Markedness drives base selection experimental evidence for the pseudo-cycle

Juliet Stanton (NYU) and Donca Steriade (MIT) stanton@nyu.edu steriade@mit.edu

Plan

- 1. Define the pseudo-cycle, for English
- 2. Lexical evidence that pseudo-cyclic stress is general in English
- 3. Behavioral evidence that pseudo-cyclic stress is productive in English

Cycle and pseudo-cycle

- The phonological cycle (Chomsky, Halle and Lukoff 1956, Chomsky and Halle 1968):
 - \circ derived phonological info is inherited from a Base by its Derivatives.
 - the B is embedded as an immediate constituent in its Ds
 - Local Base, B^L
- The pseudo-cycle (Steriade 1999, 2008 et seq, Stanton and Steriade 2018 ; cf. Breiss 2021)
 - \circ derived phonological info is inherited by Ds from a related form.
 - \circ $\,$ the related form is not embedded as an immediate constituent in the Ds $\,$

3

 \circ Remote Base, B^R

The basic phenomenon

páraffinòid	1002	páraffin	100		
gelátinòid	0102	gélatin	100	gelátinous	0100

Bases

Derivative		Local Base, B ^L		Remote Base, BR	
páraffinòid	1002	páraffin	100	*paráffinous	0100
gelátinòid	0102	gélatin	100	gelátinous	0100

Claims

Derivative		Local Base, B ^L		Remote Base, B ^R	
páraffin-òid	1002	páraffin	100	*paráffinous	0100
gelátin-òid	0102	gélatin	100	gelátin-ous	0100

• A productive process is responsible for *gelátin-oid* using the stem of *gelátin-ous*

- It is motivated by **Markedness**: 010-2 better than lapsed 100-2 (cf. Appendix 1)
- $\circ\,$ Not frequency, though frequency modulates the use of $B^R\,s$

	B: páraffin	100	BD IDSTRESS	*LAPSE _{LAT}
⊫ra.	páraffin-òid	1002		*
b.	paráffin-òid	0102	*!*	

BD Id Stress: One * for every syllable in the Base whose D correspondent differs in [±stress] *Lapse _{Lat}: One * for every 00 sequence (Latinate vocabulary)

	B ^L : gélatin	100	BD IDSTRESS	*LAPSE _{LAT}
	B^R: gelátin-ous	010-0		
Ba.	gelátin ^R -òid	0102		
b.	gélatin ^L -òid	1002		*!

	B: páraffin	100	BD IDSTRESS	*LAPSE _{LAT}
œa.	páraffin-òid	1002		*
b.	paráffin-òid	0102	*!*	

BD Id Stress: One * for every syllable in the Base whose D correspondent differs in [±stress] *Lapse _{Lat}: One * for every 00 sequence (Latinate vocabulary)

	B ^L : gélatin	100	BD IDSTRESS	*LAPSE _{LAT}
	B ^R : gelátin-ous	010-0		
ra.	gelátin ^R -òid	0102		
b.	gélatin ^L -òid	1002		*!
c.	gelátin ^L -òid	0102	*!	

	B: páraffin	100	BD IDSTRESS	*LAPSE _{LAT}
⊫ra.	páraffin-òid	1002		*
b.	paráffin-òid	0102	*!*	

BD Id Stress: One * for every syllable in the Base whose D correspondent differs in [±stress] *Lapse _{Lat}: One * for every 00 sequence (Latinate vocabulary)

	B ^L : gélatin	100	BD IDSTRESS	*LAPSE _{LAT}	CORR B ^L
	B^R: gelátin-ous	010-0			
⊯a.	gelátin ^R -òid	0102			*
b.	gélatin ^L -òid	1002		*!	
c.	gelátin ^L -òid	0102	*!		

Corr B^L: One * for D whose stem does not correspond to its local Base, B^L

Similar pairs of Latinate Ds

	B _L stress		D stress like B _L		D stress \neq B _L		BR	
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	02100				

Stress data throughout is based on OED transcriptions (stress marks, V-reduction and flapping patterns)



CORR BD = BD IDENT STRESS

Plus affix-specific M constraints



How general are pseudo-cyclic patterns in English?

