Universiteit Utrecht

mfm 2021

Modelling the exceptionality/opacity dilemma in acquiring Bedouin Arabic

Aleksei Nazarov
Utrecht University

The problem

- Phonological patterns can be partially contradicted by surface forms (McCarthy 1999)

Dutch:	/kanɔn/	/kanɔ́n/	
σ → σ́ / __σ#	[kánɔn]	[kanɔ́n]	
	‘canon’	‘cannon’	
Dutch:	/rad/	/rad-ən/	/rad ən/
d → t / __]	[rat]	[ra.dən]	[ra.tən]
d → d / else	‘guess!’	‘to guess’	‘guess a … !’

- One possible reason: Lexically stored patterns (**exceptionality**)
    - The word /kanɔ́n/ does not undergo the stress rule, the word /kanɔn/ does
- Another possible reason: Interaction between rules (**opacity**):
    - The Dutch devoicing rule does NOT consider the output of resyllabification (VC + V to V.CV)

The problem

- Phonological patterns active in a language can be partially contradicted by surface forms
- How does a learner learn whether a “contradicted” rule is lexically or grammatically conditioned?
- Here: computational approach

1. Case study of Bedouin Arabic (McCarthy 2007)
   - Multiple opacity, grammatical conditioning
2. Framework that allows lexical or grammatical analysis of “contradicted” rules via same mechanism (Nazarov 2019)
   - Optimality Theory with constraints indexed to individual segments in the lexicon
3. Learner (Round 2017):
   - Analysis grammatical conditioning can be found
   - Assumptions about dataset and underlying form learning influence success

Universiteit Utrecht

Case study: Bedouin Arabic

Al-Mozainy (1981), McCarthy (2007)

Bedouin Arabic data

- Counterfeeding between three rules

1. Syncope — delete high vowel ([i]) before Consonant+Vowel — CiCV → CCV

/kitib-at/ → kitbat	‘it_F was written’		
/kitib/ → ktib	‘it_M was written’		

2. Raising — turn low vowel ([a]) into high vowel ([i]) bef. C+V — CaCV → CiCV *→ CCV

/katab-at/ → k(i)tibat	‘she wrote’	*kitbat	(Syncope counterfed)
/katab/ → kitab	‘he wrote’	*ktab	

3. Epenthesis — insert [i] to break up CR cluster at word end — CiCC → CiCVC *→ CCVC; CaCC → CaCVC *→ CiCVC

/gabl/ → gabil	‘before’	*gibil	(Raising counterfed)
/libn/ → libin	‘clay’	*lbin	(Syncope counterfed)

Lexical vs. Grammatical analysis

- This pattern is productive (e.g., generalization to loanwords, language games; McCarthy 2007)
  - Evidence for grammatical analysis
- **Main question: How is grammatical analysis of these data found during acquisition?**
- But first: what does this analysis look like?
  - Lexical analysis:
    - Raising and Syncope each only in certain (combinations of) arbitrarily marked morphemes
  - Grammatical analysis:
    - Raising only in vowels that alternate with [a], and before C + underlying vowel
    - Syncope only for vowels that alternate with [i], and before C + underlying vowel

Universiteit Utrecht

Framework: OT with constraints tied to input segments

Indexation

- One of major approaches to lexical exceptions in phonology: diacritics or **indices** (Chomsky & Halle 1968; Pater 2000, 2010)

- Pater (2000, 2010) formulated this in Optimality Theory:
  - Morphemes may receive indices/diacritics in the lexicon
    /matste/ /napspe$_i$/ (has index *i*)
  - Constraints on all inputs: *CCC vs. Constraints on inputs with some index: *CCC$_i$

/matste/	***CCC$_i$**	***CCC**
matste		*
/napspe$_i$/		
napspe	*	*

- Round (2017): indices can be on individual segments (cf. Chomsky & Halle 1968)
  - For instance: one vowel in a word can be marked as deleting and another vowel can be marked as raising: sa$_{[+raising]}$mi$_{[+deleting]}$ʕ
  - Allows for either grammatical or lexical account of “contradictions”

Analysis similar to Turbidity (Goldrick 2001) & Coloured Containment (van Oostendorp 2008); See also Boersma (2007), Ettlinger (2008), Hauser et al. (2016),

Nazarov (2020)

