressed, for certain classes of suffixal derivatives (e.g. *–ity, –al*).
- *LAPSER$_{Suff}$: assign one * if neither of the final two syllables are stressed, for certain classes of suffixal derivatives (e.g. *–ic, –ation*).
- STRESSR$_{Suff}$: assign one * if the rightmost syllable is stressless, for certain classes of suffixal derivatives (e.g. *–ee, –esque*).

CORRB and BD-IDENT constraints:

- CORRB$_{L}$: assign a * for each derivative whose stem does not correspond with its B$_{L}$.
- BD-IDENT(stress): assign one * for each pair of correspondent syllables differing in stress.
  - Evaluated relative to the base that the derivative stands in correspondence with.

Appendix 2

Evidence for various ranking arguments, and a more complete analysis

BD-IDENT >> M

1. BD-IDENT(stress) >> *EXTLAPSE: see handout (§2) for evidence from a number of suffixes.

   See also Steriade (1999) for documentation of *EXTLAPSE violations in *–able* Ds.

2. BD-IDENT(stress) >> *CLASH, *LIGHTCLASH

   a. *–ity*: see handout (§3) for justification.

   b. *–ize:* 18 forms, including:

      sphéròidìze 132 cf. B_L *sphéròid* 12
      àntiséptìze 2013 cf. B_L *àntiséptic* 2010
      hòtélìze 213 cf. B_L *hòtél* 21

   c. *–ian*: 16 forms, including:

      lìbrárian 2100 cf. B_L *líbrary* 100
      vùlgárian 2100 cf. B_L *vúlgar* 10
      vìcárian 2100 cf. B_L *vícar* 10

   d. –ic: 30 forms, including:

      dèltáic 210 cf. B_L *délta* 10
      ìrídic 210 cf. B_L *irídium* 0100
      ìámbic 210 cf. B_L *íàmb* 12

   e. –ician (NB heavy clashes only): 7 forms, including:

      tàctícian 210 cf. B_L *táctic* 10
      semàntícian 0210 cf. B_L *semántic* 010
      pàradòxícian 20210 cf. B_L *pàradóxic* 2010

   f. *–ee:* 35 forms, including:

      expèllée 021 cf. B_L *expél* 01
      abùsée 021 cf. B_L *abúse* 01
      enròllée 021 cf. B_L *enróll* 01

   g. –al: many, including:

      spònsórial 2100 cf. B_L *spónsor 10*
      clìéntal 210 cf. B_L *clíent* 10
      èdíctal 210 cf. B_L *édìct* 12

3. BD-IDENT(stress) >> WSP

   a. *–al$_{Adj}$*: 4 forms, with speaker-dependent variation:

préfixal 100 (12/18 speakers)	cf. B_L *préfix* 10
súffixal 100 (13/18)	cf. B_L *súffix* 10
áffixal 100 (9/18)	cf. B_L *áffix* 10
áutumnal 100 (3/18)	cf. B_L *áutumn* 10

   b. *–ous*: 8 forms, including:

cávernous 100	cf. B_L *cávern* 10
chívalrous 100	cf. B_L *chívalry* 100
házardous 100	cf. B_L *házard* 10
légendous 100	cf. B_L *légend* 10

   c. *–able:* see Steriade (1999) for data. Some examples:

admínistrable 01000	cf. B_L *admínister* 0100
bállastable 1000	cf. B_L *bállast* 10
chállengable 1000	cf. B_L *chállenge* 10

   d. *–ary:* at least 3:

vòluntáry 2010	cf. B_L *vòluntéer* 102
lègendáry 2010	cf. B_L *légend* 10
sècondáry 2010	cf. B_L *sécond* 10

4. BD-IDENT(stress) >> EXTNONFIN

   a. *–al:* 5 forms, including:

forbíddal 010	cf. B_L *forbíd* 01
commíttal 010	cf. B_L *commít* 01
dismíssal 010	cf. B_L *dismíss* 01

5. BD-IDENT(stress) >> STRESSL: many suffixes. This ranking argument is generally acknowledged and accepted in recent literature (e.g. Burzio 1994, Benua 1997, Collie 2007, 2008). Two examples:

   a. *–arian:* 10 forms, including:

eqùalitárian 020100	cf. B_L *équal* 10
necèssitárian 020100	cf. B_L *necéssity* 0100

   b. *–ity:* many forms, including:

orìginálity 020100	cf. B_L *oríginal* 0100
predìctabílity 020100	cf. B_L *predíctable* 0100