Database

- Data sources: 27 Latinate derivatives from Marchand 1969
- 12 offer enough data for analysis
- Lists of Ds for each suffix: online OED, Lehnert 1979
- Stresses (for US English): online OED, Merriam Webster's
- In each Latinate D, we record:

D, B^L, B^R s: IPA, stress, OED frequency bin of D and Bs.

Stress change between B^L and D's stem;

M constraints trigger the change.

Focus on

- Ds conflicted between faithfulness to B^L and M_{Lat}

like, *gelátin-oid, páraffin-oid* not, e.g. *amýgdal-oid,* B^L *amýgdala*

• B^R s that are *optimizing* (if D is faithful to B^R , not B^L , D has better M score)

like, B^R gelátin-ous for D gelatin-oid; B^L gélatin

not, e.g. B^R pàraffin-ic for D páraffin-oid; B^L páraffin: pàraffinòid 2013 worse than páraffinòid)

• Qs for each D:

Does D have an optimizing B^R? If so, does D's stress match that B^R?

- Predictions follow from CORR BD >> M_{LAT} >> CORR B^{L}

Global result from dictionary data

- Effect of 'Better B^R' is a significant predictor of stress in 12 Latinate Ds: -*ify*, -*ee*, -*able*, -*ary*, -*ory*, -*ive*, -*ician*, -*ivity*, -*icity*, -*ism*, -*ite*, -*oid*+ root cpds (*cerébroscòpe*, *cèrebróscopy*; *polárogràph* vs. *pálatogràph*)
- Effect of 'Better B^R' can't be detected in 9 Latinate D types: (-ation, -eer, ese, -esque, -ess, -ette, -ist, -ite and -ous): too few Ds have optimizing B^Rs
- Generally, no Latinate D of English systematically deviates from model.

How productive is the pseudo-cycle?

o dictionary stress might represent just dictionary editors' judgments
o does not say how a speaker's lexicon affects their stress preferences
o does not exclude the possibility that some stress patterns are lexicalized.

We test pseudo-cyclic preferences in nonce Ds

- 3 sets of 20 nonce Latinate derivatives (stimuli in Appendix 2)
 - *-ify (humid-ify, nonce moral-ify)*
 - *-ee (employ-ee, nonce compar-ee)*
 - \circ compounds with bound 2nd members (*insect-i-cide*, nonce *antenn-i-cide*)
- M (*LAPSE, *CLASH) conflict with BD ID STRESS and/or CORRB^L in all test items
- E.g. *móral-ify* (< B^L móral, *LAPSE)? *morálify* (< B^R *morál-ity*, CORRB^L)?

We test pseudo-cyclic preferences in nonce Ds

- Experiments constructed in Experigen (Becker & Levine 2013)
- 50 subjects recruited thru Amazon's MT
 - for each condition (-*ify*, -*ee*, bound-2nd member compounds)
- Test items: half in each set have optimizing B^R : D moralify, B^L móral 10, B^R morál-ity 0100
- Test items: the other half lack any known optimizing B^R D coralify, B^L córal 10, no B^R (no *corál-ity 0100)
- 20 Fillers in each set, posing no conflict of M and BD Ident D suspénsify, B^L suspénse 01

Test, Part 1: elicit stress preference

I'm already moral enough. No reason to <i>moralify</i> me any further.
Option 1
Option 2
Which of these pronunciations do you prefer?
Option 1 Option 2
How strong is your preference?
Very weak 1 2 3 4 5 Very strong

20

Test, Part 2: is a potential optimizing B^R known?

The word is *morality*.

Pronunciation (0100)

Have you heard this word, pronounced this way, before?



Test, Part 2: is a potential optimizing B^R known?

The word is *corality*.

Pronunciation (0100)

Have you heard this word, pronounced in this way, before?