Lexical vs. Grammatical analysis of Bedouin Arabic

Lexical analysis

$/ka_j ta_k b/$ $/\text{-}a_p t/$
$/ki_m ti_n b/$
/gabl/
/libn/

- Delete vowels {m,n} before C + vowels {n,p}

$/ki_m ti_n b/ \rightarrow [kti_n b]$ $/ki_m ti_n b\text{-}a_p t/ \rightarrow [ki_m tba_p t]$

- Raise vowels {j,k} before C + vowels {k,p}

$/ka_j ta_k b/ \rightarrow [ki_j ta_k b]$ $/ka_j ta_k b\text{-}a_p t/ \rightarrow [ki_j ti_k ba_p t]$

- Vowels without markings: no raising or deletion

/libn/ → [libin] /gabl/ → [gabil]

Grammatical analysis

$/ka_L ta_{L,V} b/$ $/\text{-}a_{L,V} t/$
$/ki_{\text{-}L} ti_{\text{-}L,V} b/$
$/ga_L bl/$
$/li_{\text{-}L} bn/$

- Delete [-L]-vowels before C + [+V]-vowel

$/ki_{\text{-}L} ti_{\text{-}L,V} b/ \rightarrow [ktib]$ $/ki_{\text{-}L} ti_{\text{-}L,V} b\text{-}a_{L,V} t/ \rightarrow [kitbat]$

- Raise [+L]-vowels before C + [+V]-vowel

$/ka_L ta_{L,V} b/ \rightarrow [kitab]$ $/ka_L ta_{L,V} b\text{-}a_V t/ \rightarrow [kitibat]$

- Elsewhere, all $/V_{[+L]}/ \rightarrow [a]$, all $/V_{[\text{-}L]}/ \rightarrow [i]$

$/li_{\text{-}L} bn/ \rightarrow [libi_{\text{-}V} n]$ $/ga_L bl/ \rightarrow [gabi_{\text{-}V} l]$
(epenthetic vowels are [-V], so no raising or deletion)

Universiteit Utrecht

Learner: Contradiction leads to indexation

Learning (local) indexation

- To learn indexation analysis:
  - Start with universal constraints and unindexed inputs
  - Determine which constraints & which input segments have which index
  - Rank all constraints
- Here: Round's (2017) algorithm:
  - Operates within Biased Constraint Demotion (BCD; Prince & Tesar 2004)
  - When BCD runs into a contradiction in ranking requirements ("inconsistency"):
    - finds most generally applicable indexed constraint & corresponding indexed segments
    - add constraint to analysis and continue BCD
      (cf. Pater 2010)

Evaluating grammatical vs. Lexical analysis

- Grammars evaluated on whether analysis is **grammatical** (in my sense)
  - The vowels that are required to raise before C+V are the same vowels that are required to surface as [a] elsewhere
  - The vowels that are required to delete before C+V are the same vowels that are required to surface as [i] elsewhere
- Assessment in practice:
  - Analysis is grammatical (not lexical) if:
    - Raising segments indexed to both pro-raising constraint and pro-[a] constraint
    - Deleting segments indexed to both pro-deletion constraint and pro-[i] constraint
  - Only alternating vowels considered

Universiteit Utrecht

Simulations & Results

Simulation setup

- Two data sets (what does learner need to see?):

D_1 (no alternations)	D_2 (yes alternations)
kitab	kit**a**b ~ k(i)t**i**bat
ktib	kt**i**b ~ k**i**tbat
simiʕ	s**i**m**i**ʕ ~ s**a**mʕ**a**t
gabil	gabil
libin	libin

Constraints used:

*a, *i, *V
*CCC, *CC[+son]#
*aCV, *VCV
*aCi, *iCa,
Ident(high)
Max(V)
Max(V)/ClosedSyll
Dep(high)

- Two UR hypothesis sets (what does learner need to assume?):

UR_1 (URs with low/high Vs)	UR_2 (canonical URs only)
$/k\begin{Bmatrix}a\\i\end{Bmatrix}t\begin{Bmatrix}a\\i\end{Bmatrix}b/ \rightarrow kitab$	$/katab/ \rightarrow kitab$

For multiple URs during learning, see Jarosz (2006)

Universiteit Utrecht

Results

- Four conditions (data set x UR hypothesis) considered, 10 runs
  - All runs lead to some consistent analysis of the data
  - Table: is **grammatical** analysis found in this condition?
    - Only alternating vowels considered for determining lexical vs. grammatical analysis

	UR_1: all surf cands = URs	UR_2: "correct" URs
D_1: no alternations	0	0
D_2: yes alternations	7	0