M >> CORRB_L

1. *EXTLAPSE >> CORRB_L

   a. *–able*: see Steriade (1999) for extensive justification and discussion. Some examples:

demónstrable 0100	cf. B_L *démonstràte* 102, B_R *demónstrative* 0100
obfúscable 0100	cf. B_L *óbfuscàte*, B_R *obfúscatòry* 01020
illústrable 0100	cf. B_L *íllustràte*, B_R *illústrative* 0100

   b. *–ate*: 1 form: *pròfessóriate* 20100, cf. B_L *profféssor* 010, B_R *pròfessórial* 20100

   b. *–ous*: 1 form: *àbdóminous* 2100, cf. B_L *ábdomen* 100/102, B_R *àbdóminal* 2100

   c. *–oid*: 1 form: *càrtiláginòid* 20103, cf. B_L *cártilage* 100, B_R *càrtiláginous* 20100

2. *(LIGHT)CLASH >> CORRB_L

   a. *–ize*: 14 forms, including:

Jápanìze 102	cf. B_L *Japán* 01, B_R *Jàpanése* 201
bureáutìze 0102	cf. B_L *búreaucràt* 102, B_R *bureáucracy* 0100
àttitúdinìze 20102	cf. B_L *áttitùde* 102, B_R *àttitúdinal* 20100

   b. *–arian*: 3 forms:

àttitudinárian 2020100	cf. B_L *áttitùde* 102, B_R *àttitúdinal* 20100
plàtitùdinárian 2020100	cf. B_L *plátitùde* 102, B_R *plàtitùdinal* 20100
àltitùdinárian 2020100	cf. B_L *áltitùde* 102, B_R *àltitúdinal* 20100

   c. *–ician*: 7 forms, including:

phònetícian 2010	cf. B_L *phonétic* 010, B_Rs *phóne* 1, *phónic* 10
mèchanícian 2010	cf. B_L *mechánic* 010, B_R *méchanìsm* 2010
ecònometrícian 020010	cf. B_L *ecònométric* 02010, B_R *ecónomy* 0100

   d. *–ee*: 17 forms, including:

cònferée 201	cf. B_L *confér* 01, B_R *cónference* 1(0)0
cònsultée 201	cf. B_L *consúlt* 01, B_R *cònsultátion* 2010
prèsentée 201	cf. B_L *presént* 01, B_R *prèsentátion* 2010

   e. *–ation*: 32 forms, including:

itàlicizátion 020010	cf. B_L *itálicìze* 0102, B_R *itálic* 010
cànnibalizátion 200010	cf. B_L *cánnibalìze* 1002, B_R *cánnibal* 100
sèrializátion 200010	cf. B_L *sérialìze* 1002, B_R *sérial* 100

3. WSP >> CORRB_L

   a. *–al:* 5 forms, including:

dìamétral 2010	cf. B_L *diámeter* 0100, B_R *dìamétric* 2010
hèxamétral 2010	cf. B_L *hexámeter* 0100, B_R *hèxamétric* 2010
pàramétral 2010	cf. B_L *parámeter* 0100, B_R *pàramétric* 2010

4. EXTNONFIN >> CORRB_L

   a. *–ous*: *arómatous* 0100, cf. B_L *àromátic* 2010, B_R *aróma*

   b. *–al:* 19 forms, including:

dèmoníacal 20100	cf. B_L *démon* 100, B_R *demóniac* 0100
pròsodíacal 20100	cf. B_L *prosódiàc* 0102, B_R *prósody* 100
mercúrial 0100	cf. B_L *mércury* 100, B_R *mercúriate* 0100

CORRB_L >> M

1. CORRB_L >> STRESSL

   a. *–ity:* the well-known case of *orìgináli*ty (B_L *oríginal* 0100, B_R *órigin* 100), and others.

   b. *–arian*: 4 forms, including:

eqùalitárian 020100	cf. B_L *equálity* 0100, B_R *équal* 10
necèssitárian 020100	cf. B_L *necéssity* 0100, B_R *néed* 1
humànitárian 020100	cf. B_L *humánity* 0100, B_R *húman* 10

CorrB_{LGer} >> M

Native Germanic suffixes prohibit correspondence with B_Rs, whatever the consequences. Some examples from *–ish* and *–ing* below, where *EXTLAPSE is violated despite there being B_Rs available.

