

Live and let live: Learning to combine competing accounts for phonological opacity

Aleksei Nazarov

Utrecht University

OCP22, Universiteit van Amsterdam

Overview

- Explanations for opaque interactions:
 - Intrinsic ordering
 - Extrinsic ordering
 - Lexicalization
- What if each opaque interaction has its own explanation?
- Framework to combine explanations (intrinsic > extrinsic > lexical)
- Initial learning simulations show feasibility:
 - Canadian Raising
 - Gran Canaria Spanish

Opacity Explanations

Opacity (Kiparsky 1971)

- Informally: interaction between 2 or more processes where one process does not take output of other process into account
- *North American English: Canadian Raising (e.g., Joos 1942, Vance 1987)*

1. $aV \rightarrow \Lambda V / _C$

$/rait/ \rightarrow [rΛit]$
 $/raɪd/ \rightarrow [raɪd]$

2. $t \rightarrow r / \acute{V}_V (*)$

$/kΛt-ə/ \rightarrow [kΛrə]$

when both applicable, 1 does not take output of 2 into account
(*counterbleeding*)

$/rait-ə/ \rightarrow [rΛɪrə]$

Opacity

- Informally: interaction between 2 or more processes where one process does not take output of other process into account
- *Gran Canaria Spanish (Broś 2016)*

1. $D \rightarrow \bar{D} / [+cont]$

/rasgo/ → [razyo]
/el gato/ → [el γato]

2. $T \rightarrow D / V_-$

/frekwensia/ → [fregwensia]
/otra klase de/ → [otra glase ðe]

when both applicable, 1 does not take output of 2 into account
(*counterfeeding*)

/frekwensia/ → [fregwensia], *[freywensia]
/otra klase de/ → [otra glase ðe], *[otra γlase ðe]

Explanations

- **Extrinsic ordering:** process 1 and 2 are not inherently restricted from interacting, but there is a mechanism that does restrict their interaction (rule ordering, ordering constraints, ...)
- E.g., Serial Markedness Reduction (Jarosz 2014):
 - Harmonic Serialism tracks improvements on markedness constraints, features S(erial)M(arkedness) constraints on the order of markedness satisfaction
- *Gran Canaria Spanish:*
Both processes apply in same domain (phrase)
Constraint SM(*[+cont]D, *VT) highly ranked

/frekwensia/ → [fregwensia] <*VT> *→ [freywensia] <*VT, *[+c]D>

Explanations

- **Intrinsic ordering:** processes 1 and 2 are in different layers (strata, layers of representation, ...) for independent reasons; the layer of process 2 does not feed into the layer of process 1
- E.g. Stratal OT (Bermúdez-Otero 1999, Kiparsky 2000):
 - Phrase-level processes do not feed into word-level processes, word-level process do not feed into stem-level processes
- *North American English (Bermúdez-Otero 2003):*
Raising: word-level /raɪt/ → [rΛɪt] /laɪ (tə)/ → [laɪ (rə)]
Flapping: phrase-level /kΛt-ə/ → [kΛrə] /laɪ tə/ → [laɪ rə]
Flapping cannot influence the application of Raising!

Explanations

- **Lexicalization:** process 1 is no longer productive/overruled by lexical factors, making it impossible for process 2 to influence application of process 1
- Has been hypothesized for North American English (e.g. Vance 1987):
 - Raising is lexical – supported by existence of lexical exceptions
 $/tʌɪgə/ \rightarrow [tʌɪgə]$
 - Therefore, Flapping cannot interact with it
 $/rʌɪt-ə/ \rightarrow [rʌɪtə], *[rʌɪtə]$
- However: evidence for productivity of Raising (e.g., Idsardi 2006, Farris-Trimble & Tessier 2019)

Exclusivity vs. inclusivity

- Often, focus is on choosing one of these explanations for opacity
- However, separate arguments for each type of explanation:
 - independent evidence for Stratal organization (e.g., Bermúdez-Otero 1999) and lexical factors (e.g., Becker 2009), but some cases appear to necessitate extrinsic ordering (e.g., McCarthy 2007)
- Alternative hypothesis:
 - Different opaque phenomena may have different explanations (intrinsic, extrinsic, lexical)
 - This necessitates adjustments to theoretical and learnability assumptions

Combining mechanisms

A single framework

- To allow different opaque interactions to have different explanations: one framework that allows all such explanations
- Start with standard Optimality Theory
- Lexical factors: indexed constraints (Pater 2000, 2010)
- Extrinsic ordering: Harmonic Serialism with SM constraints
- Intrinsic ordering: Stratal OT
- Does this mean we need Stratal Harmonic Serialism with SMR and indexed constraints?