- Grammatical analysis found for dataset D_2 and UR_1
  - True alternations a~i, i~∅ necessary
  - Uncertainty about underlying form of [a] and [i] necessary (present in UR_1 but not in UR_2)

Example of grammatical analysis found

$*CC_{[+son]}\# \gg \text{Dep(high)} \gg$

$*a_A CV_V \gg$

$\text{Max(V)}_M \gg *VCV \gg \text{Max(V)} \gg$

$*i_L \gg *a \gg \text{Ident(high)}$

Showing only index markings explicitly assigned by learner;
All segments unmarked for index A are [-A] by default, ditto for L,M,V

/k $a_{A,M}$ t $\mathbf{a}_{A,L,M,V}$ b/	kitab ~ kitib+a_Vt
/k $\mathbf{i}_{A,-L,-M}$ t $\mathbf{i}_{A,-L,-M,V}$ b/	k∅tib ~ kit∅b+a_Vt
/s $\mathbf{a}_{A,L,M}$ m $\mathbf{i}_{A,-L,-M,V}$ ʕ/	simiʕ ~ sam∅ʕ+ a_Vt
/g $a_{L,M}$ bl/	gabil
/l $i_{-L,M}$ bn/	libin

Raising segments indexed to both pro-raising constraint (A) and pro-[a] constraint (L) (rule: L ⟶ A)

Deleting segments indexed to both pro-deletion constraint (-M) and pro-[i] constraint (-L) (rule: -M ⟶ -L)

Epenthesis

- No alternations w.r.t. epenthesis considered here
  - Not provided in data
  - Difficulties with implementation
- Therefore, grammatical analysis for epenthesis not assessed
- Grammatical analysis found in simulations: Raising before epenthetic vowels blocked (*$a_A CV_V$, epenthetic vowels cannot be [+V])
  - However, analysis overgenerates: underlying [-V] vowels can occur anywhere, also outside epenthesis context
  - In addition, syncope before epenthetic vowels is not blocked
- To find fully grammatical analysis of epenthesis: consider alternations in terms of epenthesis, vary URs in terms of presence of epenthetic vowels

Universiteit Utrecht

What can we conclude?

Discussion & Conclusion

- Choosing grammatical analysis of Bedouin Arabic: possible even when lexical analysis readily available
- Conditions:
  - learner has access to crucial alternations
  - learner has not yet finished determining URs of crucial segments (Hayes 2004, Prince & Tesar 2004)
- Very specific time for discovering opaque analyses in course of acquisition
- Follow-up:
  - Epenthesis in BA + other types of opacity (cf. Canadian Raising, Nazarov 2020)
  - Implement generalization of indices (L → A; -M → -L)
  - Probabilistic learner (cf. Nazarov 2018)
  - Other assumptions about URs, data, constraints, etc.

Thank you!

Universiteit Utrecht

References

- **Al-Mozainy**, Hamza Q. 1981. Vowel Alternations in a Bedouin Hijazi Arabic Dialect: Abstractness and Stress. PhD dissertation, University of Texas Austin.
- **Boersma**, P. 2007. Some listener-oriented accounts of h-aspiré in French. Lingua 117(12): 1989-2054.
- **Chomsky**, N. & M. Halle. 1968. The sound pattern of English. New York, NY: Evanston.
- **Ettlinger**, M. 2008. Input-Driven Opacity. PhD dissertation, University of California Berkeley.
- **Goldrick**, M. 2001. Turbid output representations and the unity of opacity. In M. Hirotani, A. Coetzee, N. Hall & J.-Y. Kim (eds.), Proceedings of the Northeast Linguistic Society 30, 231–45. Amherst, MA: GLSA.
- **Hauser**, I., C. Hughto & M. Somerday. 2016. Faith-UO: Counterfeeding in Harmonic Seralism. Proceedings of the 2015 Annual Meetings on Phonology. Washington, DC: LSA.
- **Hauser**, I. & C. Hughto. 2020. Analyzing opacity with contextual faithfulness constraints. Glossa 5(1): 82.
- **Jarosz**, G. 2006. Rich Lexicons and Restrictive Grammars: Maximum Likelihood Learning in Optimality Theory. PhD dissertation, Johns Hopkins University.
- **McCarthy**, J. 1999. Sympathy and phonological opacity. Phonology 16(3): 331-99.
- **McCarthy**, J. 2007. Hidden generalizations: Phonological opacity in Optimality Theory. London: Equinox.
- **Nazarov**, A. 2018. Learning within- and between-word variation in probabilistic OT grammars. In G. Gallagher, M. Gouskova & S. Yin (eds.), Supplemental Proceedings of the 2017 Annual Meeting on Phonology, Washington, DC: LSA.
- **Nazarov**, A. 2020. Bedouin Arabic multiple opacity with indexed constraints in Parallel OT. In H. Baek, C. Takahashi & A.H.-L. Yeung (eds.), Supplemental Proceedings of the 2019 Annual Meeting on Phonology, Washington, DC: LSA.
- **Nazarov**, A. 2021. Learnability of indexed constraint analyses of phonological opacity. Proceedings of the Society for Computation in Linguistics 4: 16.
- **Van Oostendorp**, M. 2008. Incomplete devoicing in formal phonology. Lingua 118: 1362–74.
- **Pater**, J. 2000. Non-uniformity in English secondary stress: The role of ranked and lexically specific constraints. Phonology 17(2): 237–74.
- **Pater**, J. 2010. Morpheme-specific phonology: Constraint indexation and inconsistency resolution. In S. Parker (ed.), Phonological argumentation: Essays on evidence and motivation, 123–54. Equinox.
- **Prince**, A. & P. Smolensky. 1993/2004. Optimality Theory: Constraint interaction in generative grammar. Malden, MA: Blackwell.
- **Round**, E. 2017. Phonological exceptionality is localized to phonological elements: The argument from learnability and Yidiny word-final deletion. In C. Bowern, L. Horn, R. Zanuttini (eds.), On looking into words (and beyond): Structures, relations, analyses, 59–97. Language Science Press.

