Productive pseudo-cyclicity and its significance

Donca Steriade (MIT) and Juliet Stanton (NYU)

0. Introduction

In their contribution to Labphon 10, Pierrehumbert and Clopper (2010:116) defined the Labphon community as a "federation of scholars" gathering in a "free-trade zone" for the purpose of exchanging certain commodities: data, methods and ideas. These federated scholars share the purpose of documenting the outward forms taken by linguistic expressions and of unpacking their underlying cognitive and physical mechanisms. The commodities originate in distinct, if overlapping, communities of phoneticians, psycholinguists, sociolinguists, and phonologists. One benefit of participating is the chance to realize which assumptions of one's own home community require additional support and unpacking before addressing the rest of the federation.

It is a privilege to come to this free-trade zone on a year that marks the 30th anniversary of the publication of the 1st Labphon proceedings. We offer for your consideration a commodity of potential relevance to those who study the structure of the lexicon, to those curious about the balance between storage and computation in processing morphologically complex words, to those who refer to influences words have on each other's shape as *analogy*, and to those interested in the effect of extra-grammatical factors, like frequency, on phonological processes.

The commodity we bring to this gathering is a phenomenon called the pseudo-cycle (Steriade 2008). We define it in relation to the standard cycle in sections 1-2; we outline in sections 3-6 findings that distinguish the pseudo-cycle from a collection of unrelated word-on-word analogies or performance errors. In the process, we sketch ideas about the shared mechanisms that underlie the pseudo-cycle and the standard cycle, and try to connect these to the concerns of the Labphon interest groups mentioned above.

1. The cycle

The cycle (Chomsky, Halle and Lukoff 1956; Chomsky and Halle 1968) is a mechanism that transmits linguistic information from one form to a structurally related one. Relevant here is the cyclic transmission of phonological information. This phenomenon can be initially diagnosed as the unexpected and systematic similarity between certain pairs of related forms. The similarity is unexpected because it is not predicted by the grammar that models the sound pattern of simple words in the language.

As an initial example, any English prefixed verb like *refít* can be said to have its stress cyclically inherited from its unprefixed base, *fít*. The simple verb is assigned stress in isolation, and its prefixed form inherits this stress. The accentual similarity between such forms is unexpected. Simple verbs, like *prófit,* avoid final stress on lighter syllables. In an OT analysis (Prince and Smolensky 2004) a Markedness constraint NONFINALITY[1] models this avoidance, making explicit why this similarity between *fit* and *refit* is unexpected: the grammar of simple words would predict a dissimilar pair, *fít-*réfit*, that better satisfies NONFINALITY.

[1] Hung 1994. Definitions of all basic accentual M constraints used below are provided in Gordon 2002.

As the stress of *fit* is predictable, the pair *fit-refít* illustrates the basic fact that the cycle transmits rule-governed features, properties that may have been assigned by the grammar.

Our analysis of cyclic phenomena rests on OT's distinction between Markedness (M) and Faithfulness (F). Unexpected, systematic similarities like that in *fít-refít* are due to F constraints, globally known as Base-Derivative Correspondence (Benua 1997), and penalizing dissimilar pairs of morphologically related surface forms. These constraints provide the mechanism of cyclic transmission: they force related forms to resemble each other. A member of this class is BASE-DERIVATIVE IDENT STRESS (BD IDSTRESS). It assigns a penalty for every syllable whose [±stress] value in the Base differs from that of its counterpart in the Derivative. In our example, BD IDSTRESS ranks above NONFINALITY. Intuitively, this says that the similarity within this class of related pairs is more important than the preference for non-final stress. Shown below is the second step in a cyclic derivation of *refít*, the step in which the prefixed verb is assigned stress based on the simple verb's stress. Note that the predictably stressed *fít* is an input in this evaluation, a fixed reference term comparable to a lexically stored word.

(1)	B: fít		BD IDSTRESS	NONFINALITY
a.	réfit	10	*!	
☞b.	rèfít	21		*

The most common ranking schema characterizing cyclicity, and the only one relevant here, is BD CORR >> M >> IO CORR. The lower ranking, M >> IO CORR characterizes the basic sound regularity that's disrupted by the cycle, here the broad dispreference for final stress. The upper ranking is a mechanism of cyclic transmission, BD Corr >> M.

Cyclic transmission has two further properties relevant to this discussion. First, it is asymmetric, in the sense that cycles operating on outer constituents inherit information from inner cycles, and never the other way around. This asymmetry is known as *Base Priority* (Benua 1997). The empirical evidence for Base Priority is, partly, that the record of M violations is never better in a cyclic derivative than it is in its base. That's because, under the ranking BD CORR >> M >> IO CORR, the derivative succeeds in maintaining similarity to its base by violating M constraints, whenever M conflicts with higher BD CORR. In a more basic sense, that's also because the Base is treated in such analyses as a stored form, a member of the derived lexicon (Halle 1973), thus immune to change. The similarity to its derivatives can be obtained only by altering *their* shape.

Second, cyclic transmission is syntactically constrained. In general, the form that transmits the information is the exponent of a morpho-syntactic constituent immediately contained in the form that inherits it. We refer to this property as *C-Containment*, as in *the base, cycle n, is immediately contained in its derivative, cycle n+1. C-* stands for 'cyclic'. We formulate it as in (2), anticipating the need to conceive of it as a violable constraint:

> 2. C- CONTAINMENT: Assign a violation to any derivative D whose base is not exponent of an immediate constituent of D

C-CONTAINMENT is initially illustrated by the accentual difference between English left-branching compounds, as in [*háy fêver*] *treàtment* 13020 vs. right-branching *bùsiness* [*crédit* *câr̂d*] 2013; or, with longer compounds, left-branching [[[*láw degrêe*] *requîrement*] *chànges*] 103040020 vs. right-and-left branching [[*làw degrêe*] [*lánguage requîrement*]] 204100300[2]. The basic generalization is that the prominence relations internal to each compound are preserved under embedding, as mandated by an F constraint we call BD IDENT RELATIVE PROMINENCE. This penalizes any inversion of prominence (sw →ws, or vice versa) between a base in isolation, here an inner compound in isolation, and a version of it embedded in a derivative, i.e. in a larger compound. Subject to BD IDRELPROM, a form of non-finality obtains in compounds: main stress on the last word is generally disfavored, NONFINALITY(WORD)[3]. Bottom ranked, but still active, is the preference for rightmost main stress (MAINSTRESSRIGHT, MSR), a version of Liberman and Prince's (1977) "right is strong". C-Containment has two effects in compound stress. First, left-branching compounds like [*háy fêver*] *treàtment* show that the weak branch of a compound remains weak under embedding, at the cost of increased violations of MSR. For this result to obtain, the input to stress assignment on the larger compound must be its *immediate* subconstituents, here [*háy fèver*], with its already defined sw relation of relative prominence, not the minimal word-sized components, *háy* and *féver*. This is exactly what C-Containment demands. And, without C-Containment, candidate (b) wins under this ranking.