1. *–ish*: many forms, including:

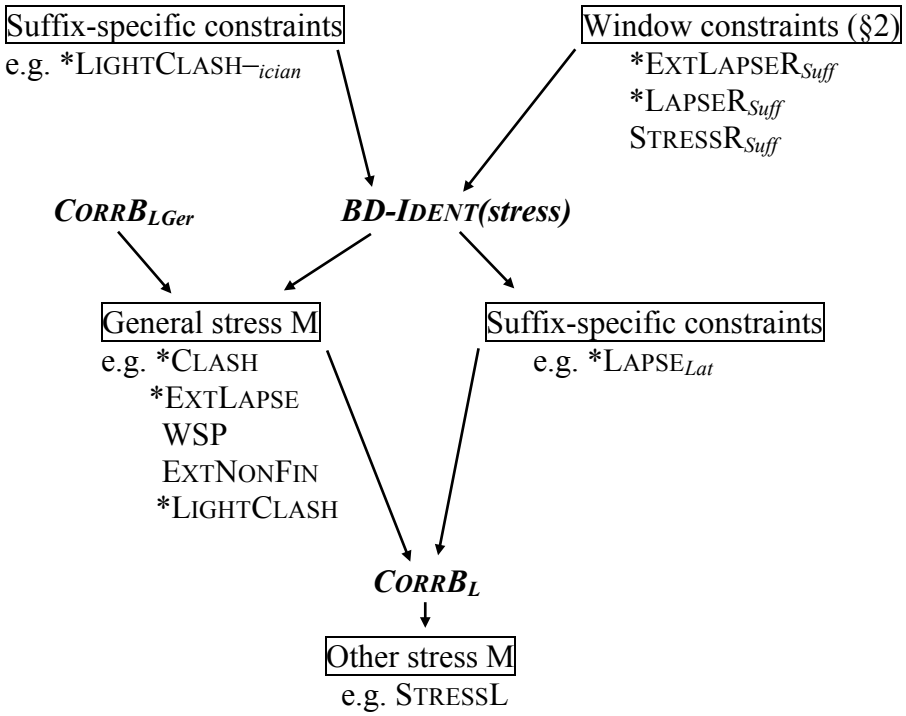
órchestra-ish 1000	cf. B_L *órchestra* 100, B_R *orchéstral* 010
éxcellent-ish 1000	cf. B_L *éxcellent* 100, B_R *excél* 01
rémedy-ish 1000	cf. B_L *rémedy* 100, B_R *remédial* 0100

2. *–ing*: many forms, including:

rémedying 1000	cf. B_L *rémedy* 100, B_R *remédial* 0100
órnamenting 1000	cf. B_L *órnament* 100, B_R *òrnaméntal* 2010
présidenting 1000	cf. B_L *président* 100, B_R *prèsidéntial* 2010

The contrast between Latinate (*–ity* and *–ive*) and Germanic (*–ing* and *–ish*) suffixes necessitates a revision: a version of CORRB_L, indexed to Germanic derivatives (CORRB_{LGer}), dominates all M.

Total ranking hierarchy



Appendix 3

Segmental identity to remote bases

1. Recall our analysis of [æ̀]tomícity: the [æ] appears in the doubly-suffixed form because *atomicity* stands in correspondence with its remote base, [ǽ]tom. (Simplified tableau below.)

[æ]tom + ic + ity B_L: [ə]tómic B_R: [æ]tom	*EXTLAPSER$_{-ity}$	BD-IDENT(stress)	*CLASH	CORRB_L
> a. [æ̀]tom$_R$ ícity				*
b. [ə]tómic$_L$ ity	*!			
c. [ə]tòmíc$_L$ ity		*!	*	
d. [æ̀]tomíc$_L$ ity		*!*		

2. Our approach can be compared to Benua (1997), where the appearance of [æ] would be attributed to a high-ranked IO-IDENT constraint preserving the vowel quality of the input. Both analyses are capable of accounting for *atomicity*, but globally, the two make distinct predictions.

   a. Benua: multiply suffixed derivatives can be faithful to only those properties of remote bases *that are present in the input*.

   b. This paper: multiply suffixed derivatives can be faithful to *any* property of a remote base, even noncontrastive properties that are not guaranteed to be part of the input.