Yes



Predictors of nonce stress preferences

- a. **B^R in OED** (binary predictor): Does the OED recognize the B^R's existence?
- b. **B^R known?** (binary predictor): Does the participant report knowing the B^R?
- c. **D frequency** (numerical predictor): OED frequency band of D
- d. **B^L frequency** (numerical predictor) OED frequency band of B^L

NB: frequency-based factors are included because they have been used to explain pseudocylic effects in earlier literature (e.g. Hay 2003, Collie 2008, Dabouis 2019).

Results for *-ify* test: nonce words



(For by-item plots, see appendix 3!)

Statistics for *-ify*

- In separate models...
 - **B**^R in OED is a significant predictor of D stress (p < .001)
 - **B**^R known is a significant predictor of D stress (p < .001)
 - (Which is the better predictor? **B**^R known.)
- In the best-fit model...
 - Frequency-related factors (D frequency, B^L frequency) not significant.

Results for *-ee* test: nonce words



Statistics for -ee

- In separate models...
 - **B**^R in OED is a significant predictor of D stress (p < .001).
 - **B**^R known is not a significant predictor of D stress (p = 0.75).
- In the best-fit model...
 - Frequency-related factors (D frequency, B^L frequency) not significant.

Insect-i-cide compounds

- 2nd member is typically a bound root: *immun-o-gen, demon-o-graph-y*
- It has fixed stress
 - On unsuffixed root: $colúmn-i-\underline{form}_A$, $icon-o-\underline{clast}$, $immún-o-\underline{gen}$ Penult/antepenult otherwise: $demon-o-\underline{graph-y}$, $reptil-i-\underline{fer-ous}$
- 1st member can be a free standing word: <u>sándal-i-fòrm</u>
- 1^{st} member satisfies *LAPSE/CLASH, if it has an optimizing B^R

cèrebr-ó-scop-y vs. *cerébr-o-scòpe* (B^L: *cérebr-um* B^R: *cerébr-al*) *pèctin-í-ferous* vs. *pèctin-o-lýtic* (B^L *péctin*, B^R *pectínic* 0 frequency)

Results for compound test: nonce words



Statistics for compounds

- In separate models...
 - **B**^R in OED is a significant predictor of D stress (p < .001)
 - **B**^R known is a significant predictor of D stress (p < .01)
 - (Which is the better predictor? B^R in OED.)
- In the best-fit model...
 - Frequency-related factors (D frequency, B^L frequency) not significant.

Summary of experimental results

- In all three tasks, preferred stress pattern of the D was influenced by a B^R: D significantly more likely to shift stress if optimizing B^R exists.
 - Results were inconsistent as to whether **B**^R in **OED** or **B**^R known was a better predictor of participant responses.
 - Probably: neither measure is a great approximation of speakers' knowledge.
- Frequency-related factors are not significant here, probably because we controlled for B^R, D, and B^L frequency. Outside of these conditions, however, frequency is likely to play a role.
- **Importantly**: the effect of markedness, i.e. of the optimizing B^R, is independent.

To conclude: what we understand

- Pseudo-cyclic stress patterns detected in dictionary data are matched by subject preferences for stress in nonce words.
- See also Breiss 2021 for independent findings in the same vein
- Results supports the general model offered

 \dots BD CORR >> M_{LAT} >> CORRB^L >> M \dots

and the hypothesis of violable CORRB^L in particular

• They lend credibility to OED's methods for eliciting stress in rare words and to analyses based on those stresses.

To conclude: what don't understand

1. What knowledge do speakers use in deciding that a word – a potential optimizing B^{R} – does or doesn't exist in the language?

Cf. *candidity* vs. *rancidity*: 0 vs. 4 frequency, but are treated as comparably good B^Rs by participants in the *-ify* test. Conjecture: an optimizing B^R is any potential word guaranteed as well-formed by the grammar. We have not yet tested this in English.