(3)	B1: háy fèver 120 B2: treátment 10	BD IDRELPROM	NONFINALITY (WORD)	MSR
a.	hày fêver treátment 23010		*!	
b.	hây féver treàtment 31020	*! (12→31)		*
☞c.	háy fêver treàtment 13020			**

Second, when a compound is right-prominent, e.g. in *Chrìstmas dínner*, that prominence is also preserved under embedding, at the expense of NONFINALITY(WORD), as in [*fàmily* [*Chrîstmas dínner*]] 231. Here too, C-Containment blocks a stress computation that would consider only individual words, missing the prominence relations that hold between subconstituents.

2. The pseudo-cycle

The *pseudo-cycle* (Stanton and Steriade 2018, Steriade 1999, 2008, Steriade and Yanovich 2015) is also a mechanism that transmits linguistic information, but, unlike the cycle, it violates C-Containment. The form transmitting information is not an immediate constituent in the target derivative. It is a lexically related form, sometimes a remote sub-unit in the target form, and sometimes a co-derivative of it. We discuss here only the English evidence for this phenomenon.

Consider first the pair of *-oid* derivatives in (4). They have identically stressed bases, and equivalent segmental composition, but their stress differs.

4. Cyclic and pseudocyclic derivatives (Ds).

páraffin-òid 1002	páraffin 100	-
gelátin-òid 0102	gélatin 100	gelátinous 0100

[2] Examples from Kösling et al. 2013 and Liberman and Prince 1977:256; see also Chomsky, Halle and Lukoff 1956. The empirical generalizations in the text reflect Kösling et al's findings.

[3] With pockets of systematic exceptions, now studied by Plag and colleagues: cf. Köslin et al. 2013

The difference in this pair is that *gelátinòid* has a co-derivative, *gelátinous*. By contrast, *páraffin* has no occasion to generate a comparable allomorph *paráffin-*, because it lacks any *-ous, -ity*, *-al* derivative that would require such a shift of stress. (*Pàraffín-ic* is known to some, but its stem has different stress.)

We claim that the speakers who know *gelatinous* can choose to transmit its stem allomorph *gelátin-* to *gelátinòid*, for the purpose of improving its stress. This transmission violates C-Containment, as the *-ous* form is not contained in *gelátinòid*, but it has other properties of the cycle: it transmits predictable information, and does so in a way that resembles Base Priority. We outline the support we have found for two of these claims: pseudo-cyclic effects are deviations from C-Containment *motivated by Markedness* and, like the cycle, they involve transmission of *predictable* properties. We emphasize the partial similarity between the pseudo-cycle and the cycle because we propose to offer a unified analysis of both.

Markedness first. The *gelátin-* stress in *gelátinous* is adopted in *gelátinòid* to avoid a stress lapse. The constraint *LAPSE is generally enforced in the inter-stress regions of Latinate words, as seen in *manípulàte, elíminàte, metábolìsm, apócalỳpse*. Deviations from it, as in *páraffinòid*, the form that lacks an alternative allomorph, indicate that BD IDSTRESS outranks *LAPSE (see (6)).

How is the stress of *gelátinous* transmitted to the structurally unrelated *-oid* form? Suppose C-Containment is cast as a violable constraint. Suppose further that derivatives can access a larger set of lexically related forms as possible bases. By lexically related we mean having the same root and identical or related lexical semantics. In this set, one form satisfies C-Containment, as the exponent of an immediate constituent of the derivative. Other potential bases don't, but the ranking M >> C-CONTAINMENT makes these other bases potentially relevant. Under this proposal, *gelátin-ous* is a possible source of information for *gelátin-oid,* i.e. a possible base, in addition to the base *gelatin*. M constraints will choose the better base among them.

The analyses below illustrate how the ranking BD IDSTRESS >> M >> C-CONTAINMENT differentiates the two *-oid* forms in (4). B^L, or the *local base*, refers to the base that satisfies C-Containment. A superscript ^{L} on the stem of a candidate indicates correspondence to this local base. Any other base is a *remote base*, B^R. The superscript ^{R} on the stem of a candidate indicates correspondence to a B^R. We simplify by listing just the minimal number of potential bases. The annotation $_{LAT}$ on $*LAPSE_{LAT}$ anticipates the need to distinguish Latinate from non-Latinate derivatives: we index the M constraints evaluating Latinate words.

(5)	B^L: gélatin 100 B^R: gelátin-ous 010-0	BD IDSTRESS	$*LAPSE_{LAT}$	C-CONTAINMENT
☞a.	gelátin^R-òid 0102			*
b.	gélatin^L-òid 1002		*	
c.	gelátin^L-òid 0102	*!*		

Any BD CORR constraint evaluates the match between a candidate and the base identified by the superscript: the result is that both top candidates in (5) satisfy BD IDSTRESS, but not (5.c).

The analysis of *páraffinoid* in (6) shows that what looks like cyclic transmission – the intact preservation of the predictable base stress in a derivative, at the expense of Markedness – can be generated in this analysis alongside pseudo-cyclic effects. Generalizing, any analysis in which C-Containment is dominated by M or F, or both, will produce unexpected similarities between B^Rs and target derivatives in some forms, and will continue to produce the similarities between B^Ls and target forms in others, exactly like the cycle. The differences between the forms like *páraffinoid* and *gelátinoid* will be predictable from the structure of their extended lexical families, not a stipulated difference between the grammars that generate them.

(6)	B^L: páraffin	100	BD IDSTRESS	*LAPSE$_{LAT}$	C-CONTAINMENT
☞a.	páraffin^L-òid	1002		*	
b.	paráffin^L -òid	0102	*!*		

A remark about predictability. While the existence of an *-ous* derivative of *gélatin* may be hard to predict – though this hasn't been attempted – the stress in *gelátinous* is predictable. Like most Latinate words, *-ous* adjectives don't tolerate extended lapses, 000 sequences, on their right edge. This plus the general English dispreference for stress on the last two syllables cause stress to advance from the initial in *gélatin* to the peninitial in *gelátinous*. In this sense, B^R *gelátinous* functions like a first cycle output: it transmits predictable properties. This distinguishes the pseudo-cycle from lexical selection among unpredictable allomorphs (Kager 1996).