Utrecht University

DISCLAIMER

Although every effort has been made to ensure that all information in this presentation is correct and up to date, Utrecht University cannot be held liable for any false, inaccurate or incomplete information presented herein.

Appendix

Indexation

- Indexation: universal constraints may have variants specific to some inputs (morphemes/words) only (Pater 2000)

  *[+voice]: no voiced segments
  $*[+voice]_i$: no voiced segments in *i* words

- *Extended indexation*: indices are local to specific segments (Round 2017) and represented as binary; <u>basically same as SPE diacritics</u>

  $*[+voice]_{[+i]}$: no voiced [+i] segments ✓$[d]_{[-i]}$ $*[d]_{[+i]}$

Grammatical analysis tableaux (real words; Nazarov 2020)

/ka$_{[+L, +V]}$ta$_{[+L, +V]}$b/	*CC$_{[+son]}$#	MAX(C)	*V$_{[-L]}$CV$_{[+V]}$	DEP	*a$_{[+L]}$CV$_{[+V]}$	MAX(V)	MAX(V) /CLSSYLL	IDENT(hi)
a. ka$_{[+L, +V]}$ta$_{[+L, +V]}$b					*!			
b. ☞ ki$_{[+L, +V]}$ta$_{[+L, +V]}$b								*
c. k∅$_{[+L, +V]}$ta$_{[+L, +V]}$b						*!		
d. ka$_{[+L, +V]}$ti$_{[+L, +V]}$b					*!			*
e. ka$_{[+L, +V]}$t∅$_{[+L, +V]}$b						*!	*	
f. ka$_{[+L, +V]}$t . ta$_{[+L, +V]}$b				*!				

/ga$_{[+L, +V]}$bl/	*CC$_{[+son]}$#	MAX(C)	*V$_{[-L]}$CV$_{[+V]}$	DEP	*a$_{[+L]}$CV$_{[+V]}$	MAX(V)	MAX(V) /CLSSYLL	IDENT(hi)
a. ga$_{[+L, +V]}$bl	*!							
b. ☞ ga$_{[+L, +V]}$bi$_{[-L, -V]}$l				*				
c. gi$_{[+L, +V]}$bi$_{[-L, -V]}$l				*				*!
d. ga$_{[+L, +V]}$b		*!						