3. A testing ground for the two hypotheses: the behavior of *–ee* derivatives.

   a. Albright (2006): the release quality of the pre-*ee* stop is linked to the release quality of the *–ee* derivative's base in isolation

      i. NT clusters are likely released, thus paint yields pàin[t^h]ée.
      ii. Final singleton Ts are likely unreleased or glottalized, thus eat yields èa[t]ée (*èa[t^h]ée).

   b. Albright analyzes this correspondence in release quality as an effect of an OO identity constraint; let's call it BD-IDENT(release). Tableau for *eatee* is below. (*TV́ = aspirate before stresses.)

eat + ee B_L: éa[t̃]	BD-IDENT[release]	*TV́
> a. èa[t]ée		*
b. èa[t^h]ée	*!	

4. We focus here on *–ee* derivatives of four verbs: permute, salute, permit, and sublet.

   a. All four verbs end in final singleton Ts; we expect all four *–ee* derivatives to lack aspiration.
   b. But this does not bear out: aspiration is preferred in *pèrmutée* and *sàlutée* (but not the others).

Aspiration strongly preferred		*No strong preference*
pèrmu[t^h]ée	(*pèrmu[t]ée)	$^{??}$pèrmi[t]ée, $^{??/*}$pèrmi[t^h]ée
sàlu[t^h]ée	(*sàlu[t]ée)	$^{??}$sùble[t]ée, $^{??/*}$sùble[t^h]ée

5. The relevant difference between the *permutee*-type and the *permitee*-type –ee derivatives: *permute* and *salute* have related forms with aspirated stops (pèrmutá[t^h]ion, sàlu[t^h]átion). The others don't.[12]

   a. For *permute*: when the *–ee* derivative is faithful to the *–ation* form's stress, to avoid a violation of *CLASH, it is faithful to its segmentals as well. This includes the aspirated T.

permute + ee B_L: permú[t̄] B_R: pèrmu[t^h]átion	BD-IDENT[release]	*CLASH	*TV́[13]	CORRB_L
> a. pèrmu[t^h]$_R$ée				*
b. pèrmu[t]$_R$ée	*!		*	*
c. permù[t]$_L$ée		*!	*	

   b. For *permit*: corresponding with *pérmit* (n.) allows satisfaction of *CLASH. Faithfulness to the segmentals of *pérmit* (n.) does not license aspiration, because the stop is likely unreleased.

permit + ee B_L: permí[t̄] (v.) B_R: pérmì[t̄] (n.)	BD-IDENT[release]	*CLASH	*TV́	CORRB_L
> a. pèrmì[t]$_R$ée			*	*
b. pèrmì[t^h]$_R$ée	*!			*
c. permì[t]$_R$ée		*!	*	

6. Implications: what kind of model can analyze segmental identity with *non-contrastive properties of remote bases*?

   a. Two things the model must allow: (i) access to remote bases, and (ii) surface correspondence.

   b. Our model allows both. Other current, competing models of cyclicity do not.

      i. Stratal OT (Kiparsky 2000, Bermúdez-Otero 2011): no allowed access to remote bases[14] or guaranteed access to surface properties of remote bases.
      ii. Standard OO correspondence (e.g. Benua 1997): no access to remote bases.
      iii. Cyclic locality (e.g. Embick 2010): no access to either remote bases or surface properties.

[12] Another difference between the two sets lies in the pre-/t/ vowel quality: in *permute* and *compute* the vowel is tense, and in *permit* and *sublet* the vowel is lax. In English, release of final singleton stops is more likely following tense vowels (see Kang 2003: 240-242), and this asymmetry shows up in –ee derivatives: in VT-ée contexts, the pre-ee stop is more likely to be aspirated if V is tense (A. Albright, p.c.). But it is doubtful that the vowel quality difference between *permute* and *permit* (for example) is responsible for the difference in aspiration: permù[ɾ??]ée, with the stress of *permúte*, cannot have an aspirated T.

[13] Evidence that *TV́ >> CORRB_L comes from forms like elìci[t^h]ée, (cf. elícit̄, elìci[t^h]átion), but this isn't crucial.

[14] Collie's (2007, 2008) proposal of "fake cyclicity", attributed to Bermúdez-Otero (in prep), allows that multi-layered forms can access the phonology of subconstituents. This proposal differs from ours. First, fake cyclicity does not allow co-derivatives to act as remote bases. Second, it predicts that B_Rs are accessed probabilistically, depending on their frequency; we claim that access to B_Rs is predicted by M >> CORRB_L.