The examples in (7) extend this discussion to a class of derivatives whose stress matches that of a constituent that's embedded inside the B^L. As in (4), we compare minimal pairs of derivatives. In each pair, B^Ls have identical stress. The derivatives have the same suffix, but different stress. Stresses deviating from the B^L are less marked and resemble, in each case, a related form, a B^R.

(7)	B_L stress		D stress like B_L		D stress ≠ B_L		B^R	
i.	*pàcific-átion*	20010			*pacífic-atòry*	010020	*pacífic*	010
ii.	*clàssific-átion*	20010	*clássific-atòry*	100020				
iii.	*àpostól-ic*	2010			*apòstol-íc-ity*	020100	*apóstle*	010
iv.	*àlcohól-ic*	2010	*àlcohòl-íc-ity*	203100				
v.	*refléx-iv-e*	010			*rèflêx-ív-ity*	23100	*réflèx*	12
vi.	*refléct-ive*	010	*reflèct-ív-ity*[4]	02100				

Each pair of Ds in (7) is such that, if they were both faithful to their B^L, they would both violate a certain M constraint: *EXTENDED LAPSE in (i-ii), *CLASH in (iii-iv), and the trigger of the word-internal Rhythm Rule (Kiparsky 1979; Hayes 1982) in (v-vi). But the internal structure of the B^Ls differs in each pair and this provides one member of the pair with a B^R that can optimize its stress. Thus, *pàcificátion* contains *pacífic* while *clássificàtion* only contains *clássify*: there is no **classífic*. The stress of *pacífic* improves that of *pacíficatòry:* it avoids an extended lapse. The stress of *clássificatòry* can't be remedied in similar ways. In (iii), B^R *apóstle* helps avoid a clash in *apòstolícity*. Minimally different *àlcohôlícity* can use *álcohòl*, but its stress can't avoid the

[4] The contrast between *rèflêxívity* and *reflèctívity* is, as most other data in this study, based on OED's transcriptions. Speakers we consulted can also neutralize it, accepting 23100 for *reflectivity*, an output of the Rhythm Rule. But they also report the predicted contrast.

clash with *-ícity*. In (vi) *réflèx* allows improved rhythm in *rèflêx-ív-ity*, while *reflèctívity* can't be improved, because trochaic **réflèct* is missing. The analysis of such cases is identical to (5-6):

(8)	B^L: pàcific-át-ion B^R: pacífic	20010 010	BD IDSTRESS	*EXTLAPSE$_{\text{LAT}}$	C-CONTAINMENT
☞a.	pacífic^R-at-òry	010020			*
b.	pácific^L-at-òry	100020		*!	
c.	pacífic^L-at-òry	010020	*!*		

(9)	B^L: clàssific-át-ion	20010	BD IDSTRESS	*EXTLAPSE$_{\text{LAT}}$	C-CONTAINMENT
☞a.	clássific^L-at-òry	10002		*	
b.	classífic^L-at-òry	010020	*!*		

A complete model of English cyclicity must reflect two deviations from the schema CORR BD >> M >> C-CONTAINMENT. First, only Latinate (roughly, 'Level 1') derivatives allow violations of C-CONTAINMENT. Non-Latinate derivatives like *párenting, dísciplining,* or *rémedying* violate M constraints conflicting with CORR BD, even when optimizing B^Rs, like *paréntal, dìsciplínary* or *remédial,* exist. B^R-based **parénting*, **dìsciplíning, *remédying* are impossible. By contrast, Latinate *parénticide*, *remédiable, dìsciplínable* improve their M score by using these B^Rs.

This split between derivatives that never optimize their M score at the expense of C-CONTAINMENT and those that do can be characterized in several ways. Here we use M constraints indexed to groups of affixes, as in Pater 2008[5]. Thus *EXTLAPSE is split into an indexed version *EXTLAPSE$_{\text{LAT}}$, which evaluates only Latinate derivatives, and a lower ranked general version. C-CONTAINMENT is outranked by *EXTLAPSE$_{\text{LAT}}$, as seen in (8), but outranks the general *EXTLAPSE, shown in (10). This split ensures that B^Rs are inert outside the Latinate set:

(10)	B^L: dísciplin B^R: dìsciplín-ary	100 20100	BD IDSTRESS	C-CONTAINMENT	*EXTLAPSE
☞a.	dísciplin^L-ing	1000			**
b.	dìsciplín^R-ing	2010		*!	

A further distinction obtains in the Latinate set, between derivatives that enforce certain M constraints at the expense of BD CORR, vs. all others. A set of affixes that include *-ian, -ity, -ous, -al* prohibit final extended lapses; *-ic, -id* generally require penult stress; *-ation* tends to prohibit clashes. These conditions can be imposed in violation of BD IDSTRESS. Thus, novel *-ian* and *-ic* adjectives like *Beckéttian* or *Titiánic* shift stress without benefit of any B^R. We take this as support for ranking above BD IDSTRESS a small set of indexed constraints like *EXTLAPSER$_{\text{ian}}$ (for the entire class composed of *-ity, -al, -ous -ian*), and *LAPSER$_{\text{IC}}$ (for the small *-ic* class).

Other Latinate derivatives, like *-able*, differ. They satisfy M constraints like *EXTLAPSE$_{\text{LAT}}$ only when a B^R supplies the better stress profile, as seen in (8) vs. (9). Absent an optimizing B^R, *-able* adjectives, like those in *-òry*, violate *EXTLAPSE$_{\text{LAT}}$: e.g. *jéttisonable, párodiable.*

---

[5] For a restrictive alternative to affix-indexed constraints, see Becker and Gouskova 2016, a model we have not yet experimented with. For ideas that pertain to the Latinate-Germanic split in English, see Pierrehumbert 2005.