A single framework

- Instead: parallel OT with various extensions
- Lexicalization: induce lexically indexed constraints on demand
- Intrinsic ordering: induce SM constraints on demand (+ include HS derivations only if necessary)
- Extrinsic ordering: induce constraints specific to Stratal levels on demand (+ include Stratal derivations only if necessary)
- Worst-case scenario still Stratal HS-SMR with indexation, but entire range of possibilities does not need to be used

Approximations

- To avoid actual Stratal HS-SMR and associated learnability challenges, Stratal and SMR elements approximated in Parallel OT
- Stratal constraints:
 - C_{Word} or C_{Stem} are “indexed” versions of constraint C only applicable with the Word *casu quo* Stem domain
 - Only covers some functions of Stratal ranking discrepancies, as no intermediate derivational level available

Illustration

- Stratal approach: Canadian Raising

<i>/raɪt-ə̄/</i>	*VTV	Ident-C	*aV_hC_oStem	Ident-V	*aV_hC
raɪtə̄	*!		*		*
rʌɪtə̄	*!			*	
raɪrə̄		*	*!		*
☞ rʌɪrə̄		*		*	

<i>/laɪ tə̄/</i>	*VTV	Ident-C	*aV_hC_oStem	Ident-V	*aV_hC
laɪ tə̄	*!				*
lʌɪ tə̄	*!			*	
☞ laɪ rə̄		*			*
lʌɪ rə̄		*		*!	

Approximations

- To avoid actual Stratal HS-SMR and associated learnability challenges, Stratal and SMR elements approximated in Parallel OT
- SM constraints:
 - $M1_{M2}$ is an “indexed” version of markedness constraint $M1$, which indicates the violations $M1$ had when $M2$ was last satisfied, if this was before the last step
 - No intermediate derivational level available, so derivations are estimated for each surface candidate, sometimes considering multiple derivations

Illustration

- SMR approach: Gran Canaria Spanish

/el gato/	*[+c]D _{*_{VT}}	* _{VT}	*[+c]D	Ident
el gato			*!	
☞ el yato				*

/frekwensia/	*[+c]D _{*_{VT}}	* _{VT}	*[+c]D	Ident
frekwensia		*!	*	
☞ fregwensia			*	*
fre(g→)ywensia	*!			*

Order of learning

- Logically: consider most restrictive hypothesis first (Subset Learning)
- Lexical explanation least restrictive: any process could be lexicalized given training data
- Extrinsic ordering more restrictive: process must be in grammar, but ordering is unrestricted
- Intrinsic ordering most restrictive: process ordering must correspond to morphosyntactic application domains
- Therefore, order of hypothesis consideration:
Intrinsic > Extrinsic > Lexical

Learnability experiments

- What happens if we adopt such an “on demand” approach?
- The hope:
 - Canadian Raising without lexical exceptions: Stratal constraints only
 - Gran Canaria Spanish: SM constraints only
 - Canadian Raising with lexical exceptions: Stratal and lexical constraints
- Let’s implement this in a formal learner and try it out!

Learning

Learning framework

- Recursive Constraint Demotion (Tesar 1996)
- Constraint induction “on demand” in the spirit of Becker (2009), Pater (2010), Round (2017)
- Whenever contradictory ranking requirements:
 - Add Stratal constraints, see if inconsistency is resolved
 - If not, add SM constraints, see if inconsistency is resolved
 - If not, add lexical constraints
- Constraints induced based on those constraints not yet inserted in the ranking

Canadian Raising case studies

- Dataset inspired by Nazarov & Pater (2017):
 - l[ʌɪ]fe, l[ʌɪ]fer, l[aɪ] for, l[ai]ve – surface candidates differ in diphthong height (ai/ʌɪ) and voicing (f/v)
 - l[ʌɪ]t, l[ʌɪr]er, l[ai r]o, l[ai]d – surface candidates differ in diphthong height (ai/ʌɪ) and sonorancy/voicing (t,d,ɾ,r)
 - With lexical exceptions: add c[ai]der, sp[ʌɪ]der
- Constraints:
 - Pro-flapping: *VTV, *ɾ
 - Pro-raising: *aGɾ, *aGC
 - Faithfulness: Ident(voice), Ident(son), Ident(low)

Gran Canaria Spanish case study

- Data taken from Broś (2016):
 - /T/ interacting with final C deletion: pensar[$\rightarrow \emptyset$] [t]ontorias
 - /T/ in interacting environment: de[b]artamiento, yo [b]ienso
 - /D/: lle[β]o, la [β]oca
 - Surface candidates: for each target C, consider {p,b,φ,β,∅} or {t,d,θ,ð,∅}
- Constraints:
 - Pro-voicing: *VT, *{φ,θ,x}
 - Pro-spirantization: *[+cont]D, Onset
 - Pro-final C deletion: *FinalC
 - Faithfulness: Ident(voice), Ident(cont), Max

Results

- Canadian Raising without lexical exceptions:
 - As expected, only Stratal constraints active
 $*VTV, *f >> Id(voice), Id(son) >> *aGC_{\circ Stem} >> Id(low) \dots (>> rest)$
- Gran Canaria Spanish:
 - Unlike expectation, Stratal AND SM constraints active
 $*\phi\theta x, \text{Onset} >> \text{Max} >> *V\text{T}_{\text{Word}} >> *[+c]\text{D}_{*V\text{T}} \dots >> *V\text{T}, *[+c]\text{D} \dots >> \text{Faith}$
- Canadian Raising with lexical exceptions
 - Unlike expectation, Stratal AND SM AND lexical constraints active
 $*VTV, *f >> Id(v), Id(s) >> *aGC_{\circ Stem} >> *VTV_{*aGC} >> *aGC_{spider} >> \dots$

Discussion & Conclusion

What happened?