2. What mechanism ensures that optimizing B^Rs are available as inputs in the computation of a Derivative's phonology?

Thank you

References

- Benua, Laura. 1997. Transderivational identity: Phonological relations between words. Ph.D. thesis. UMassAmherst.
- Breiss, Canaan 2021 Inside the wug test https://www.cuny2021.io/wp-content/uploads/2021/02/CUNY_2021_abstract_350.pdf
- 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. The Sound Pattern of English. Harper and Row. 1968
- Collie, Sarah 2008 English stress preservation: the case for 'fake cyclicity'. PhD dissertation, University of Manchester.
- Dabouis, Quentin 2019. When accent preservation leads to clash. English Language & Linguistics 23. 363-404.
- Hay, Jennifer 2003 Causes and Consequences of word structure. Routledge
- Hayes, Bruce 1982 Extrametricality and English stress. Linguistic Inquiry. 13:227-276.
- Lehnert, Martin (1971). Reverse dictionary of present-day English. VEB Verlag Enzyklopädie.
- 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.
- Prince, A. and P. Smolensky 2004 (1993) Optimality Theory: Constraint interaction in generative grammar. Wiley-Blackwell
- Stanton, Juliet and Donca Steriade 2018, in prog. Stress and the cycle, in English and elsewhere, MIT and NYU ms.
- Steriade, Donca. 1999. Lexical conservatism in French adjectival liaison, in B. Bullock, M. Authier, and L. Reed (eds.), *Proceedings 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 2019 CiV Lengthning and the length of CV, http://lingphil.mit.edu/papers/steriade/11_Steriade_2019_CiV-lengthening-and-the-weight-of-CV.pdf
- 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.
- Steriade, Donca and Juliet Stanton Productive pseudocyclicity, talk at Labphon 2020 http://lingphil.mit.edu/papers/steriade/Steriade_and_Stanton_2020_Labphon.pdf

Appendix 1: Latinate Lapse

- In simple, non-Latinate words: STRESSL >> *LAPSE àbracadábra, Winnepesáukee, hùllabalóo, càtamarán etc. (Hayes 1982)
- In Latinate words lacking a free Base: *LAPSE_{LAT} >> STRESSL metábolism, Epàminóndas, manípulàte, elíminàte, accéleràte, other -ate apócalýpse, acétylène, apólogète, apóthecàry (Liberman & Prince 1977:274ff)
- Deviations from *LAPSE_{LAT} happen to satisfy BD IDENT STRESS: *Áfrican-ist, èlephant-íasis, ècono-métrical,* etc.
- or involve iV (amélioràte, cértioràri) which tolerates lapse (Steriade 2019).

Appendix 2a: items for -ify

- 1. I am going to address you in Pig Latin. I am going to Pig *Latinify* this whole speech.
- 2. Let's talk muffins. And zucchinis. Zucchini bread is too big to carry, but you can *muffinify* it to make it portable.
- 3. Fats go rancid quickly in this temperature. In fact, they *rancidify* overnight.
- 4. At first, the recorded discussion won't be candid. But when the cameras stop, the banter will instantly *candidify*.
- 5. The first goal of this German settlement is to *Germanify* the entire area.
- 6. You want to make a proper turban? Then you should learn to *turbanify* a heavy cloth, of about 3 feet by 4.
- 7. I'm already moral enough. There's no need to *moralify* me any further.
- 8. Coral reefs are made up of skeletons from coral colonies. As other chemical elements are incorporated into calcium carbonate deposits, the skeletons *coralify*.
- 9. Liquid hydrogen is too exotic as a fuel. What we should *hydrogenify* are the electrical power plants, not the cars.
- 10. Afraid of too much estrogen? Cut the soy milk, because it's sure to *estrogenify* you.

Appendix 2a: items for -ify

- 11. A nomad's life is what you want? You can *nomadify* your possessions if they fit in this bag.
- 12. If you prefer salad, then you can *saladify* this chili by adding these greens.
- 13. We compared the human- and computer-generated solutions and decided it was best to *humanify* all of them.
- 14. He was not an original shaman. He had a friend *shamanify* him at the Happy Daze fest.
- 15. To look like a hardcore metal fan, it's not enough to just *metalify* your makeup and hairspray.
- 16. This bike doesn't come with pedals, but you can easily *pedalify* it in one visit to the shop.
- 17. Bush was the victor, but he really needed Rove's tactics to *victorify* himself.
- 18. He will become a rector, but do you know who will *rectorify* him?
- 19. A platform that connects tutors with students who need them is *tutorify*.inc
- 20. I have never been a debtor, but you could say that these repairs have managed to *debtorify* me.