These examples suggest an articulated hierarchy of M constraints, separated by accentual base faithfulness and the base preference, C-CONTAINMENT. The more restricted stress pattern of mono-morphemic forms is captured by ranking IO CORR at the very bottom, outranked by all active M constraints on stress. Different sub-rankings in this schema correspond to different classes of derivatives:

11. A ranking schema for varieties of English stress using affix-indexed constraints

M_1 >> CORR BD >> M_2 >> C-CONTAINMENT >> M_3 >> IO CORR
{*EXTLAPSER$_{IAN, \ldots}$} {*EXTLAPSE$_{LAT}$} {*EXTLAPSE$_{\ldots}$}

invariably stress-shifting pseudo-cyclic echt-cyclic monomorphemes

Latinate

There are rankings internal to M_2 and M_3, left out of this discussion for space reasons. We turn now to a summary of the empirical results that support this architecture.

3. A dictionary study of English pseudo-cyclicity

We have explored the lexical evidence for the use of B_Rs in American English, by extracting from the online Oxford English Dictionary (OED) derivatives with known bases, formed with 27 of the Latinate suffixes listed in Marchand 1969. The stress information is inferred from transcription details of the OED entries identified as US pronunciations. These details include stress marks and patterns of reduction and flapping. We consulted the OED recordings wherever available. Occasionally, OED's record was amended, based on evidence from Merriam Webster's online dictionary, from youglish.com and other sources. The resulting database records, for each derivative D, the stress of its stem, if this stress differs from that of its B^L; if this D has an accentually optimizing B^R (i.e. if D has some Base B, different from the B^L, such that, if D were faithful to B, not to B^L, it would better satisfy the Markedness hierarchy); and whether this better B^R, if it exists, is actually used as the correspondent of D's stem, i.e. if D's stem is more similar in stress to the B^R than to the B^L.

We considered only evidence provided by constraints belonging to the M_2 class in (11). This means, for instance, that we recorded accentual M and F violations for the *-ian* adjectives but not those relating to the enforcement of *EXTLAPSER$_{ian}$: this M constraint ranks above BD CORR in (11), and thus is uninformative on the hypothesis that M outranks C-CONTAINMENT.

With this exclusion, the database records all actual and potential violations of *EXTLAPSE$_{LAT}$, *LAPSE$_{LAT}$, *CLASH$_{LAT}$, and the trigger of the Rhythm Rule[6]. *Potential M violations* are those expected to surface if the derivative were faithful to its B_L. We are interested only in derivatives that are potential M violators in this sense. Thus, we are interested in *gelatinoid*, because the candidate faithful to B^L *gélatin* violates *LAPSE$_{LAT.}$ (This candidate is *gélatinoid*, an unattested

[6] We define the trigger of the word-internal Rhythm Rule as follows: "A stressed syllable is either a trough or a peak. Assign a * to each stressed syllable that is neither. Stress peak =$_{Def}$ a syllable with a higher grid column than adjacent syllables on either side. Stress trough=$_{Def}$ a syllable with a lower grid column than adjacent syllables on either side." This penalizes, for instance, 021, 321 and is satisfied by 201, 231

pronunciation.) We are not interested in forms like *alkaloid*, because the candidate faithful to its B^L *álkalì*, *álkalòid*, the attested form, is accentually optimal, in addition to being faithful.

What findings would support (11)? First, finding that some markedness constraint M, which is potentially violated in a derivative D, as defined above, is in fact satisfied in D when D possesses a B^R whose stress contour satisfies M. That's the case in the examples analyzed in (5) and (8) and in the odd-numbered rows of (7). (11) is also consistent with a finding that M, which is potentially violated in a D, is actually violated in D, when that D has no B^R whose stress contour satisfies M. That's the case of examples (6), (10) and of the even-numbered rows of (7).

What findings would contradict (11)? Finding that some D violates some M despite the existence of a B^R whose stem allomorph, if adopted in D, would satisfy M. (11) is also contradicted when D satisfies an M_2 constraint by violating BD CORR, i.e. by changing its stem stress in ways that don't match either the B^L or any B^R. Examples all of such cases follow. Evidence from *-ify* verbs (where the relevant member of M_2 is *LAPSE) and *-ee* nouns (for *CLASH) is presented below.

There are 42 *-ify* verbs in the OED which would violate *LAPSE if they remained faithful to their B^L. 37 among them have an optimizing B^R and all but one use it to avoid this potential lapse. Forms lacking a lapse-avoiding B^R preserve the lapse:

12.	Optimizing B^R	No optimizing B^R
*LAPSE satisfied (stress shifts R wrt B_L)	36 *Germán-ify*; B^L *Gérman,* B^R *Germán-ic*[7]	0
*LAPSE violated (stress same as B_L)	1 *tútor-ify*; B^L *tútor*; B^R *tutór-ial*	5 *Énglish-ify;* B^L *Énglish*)

Effect of a 'Better B^R' factor on the rightward shift of stress in X-*ify*, wrt B^L
$p < .001$, Fisher's Exact Test

The pattern of *-ee* nouns is comparable, if noisier. There are 101 *-ee* nouns expected to violate *CLASH if they remain faithful to their B^L, as in *emplóy, emplòy-ée*. All B^L-faithful forms would also violate the trigger of the Rhythm Rule, as the B^L in all these cases is a verb with final main stress. We report only clash resolution results. (To decide if the Rhythm Rule has applied in a derivative we need more data than the OED's transcriptions, which don't always distinguish 23 from 32 contours. For *-ee* nouns, this information is too sparse.)

13.	Optimizing B^R	No optimizing B^R
*CLASH score improved (stress shifted L wrt B_L)	12 (*òrdin-ée;* B^L *ordáin,* B^R *òrdinátion*)	1 (*pàrolée,* B^L *paróle*)
*CLASH violated (stress as in B_L)	13 (*provók-èe;* B^L *provóke,* B^R *pròvocátion*)	75 (*divòrcée,* B^L *divórce*)

Effect of a 'Better B^R' factor on the leftward shift of stress in X-*ee*, shift wrt B^L
$p < .001$ (Fisher's Exact Test)

[7] We take *Gérman* not *Germánic* to be the B^L of *germánify* in part because the texts supplied by the OED point to *germánify* as meaning 'make something German or relevant to Germans or Germany' and not 'make something Germanic'. This illustrates some factors we consider in deciding if a form is the B^L or the B^R of a derivative.

Overall, 12 derivative types (*-ify, -ee, -able, -ary, -ory, -ive, -ician, -ivity, -icity, -ism, -ite, -oid*) and the class of root compounds (e.g. *cerébroscòpe, cèrebróscopy; polárogràph* vs. *pálatogràph*), provided evidence that potential M violations are significantly more likely to be resolved in derivatives that can adopt the stress profile of a B^R that doesn't violate M than otherwise. For 9 other types there is insufficient data to assess a B^R effect, most frequently because too few derivatives have optimizing B^Rs: *-ation, -eer, -ese, -esque, -ess, -ette, -ist, -ite* and *-ous* belong here. Thus, significant positive evidence supports (11), and no Latinate derivative of English appears to systematically deviate from this ranking.