Grammatical analysis tableau (ROTB; Nazarov 2020)

/$\text{ba}_{[-L,+V]}\text{ta}_{[-L,+V]}$/	$*CC_{[+son]}\#$	$*V_{[-V]}$	$*V_{[-L]}CV_{[+V]}$	MAX(V)	$*a_{[+L]}CV_{[+V]}$	$*[-hi]_{[-L]}$	$*[+hi]_{[+L]}$	IDENT(hi)
a. $\text{ba}_{[-L,+V]}\text{ta}_{[-L,+V]}$			*!			**		
b. $\text{b}\varnothing_{[-L,+V]}\text{ta}_{[-L,+V]}$				*		*!		
c. ☞ $\text{b}\varnothing_{[-L,+V]}\text{ti}_{[-L,+V]}$				*				*
/$\text{ti}_{[+L,+V]}\text{ki}_{[+L,+V]}$/								
d. $\text{ti}_{[+L,+V]}\text{ki}_{[+L,+V]}$							**!	
e. $\text{ta}_{[+L,+V]}\text{ka}_{[+L,+V]}$					*!			**
f. ☞ $\text{ti}_{[+L,+V]}\text{ka}_{[+L,+V]}$							*	*
/$\text{bi}_{[+L,+V]}\text{ki}_{[+L,-V]}$/								
g. $\text{bi}_{[+L,+V]}\text{ki}_{[+L,-V]}$		*!						
h. ☞ $\text{bi}_{[+L,+V]}\text{k}\varnothing_{[+L,-V]}$				*				
/bn/								
i. bn	*!							
j. ☞ $\text{bi}_{[-L,-V]}\text{n}$		*						

Current model vs. previous proposals

- Current model depends on UR information (indices) for dealing with opacity (Nazarov 2019, 2020)
  - Unpronounced information present in SRs like in:
    - Turbidity Theory (Goldrick 2001)
    - Coloured Containment (Van Oostendorp 2008)
    - Diagonal Correspondence Theory (Ettlinger 2008)
    - Bidirectional OT (see Boersma 2007 for an account of opacity)
  - Refers directly to some information from UR as in UO-Faith approach (Hauser et al. 2016)
    - Current approach does not need dedicated type of constraint for opacity only (though see Hauser & Hughto 2020 for an update): indexation independently needed for exceptions

Cannot handle multi-level opacity like in Bedouin Arabic

Do not automatically allow account of exceptions

- Main point: current model has similar solutions for opacity and exceptions
- Choice between lexical & grammatical analysis can be modelled with minimal confounds

Round’s (2017) learner in more detail

- Round (2017): model to learn segmentally local indexation from winner-loser pair data
  1. Based on Biased Constraint Demotion (BCD, Tesar and Smolensky 2004)
  2. Whenever two inputs in the data have conflicting ranking requirements (= inconsistency): induce some indexed constraint (Pater 2010)
  3. Which indexed constraint assigned to which segments? *(new contribution)* Selected based on number and location of Winner-preferring violation loci

Biased Constraint Demotion

- Version of Recursive Constraint Demotion (Tesar 1995) with a Markedness-over-Faithfulness bias
  - Start with no ranking
  - At each step, select only those constraints that prefer no losers = PNL
  - Out of PNL, take just the Markedness constraints and install them at the bottom of the ranking
    - If there are no Markedness constraints, select the smallest set of Faithfulness constraints that will “free up” a Markedness constraint at the next step
  - Remove from consideration all winner-loser pairs that have a W mark for one of the freshly installed constraints

Inconsistency

- BCD is dependent on constraints without L marks
- When there are no such constraints, this means something's wrong
  - Cues mutually **inconsistent** rankings, e.g.,
    input 1 wants A >> B, input 2 wants B >> A: mutually incompatible

- Pater (2010): when you encounter inconsistency, induce some indexed constraint
  - Gets you out of inconsistency:
    input 1 wants A*i* >> B, input 2 wants B >> A: mutually compatible!

Indexed constraint selection

- Round (2017) wants model that infers segmentally local indices
- Therefore: violations track segment instances ("loci")

  **[+voice] has a W violation in the second "b" for W-L pair 1, but a L violation in the first "v" for W-L pair 2*

- For each constraint, compute:
  - Φ_W: Set of segment instances that get a W violation of that constraint
  - Φ_L: Set of segment instances that get a L violation of that constraint
  - $\Phi_W - \Phi_L$: Set of segment instances that get a W violation but never a L violation of that constraint

Indexed constraint selection

- Round (2017) wants model that infers segmentally local indices

- For each constraint, compute:
  - Φ_W: Set of segment instances that get a W violation of that constraint
  - Φ_L: Set of segment instances that get a L violation of that constraint
  - $\Phi_W - \Phi_L$: Set of segment instances that get a W violation but never a L violation of that constraint

- Make [+i] indexed version of constraint that has the greatest $\Phi_W - \Phi_L$:
  - $\Phi_W - \Phi_L$ become [+i], all other segments instances become [-i] *(binarity: AN)*

Utrecht University

DISCLAIMER