- SM and lexical constraints not induced when only “earlier” type is needed
 - *Canadian Raising without exceptions*: only Stratal, no SM or lexical constraints
 - *Gran Canaria Spanish*: only Stratal and SM constraints, no lexical constraints
- Unnecessary constraints are used (high-ranked) if learner goes through stage of inducing them:
 - *Gran Canaria Spanish*: Stratal constraints used when only SM constraints needed
 - *Canadian Raising with exceptions*: SM constraints used when only Stratal and lexical constraints needed

Learning framework or concept?

- Recursive Constraint Demotion inserts constraint into ranking whenever constraint correctly predicts winner-loser pairs not yet accounted for
 - Does not compare “ham” and “spam” (relevant vs. useful but irrelevant Cs)
 - “On demand” constraint induction not based on entire grammar, but on grammar fragment built so far
- If we change to a learner without these properties (e.g., based on Nazarov & Smith 2023), will this change?
 - Does this depend on specific case studies?
 - Is there something inherent in the concept that leads to overgenerating hypotheses?

Conclusions

- Proof of concept: possible to mix Stratal, SM, lexical explanations for opacity “on demand”
- Case studies:
 - Canadian Raising with/without exceptions (Stratal +/- lexical)
 - Gran Canaria Spanish (SM)
- Full range of model (Stratal HS-SMR with indexed constraints) not always needed
- However, some unnecessary constraints still induced – needs further study (inherent in problem or artifact of implementation?)

Thank you!

References

- Becker, Michael. 2009. Phonological Trends in the Lexicon: The Role of Constraints. PhD dissertation, University of Massachusetts Amherst.
- Bermúdez-Otero, Ricardo. 1999. Constraint interaction in language change: quantity in English and Germanic. PhD dissertation, University of Manchester.
- Bermúdez-Otero, Ricardo. 2003. The acquisition of phonological opacity. In Jennifer Spenader, Anders Eriksson & Östen Dahl (eds.), *Variation within Optimality Theory: Proceedings of the Stockholm Workshop on Variation within Optimality Theory*. Stockholm University, 25-36.
- Broś, Karolina. 2016. Stratum junctures and counterfeeding: Against the current formulation of cyclicity in Stratal OT. In Christopher Hammerly and Brandon Prickett (eds.), *Proceedings of the 46th Annual Meeting of the North East Linguistic Society*, 157-170.
- Farris-Trimble, Ashley and Anne-Michelle Tesser. 2019. "The effect of allophonic processes on word recognition: Eye-tracking evidence from Canadian raising" *Language*. 95(1). e136-e160.
- Idsardi, William J. 2006. Canadian raising, opacity, and rephonemization. *Canadian Journal of Linguistics* 51: 21–28.
- Jarosz, Gaja. 2014. Serial Markedness Reduction. *Proceedings of the 2013 Annual Meeting on Linguistics*. Washington, DC: LSA.
- Joos, Martin. 1942. A phonological dilemma in Canadian English. *Language* 18:141–144.
- Kiparsky, Paul. 1971. Historical Linguistics. In Dingwall, William O. (ed.), *A Survey of Linguistic Science*. College Park: University of Maryland, 576–649.

References

- Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistics Review*, 17 : 351-366.
- McCarthy, John. 2007. Hidden generalizations: Phonological opacity in Optimality Theory. London: Equinox.
- Mielke, Jeff, Mike Armstrong, and Elizabeth Hume. 2003. Looking through opacity. *Theoretical linguistics*, 29(1-2):123–139
- Pater, Joe. 2000. Nonuniformity in English stress: the role of ranked and lexically specific constraints. *Phonology*, 17(2): 237-274.
- Pater, Joe. 2010. Morpheme-Specific Phonology: Constraint Indexation and Inconsistency Resolution. In Steve Parker, (ed.) *Phonological Argumentation: Essays on Evidence and Motivation*. London: Equinox. 123-154.
- Nazarov, Aleksei and Joe Pater. 2017. Learning opacity in Stratal Maximum Entropy Grammar. *Phonology*, 34(2): 299-324.
- Round, Erich. 2017. Phonological exceptionality is localized to phonological elements: The argument from learnability and Yidiny word-final deletion. In Claire Bowern, Larry Horn, Raffaella Zanuttini (eds.), *On looking into words (and beyond): Structures, relations, analyses*. Language Science Press, 59–97.
- Tesar, Bruce. 1996. Computational Optimality Theory. PhD dissertation, University of Colorado Boulder.
- Vance, Timothy. 1987. ‘Canadian Raising’ in some dialects of the northern United States. *American Speech*, 62:195–210.