4. *Frequency*

We have considered alternative interpretations of this data, in which frequency-based factors decide if a derivative resembles any of its Bases, and if so, which one. We anticipated, following Hay 2003, that the relative frequencies of the derivative and its bases could modulate the effects of markedness improvement. Absolute frequencies could also play a role: infrequent B^Rs can be harder to access, altogether unknown to some speakers, or believed by speakers to be inaccessible to their listeners, and avoided on that score. For any of these reasons, infrequent B^Rs would be less likely to affect stress judgments.

A different frequency-based hypothesis is offered by Collie 2007:288-290, Collie 2008 and Dabouis 2019, who propose that the frequencies of the derivative, of the B^L, and of the B^R are the only factors that select a base when several exist. In these proposals, Markedness plays no role in the stress of derivatives. The data sets considered by Collie and Dabouis differ from ours: they studied UK pronunciations, while we focus on their US versions, which differ in many critical cases. But the same frequency factors could be relevant to the behavior of both populations, so we explored three frequency hypotheses inspired by Collie and Dabouis. (14) lists the possibilities tested: (14.b-d) are frequency-based hypotheses; (14.a) is the markedness improvement hypothesis corresponding to the M >> C-CONTAINMENT part of (11).

14. Hypotheses tested
    a. M >> C-Containment: If D has an optimizing B^R, its stress matches that B^R
    b. D frequency: a frequent D optimizes its stress regardless of how its Bs are stressed.
    c. Relative frequency of B^L vs. D: the more frequent the B^L relative to the D, the more likely the D is to resemble it.
    d. Relative frequency of B^R vs. B^L: if some B^R is more frequent than the B^L, the D is more likely to resemble that B^R.

We used logistic regression analyses to test the significance of these factors in the stress of nine derivative types: *-ee, -able, -ify, -ician, -icity, -ive, -ivity, -oid and -ory*. The dependent variable is whether the stress on D's stem matches the stress of the B^L. As before, we consider only Ds whose ability to satisfy M_2 conflicts with faithfulness to their B^L. Predictors correspond to the four hypotheses in (14):

15. Predictors

Better B^R (14.a)	If D has a B^R whose stress is optimizing, assign a 1; else assign a 0
Frequent D (14.b)	Frequency of D; value of 0-8, from OED's frequency bins[8].
Frequent B^L - D (14.c)	Frequency B^L minus of frequency of D; from OED's frequency bins.
Frequent B^R (14.d)	If D has a B^R more frequent than the B^L, assign a 1; else assign a 0

The overall result, in (16), is that the Better B^R predictor is significant in all nine data sets, while the frequency predictors don't have consistent effects. 'No-effect' cells are bolded.

16. Results

D type (# of forms)	Better B^R	Frequent D	Frequent B^L - D	Frequent B^R
	If better B^R exists,	As $Freq_D$ grows,	As $Freq_{BL}$ increases wrt $Freq_D$	If B^R is more frequent than the B^L
-able (n=397)	match D-B^L less likely ($p <.01$)	match D-B^L less likely ($p =.08$)	match D-B^L more likely ($p =.05$)	match D-B^L less likely ($p =.09$)
-ee (n=101)	match D-B^L less likely ($p <.001$)	match D-B^L less likely ($p =.05$)	no effect ($p>.1$)	no effect ($p>.1$)
-ician (n=55)	match D-B^L less likely ($p <.01$)	no effect ($p>.1$)	no effect ($p>.1$)	no effect ($p>.1$)
-icity (n=65)	match D-B^L less likely ($p <.05$)	no effect ($p>.1$)	no effect ($p>.1$)	no effect ($p>.1$)
-ify (n=42)	match D-B^L less likely ($p <.01$)	no effect ($p>.1$)	no effect ($p>.1$)	no effect ($p>.1$)
-ive (n=449)	match D-B^L less likely ($p <.001$)	no effect ($p>.1$)	match D-B^L less likely ($p <.06$)	no effect ($p>.1$)
-ivity (n=65)	match D-B^L less likely ($p <.05$)	no effect ($p>.1$)	no effect ($p>.1$)	no effect ($p>.1$)
-oid (n=113)	match D-B^L less likely ($p <.001$)	no effect ($p>.1$)	no effect ($p>.1$)	no effect ($p>.1$)
-ory (n=207)	match D-B^L less likely ($p <.001$)	match D-B^L less likely ($p <.01$)	no effect ($p>.1$)	no effect ($p>.1$)

We have seen that frequency factors may play a role in the phonology of a few Latinate derivative types, but they are insufficient as an alternative explanation of the patterns we have found: the M >> C-CONTAINMENT ranking we propose offers the better fit to this data.

5. *Beyond the dictionary*

Because this work started out as a dictionary study, we sought to understand the source of our data. How do dictionaries obtain the accentual data they report?

What is most relevant for us is that the bulk of the data we analyzed consists of *very* infrequent words. For example, the median OED frequency band to which the 390 *-able* forms we studied belong is 2. This corresponds to fewer than 0.0099 occurrences per million words. Similar frequency patterns hold for all other derivatives studied. Examples of words belonging to OED's frequency band 2 are *gelátinoid, pacíficatory* and *medícinable.* We find it unlikely that the

[8]Frequency bands are defined at https://public-oed-com.libproxy.mit.edu/how-to-use-the-oed/key-to-frequency/

OED's pronunciation editors have *heard* such forms and recall the stress of forms they heard. It seems more plausible that speakers generate their stress online, when the need arises.

If so, a significant portion of our dictionary data was obtained in a manner similar to a *wug* test: a linguist instructs a native speaker to produce a nonce word. The speaker uses their grammar and lexicon to generate it. The dictionary transcription results, in part, from this introspective exercise. A consultation with Catherine Sangster, OED's Head of Pronunciation, confirms that, for rare words, missing on Youglish.com, this is the central part of the procedure.