Appendix 2b: items for -ee

- 1. I wasn't sure what to call someone who reserves the right to live and own property on Indian land, but federal law usually calls them a *reservee*.
- 2. He's giving dues to the person who deserves it. Or should we say, he's giving dues to the *deservee*.
- 3. If you receive a letter that is not addressed to you, please write 'return to sender' on it. This will inform your letter carrier that the *addressee* does not live there.
- 4. History shows us that the oppressor and the *oppressee* can never be equal.
- 5. His remark connoted that an employee had not been doing their job, though it was not clear who the *connotee* was.
- 6. Individuals who are promoted will be on probation for 12 months. Then, the *promotee* will be evaluated by their supervisor.
- 7. Informal processes are appealing for the accuser. But both the accuser and the *accusee* have to agree to such a process.
- 8. Most abusers don't seem to think they could be the problem. So it's usually up to the *abusee* to stop the cycle of violence.
- 9. In this ceremony, the candidate confirms promises made at their baptism. While this is not a sacrament, the *confirmee* must be a member of the Church.
- 10. The pianist is silent. This makes the audience grunt, cough, murmur, and rattle seats. The traditional roles of performer and *performee* have been inverted.

Appendix 2b: items for -ee

- 11. One can think of the embryo constructing itself, but this separation of a developing organism into constructor and *constructee* is difficult.
- 12. The Naughty Grammarian instructs on the finer points of English usage. It is fun for the instructor, but not for the *instructee*.
- 13. In most relationships that break up, there is a rejector and a *rejectee*.
- 14. If your boarding pass has the letters SSSS on it, it means you've been selected for a thorough examination. Your name is on a highly classified list: the *selectee* list.
- 15. We tried to infuriate the others residing in the room. The method was to occupy the doorway of the room, taunting each *residee*.
- 16. You could say that a doctor is a provider, and each patient is a *providee*.
- 17. A long time ago, a church member was suspected of heresy. The pastor was the suspecter, so he got to interrogate the *suspectee*.
- 18. The degree of respect conveyed by such expressions depends on the speaker's age and gender, but also on the *respectee*'s social status.
- 19. Each tax preparer must give this information to their client, the tax *preparee*, before it is disclosed to others.
- 20. The pastor often likes to compare new church members to old ones. I guess you could say that all church members are used to being a *comparee*.

Appendix 2c: items for compounds

- 1. A ruler who is kept in power by demons is a *demonocrat*.
- 2. A ruler of planets is a *planetocrat*.
- 3. A substance that kills arbors is *arboricidal*.
- 4. Fear of calculus is *calculophobia*.
- 5. Someone who loves history is a *historiphile*.
- 6. A device for locating the margins of a solar system is a *marginoscope*.
- 7. A plot that results in the killing of a despot is called *despoticide*.
- 8. The skill of inferring a medical condition from its symptoms is *symptomatology*.
- 9. An obsessive love of icons is known as *iconomania*.
- 10. An irrational fear of flower petals is *petalophobia*.

Appendix 2a: items for compounds

- 11. A being that looks like an angel is *angeliform*.
- 12. A receptor shaped like the photoreceptors on the retina is *retiniform*.
- 13. The amount of acid in this substance is read by an *acidograph*.
- 14. If we had a measure of academic merit, we'd have a *meritograph*.
- 15. The technical term for this kind of plant is *botanophyte*.
- 16. When water becomes vapor, its state is *vaporiform*.
- 17. A substance that generates calories is a *calorigen*.
- 18. A substance that generates odors is an *odorigen*.
- 19. A process that generates metals can be referred to as a *metallogen*.
- 20. Experiments that record the contact between the tongue and the palate are *palatographic*.

Appendix 3a: by-item plot for -ify



Appendix 3b: by-item plot for -ee



Appendix 3c: by-item plot for compounds