We report now on an effort to verify that a real *wug* test produces results that converge with the dictionary study. A distinct goal of the experiment was to verify the role of B^Rs in the stress computations of individual speakers. We wanted to check that the participants who chose a stress pattern attributable to an optimizing B^R – e.g. *gelátinous* for *gelátinoid*, as in (5) – actually know *gelátinous*. Conversely, we sought to verify that speakers who fail to use a potential B^R in computing the stress of some word fail to do so because they don't know, or reject, that B^R. Thus, if they reject *paráffinoid*, we wanted to check that they also reject *paráffinous,* a potential and sufficient source for the shifted stress in *paráffinoid*.

We selected for the *wug*-test two derivative types, -*ee* nouns and -*ify* verbs. Their pseudo-cyclic stress patterns in the dictionary data have been summarized above. For each derivative type, we constructed 20 test items and 20 fillers. All were nonce words. Each set of test items contained nonce derivatives that would violate *LAPSE (for -*ify*), or *CLASH (for -*ee*), if the stress of each derivative were to remain faithful to its B^L. In addition, each set of test items was split into 10 stimuli that have a potential optimizing B^R and 10 that do not. An effort was made to pair each derivative endowed with a B^R with a similar derivative that lacks one. Thus, in the pair of nonce verbs *moralify* (on B^L *móral*) and *coralify* (on *córal*)*,* the former has a potential B^R in *morálity,* which might license lapse-free *morálify*, while the latter lacks any B^R. In the pair of nonce nouns *reservee* (on B^L *resérve*) and *deservee* (on B^L *desérve*), only the former could benefit from *rèservátion,* whose stress avoids clash in *rèservée.* Otherwise, these paired items are matched in phonological shape and, as much as possible, in the frequencies of their B^Ls. Filler items differed in not posing any conflict between accentual markedness and faithfulness to the B^L: they included items like *raccoonify* and *pesteree.* Sample target stimuli are seen in (17). The full list is readable in the Appendix. Numbers in parentheses are OED's frequency bands. A parenthesized zero indicates the form does not occur in the dictionary: a nonce test word or a hypothetical B^R.

17. a. Some of the -*ify* stimuli. D and B^L were heard in Part 1, B^R in Part 2

Derivative	Stress profiles for D	B^L	B^R
Germanify (2)	0102, 1002	Gérman (7)	Germánic (5)
turbanify (0)	0102, 1002	túrban (5)	turbánic (0)
moralify (0)	0102, 1002	móral (6)	morálity (6)
coralify (0)	0102, 1002	córal (5)	corálity (0)
hydrogenify (0)	21002, 10002	hýdrogen (6)	hỳdrógenàte (3)
estrogenify (0)	21002, 10002	éstrogen (6)	estrógenàte (0)
nomadify (0)	0102, 1002	nómad (5)	nomádic (5)
saladify (0)	0102, 1002	sálad (5)	sládic (0)

b. Some of the *-ee* stimuli. D and B^L were heard in Part 1, B^R in Part 2

Derivative	Stress profiles for D	B^L	B^R
abusee (2)	201, 021	abúse (5)	àbusátion (0)
accusee (0)	201, 021	accúse (6)	àccusátion (5)
performee (0)	231, 021	perfórm (7)	pèrfôrmátion (0)
confirmee (2)	231, 021	confírm (6)	cònfirmátion (5)
instructee (0)	231, 021	instrúct (6)	ínstrùct (0)
constructee (0)	231, 021	constrúct (6)	cónstrùct (5)
selectee (3)	231, 021	seléct (6)	séléct (0)
rejectee (3)	231, 021	rejéct (6)	réjèct (5)

In the first part of the experiment, participants read frame sentences, which contained a nonce derivative and its B^L. They were then asked to choose between two pronunciations of each derivative, e.g. *morálify* vs. *móralify*; *dèservée* vs. *desèrvée*, and to indicate the strength of each preference on a 5-point scale. A sample frame accompanied by instructions appears below:

18. Fats go rancid in this temperature. In fact, they *rancidify* overnight.

Option 1 (a recording of *rancídify*)

Option 2 (a recording of *ráncidify*)

Which of these pronunciations do you prefer?

How strong is your preference?

Very weak 1 2 3 4 5 Very strong

As the examples above suggest, the alternate pronunciations presented for test items differed in whether a Markedness constraint, *LAPSE in some cases, *CLASH (and Rhythm Rule) in others, was satisfied or not. In the filler items, the choice offered was between forms that satisfy both M and F constraints (*pèsterée*) vs. forms that violate both (*pestèrée*).

In Part 2 of the experiment, subjects were asked to listen to a recorded sentence containing a word, not heard in Part 1, but which could have functioned as a B^R for a test item presented in Part 1. Corresponding to (18), the word subjects would hear in Part 2 would be *rancídity*. After hearing a sentence that contained such a word, the subjects were asked if they remembered ever hearing it before, with that pronunciation. The stimuli in this second part included both items that might be known to participants (*rancídity* belongs to OED's frequency band 4) and non-existent items, or items too rare be found even in the OED (e.g. *candídity*). This second class of stimuli were also potential B^Rs for test items in Part 1, e.g. for *candidify,* presented in Part 1 in a format parallel to (18).

50 participants completed the experimental tasks with *-ify* stimuli and related words. 50 completed the tasks involving *-ee*. The participants were recruited through Amazon's Mechanical Turk, had US IP addresses, and reported being native speakers of American English in a demographic survey completed after the experiment. They all had a 95% acceptance rate for previous HITs.

6. Experiment results

Participants overwhelmingly preferred the predicted stress patterns of filler items (e.g. *pèsterée*) compared to alternatives (*pestèrée.)* Recall that the predicted stress of fillers satisfies M and F constraints, as well as C-CONTAINMENT. Preference rates for the predicted stress exceeded 65% of responses to each of the *-ify* stimuli and 75% of responses to two thirds of these stimuli. For filler items in *-ee,* the preference rates for the predicted stress were above 75% of responses in each case[9]. All these results align with the predictions of the analysis in (11).

For test items, item-by-item graphs are found in the Appendix. We compare first preference rates for derivatives that have a known optimizing B^R vs. parallel derivatives that lack such a B^R. This means, for instance, comparing the preference rate for *móralify* vs. *morálify* to that for *córalify* vs. *corálify.* These pairwise comparisons show that the preference for B^L-faithful stress in derivatives that lack a B^R (*córalify,* no *corál-*) generally exceeds the preference for B^L-faithful stress in a derivative that *has* an optimizing B^R (*móralify;* B^R *morálity*).

This aspect of our results confirms that speakers are reluctant to assign a stress profile that's unattested in a lexical family, compared to one that is instantiated in a member of the family. This reluctance corresponds in our model to the undominated status of the BD CORR constraint.

Consider now just the *-ify* derivatives that have the benefit of better B^Rs. All but one of these elicit some preference for shifted stress (*morálify*) and against B^L-faithful stress (*móralify*). For 7/10 items this preference is expressed in over 70% of the responses. The fact that optimizing B^Rs have this systematic effect on stress ratings suggests that accentual Markedness has greater weight than the preference to match the B^L stress. In our model, this corresponds to the ranking M >> C-Containment.

The pattern of responses to *-ee* items is more complex. Here too, the preference to retract stress when a B^R licenses this change (e.g. B^L *resérve*, D *rèservée*, B^R *rèservátion*) is greater than the preference to retract in the absence of a B^R (e.g. B^L *desérve*, D ??*dèservée*, no B^R), as predicted. However, a clear preference for stress retraction is found in only one *-ee* item (*cònnotée*, B^R *cònnotátion*). Preference rates for stress retraction in four other comparable items hover around 50%; and 5/10 *-ee* nouns with optimizing B^R's elicited more responses indicating faithfulness to the B^L, in violation of *CLASH. This is not predicted by the M >> C-Containment ranking. An interpretation is suggested below.

---

[9] We realized too late that the filler list included just one item (*advérsify*) in possession of B^Rs with different stress. These B^Rs are the Rhythm Rule variant *ádverse*; plus *ádversàry* and its derivatives. The preference rate for *advérsify* and against *ádversify* is above 75%, in line with most other results for *-ify* fillers. This item suggests that the *non-*optimizing B^Rs are ignored by our participants when they calculate the stress of nonce words. This aligns with the results of the dictionary study, but the experimental data is insufficient to establish this point with confidence.

In the responses to Part 2 of the experiment, participants identified which potential B^Rs for the derivatives in Part 1 were known to them. Indeed, 9/10 of the B^R's to *-ify* verbs that we had conjectured would be known to speakers elicited substantial rates of 'known' responses, mostly above 90%. By contrast, imaginary B^Rs like *muffínity* and *turbánic* (for *muffinify* and *turbanify*) were reported 'known' at much lower rates, generally well below 30%. The lower rates are in line with predictions.

Potential B^R's to *-ee* nouns present a parallel contrast, but with a difference. Forms like trochaic *ínstrùct*, which we had anticipated, based on the dictionary data, to be entirely imaginary sounded familiar to our participants more frequently than expected. All non-occurring forms in this class were reported as 'known' at rates of above 30%; trochaic *ínstrùct* and two others at rates above 80%.

These responses suggest false lexical memories that must be explained, and they seem to relate to the frequent reluctance to shift stress in *-ee* nouns relative to the B^L. A possible factor in all these unexpected outcomes is that some of our participants might distinguish less reliably the prominence contours of all-stress sequences, i.e. 21 from 12, e.g. *ínstrùct* vs. *ìnstrúct*, compared to the prominence contours found in alternating stress sequences, i.e. 0102 vs. 1002. The difficulty in telling apart 21 from 12 can explain why non-existent *ínstrùct* is reported as known. Perhaps the same difficulty explains the fact that 231 (*cònstrûctée*) is not as frequently preferred to 321 (*cônstrùctée*) as one might expect: no clear preference can exist between hard-to-distinguish sequences. The only constant preference in such cases might be to repeat the stress pattern primed earlier in the frame sentence by the B^L. If indeed the 231 is hard to discriminate from 321, this issue was pervasive in *-ee* data, as 11/20 derivatives in *-ee* contained 231 contours or 321 or both.

A mixed effects logistic regression was fit to the results of the test items in each one of the *-ify* and *-ee* experiments. Among the predictors tested, one reflects the basic ranking we propose for pseudo-cyclic effects: BD CORR >> M >> C-CONTAINMENT. This binary predictor is dubbed 'B^R known' and records whether the participant stated, in Part 2, that they know the B^R of a derivative whose stress they had evaluated in Part 1. Recall that the B^Rs we discuss are all optimizing: they license a shift of stress relative to the B^L, a shift that improves the markedness of the derivative. Thus, knowing a B^R means being able to use the ranking BD CORR >> M >> C-CONTAINMENT, to get a better stress for the derivative, based on this B^R, than the B^L offers.

We also tested three numerical, frequency-related predictors: B^R frequency (the OED frequency band of the B^R), B^L frequency and D frequency, defined in parallel ways. An initial model for the fixed effects structure of the *-ify* results used all four predictors – B^R known, B^R frequency, B^L frequency and D frequency – but it proved to offer no significantly better fit to the data than the two-predictor model in (19):

19. Results from the *-ify* model

Predictor	Coefficient	*z* value	Significant?
B_R frequency	-0.26	-4.18	Yes ($p < .001$)
B_R known?	-0.68	-3.37	Yes ($p < .001$)

For the *-ee* data, the simplest model for the fixed effect structure had one predictor, B_R frequency. The predictor "BR known?" did not emerge as significant. A possible factor in "BR known" not reaching significance in the *-ee* data is that a quarter, 6/20, B^Rs of frequency 0 (e.g. *ínstrùct*) may have been hard to distinguish from real words with opposite stress (i.e. *ìnstrúct*).

20. Results from the *-ee* model

Predictor	Coefficient	*z* value	Significant?
B_R frequency	-0.28	-3.61	Yes ($p < .001$)

7. Discussion

The results of this initial experiment converge with the dictionary study and add information missing from it. The results of both two studies converge on showing a preference for shifting stress relative to the B^L (e.g. *morálify* vs. B^L *móral*), just in case the desirable stress resulting from shift improves markedness and is independently attested in the same lexical family (e.g. B^R *morálity*). The experiment begins to verify that 'independently attested' means 'known to the speaker who prefers shifting stress'. We're dealing here with a link between what individual speakers know about their lexicon and what they prefer to do with that knowledge.

The frequency effect that plays a role in the experimental results is the one we had anticipated based on the dictionary study: rare B^Rs are less likely to license a stress shift. That could be because they're unknown, or harder to activate, or because speakers anticipate that they yield harder-to-interpret words. The other frequency effects remain elusive in this domain.

Not enough data could be brought to bear here on an important issue of interest, that of lexical listing. In cyclic inheritance effects, a cyclic base can be a never-before encountered expression, one whose properties can and must be computed online. The compound examples mentioned in section 1 provide examples: e.g. Liberman and Prince's nonce *law degree requirement changes,* with its cyclic stress, can be the base of yet another compound, which will inherit the predictable stress of the inner form, and so forth. Nothing needs to be lexically listed in such cases. Does pseudo-cyclic inheritance differ? Do remote bases have to have been heard to license a deviation from the B^L? Or is it sufficient to infer the stress of a potential B^R? The simple model proposed in section 2 derives cyclic and pseudo-cyclic effects from the same hierarchy and thus predicts that whatever is true of the cycle will be true of the pseudo-cycle, but real evidence is missing.

However, among the results of the *-ify* experiment there is a suggestive datum that could bear on this question: 68% of the responses to *candídify* expressed a preference for the shifted stress, 2102, against the B^L-faithful stress, 1002. The only potential B^R for this item is the very rare *candídity*. Only 38% of the subjects declared they know it. It is possible that what licenses the stress shift in *candídify* is not direct acquaintance with *candídity* but awareness of an island-of-reliability (Albright 2002) in English morphology, which allows one to predict that any *-id* adjective has a legitimate potential *-ity* derivative. If this fact is known, the stress on *candídity* follows, and, from it, the stress on *candídify*. This reasoning, while explaining the otherwise isolated preference to shift stress in *candídify*, opens up a host of new questions: failure to find a form in the OED is no longer a guarantee that the speakers can't predict its potential existence and its potential phonological properties. Then, if the analyses that precede are on the right track, there is considerably more to find out about how speakers imagine some potential forms and build inferences on their properties, while ruling out many others.

We hope to return in the future to this trade-free zone, bringing commodities that could bear on this last point.

References

Albright, A., 2002. Islands of reliability for regular morphology: Evidence from Italian. *Language*, pp.684-709.

Bachrach, A. and Nevins, A. eds., 2008. *Inflectional identity* (No. 18). Oxford University Press.

Becker, M. and Gouskova, M., 2016. Source-oriented generalizations as grammar inference in Russian vowel deletion. *Linguistic inquiry*, *47*(3), pp.391-425.

Benua, L., 1997. *Transderivational identity: Phonological relations between words* (University of Massachusetts, Amherst Doctoral dissertation).

Burzio, Luigi, 1998. Multiple Correspondence, *Lingua* 103, 79-109

Burzio, L., 2005. Sources of Paradigm Uniformity. Paradigms in Phonological Theory, ed. by Laura J. Downing, Tracy A. Hall and Renate Raffelsiefen, 65-106.

Chomsky, Noam, Morris Halle, and Fred Lukoff, 1956. "On accent and juncture in English", in M. Halle et al. (eds.), *For Roman Jakobson*, 65-80. The Hague: Mouton

Chomsky, Noam and Morris Halle, 1968. *The Sound Pattern of English*. Harper and Row.

Collie S. 2007. *English stress preservation and stratal optimality theory*. The University of Edinburgh Doctoral dissertation.

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

Dabouis, Q., 2019. When accent preservation leads to clash. *English Language & Linguistics*, *23*(2), pp.363-404.

Fougeron C., Kuehnert, B., Imperio, M. and Vallee, N. eds., 2010. *Laboratory phonology 10* (Vol. 4, No. 4). Walter de Gruyter.

Ito, Junko & Armin Mester 1999. The phonological lexicon. In Natsuko Tsujimura (ed.) The Handbook of Japanese Linguistics, Oxford: Blackwell. 62–100.

Gordon, M., 2002. A factorial typology of quantity-insensitive stress. *Natural Language & Linguistic Theory*, *20*(3), pp.491-552.

Halle, M., 1973. Prolegomena to a theory of word formation. *Linguistic inquiry*, *4*(1), pp.3-16.

Hay, Jennifer, 2003. *Causes and Consequences of word structure*. Routledge

Hayes, Bruce, 1982. Extrametricality and English stress. *Linguistic Inquiry*. 13:227-276

Hung, H.J., 1994. *The rhythmic and prosodic organization of edge constituents*. Brandeis University Doctoral dissertation.

Kager, R.W.J., 1989. *A metrical theory of stress and destressing in English and Dutch*. Doctoral dissertation, University Utrecht.

Kager, R.W.J., 1996. On affix allomorphy and syllable counting. *Studia Grammatica*, *41*, pp.155-171.

Kiparsky, Paul, 1979. Metrical structure assignment is cyclic. *Linguistic Inquiry*. 10:421-441.

Kösling, K., Kunter, G., Baayen, H. and Plag, I., 2013. Prominence in triconstituent compounds: Pitch contours and linguistic theory. *Language and Speech*, *56*(4), pp.529-554.

Liberman, M. and Prince, A., 1977. On stress and linguistic rhythm. *Linguistic inquiry*, *8*(2), pp.249-336.

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

McCarthy, John and Alan Prince, 1995. Faithfulness and reduplicative identity. In Beckman, Dickey and Urbanczyk (eds.) *Papers in Optimality Theory*. University of Massachusetts Occasional Papers, 18, p. 249-384. GLSA, UMass Amherst.

Pater, Joe, 2008. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In Steve Parker (ed.) *Phonological argumentation: Essays on Evidence and motivation*, London: Equinox. 123–154.

Pierrehumbert, J.B., 2006. The statistical basis of an unnatural alternation. *Laboratory phonology*, *8*, pp.81-107.

Pierrehumbert, Janet and Cynthia Clopper, 2010. What is Labphon? And where is it going? In Fougeron et al. (eds.) 2010, pp. 113-132.

Raffelsiefen, R., 2004. Absolute ill-formedness and other morphophonological effects. *Phonology*, pp.91-142.

Stanton, Juliet and Donca Steriade, 2018. *Stress and the cycle, in English and elsewhere*. Ms. in progress, NYU and MIT.

Steriade, Donca. 1999. Lexical conservatism in French adjectival liaison, in B. Bullock, M. Authier, and L. Reed (eds.), *Procedings of LSRL 25*. Benjamins, pp. 243–71.

Steriade, Donca 2008 A pseudo-cyclic effect in the Romanian declension, A.Bachrach and A.Nevins (eds.) *Inflectional Identity*, Oxford University Press, pp. 313-361

Steriade, Donca and Igor Yanovich 2015 Accentual allomorphs in East Slavic: an argument for inflection dependence. in E. Bonet, M.R Lloret, J. Mascaró (eds) *Understanding Allomorphy*, 254-314.

Appendix: Results of -ify and -ee test items