A Computational Method for Evaluating Theories of Phonological Representation

Nick Danis, Eileen Blum, Luca Iacoponi, Hazel Mitchley, and Adam Jardine

RUTGERS

Chicago Linguistic Society 53
May 26, 2017

Overview

- Phonologists have long argued that phonological grammars should express **natural** generalizations as a result of **simple** rules or constraints
- We define 'simple' in terms of an independently motivated notion of computational complexity
- In formal language theory, simple means **small**, **connected** substructures

Overview

- We ask of a representational theory:
  - Does it express natural generalizations with small, connected substructures?
  - Does it express **un**natural generalizations with larger structures?
- We establish that features and tiers differentiate natural and unnatural processes by this metric
- Larger point: unifying computation and phonological information
  - How are representations organized such that salient properties of sounds are connected?

Naturalness vs. simplicity

- Phonologists have long argued that phonological grammars should express **natural** processes as a result of **simple** rules or constraints

Naturalness

- **Empirical** property
- Typologically frequent
- Phonetically grounded
- e.g. assimilation, dissimilation

Simplicity

- **Representational** property
- ‘Fewer symbols’
- Restricts arbitrariness

Naturalnes vs. simplicity

- Halle (1962, p. 381–2):

$$R_1\text{: k} \rightarrow \text{ tʃ / \_\_} \left\{ \begin{array}{c} \text{i} \\ \text{e} \\ \text{æ} \end{array} \right\} \quad R_2\text{: k} \rightarrow \text{ tʃ / \_\_} \left\{ \begin{array}{c} \text{p} \\ \text{r} \\ \text{a} \end{array} \right\}$$

- Hyman (1975, p.104): "[S]implicity can be quantified by counting features, and only a theory which requires that segments are composites of features will differentiate between real and spurious generalizations"

Simplicity (in detail)

- What and how should we be counting?
- Minimum description of pattern depends on how grammar is encoded (Rogers et al., 2013)
- How do we encode **non-linear representations** (Kornai, 1995)?
- Complexity classes of patterns offer an encoding-independent notion of simplicity (Rogers et al., 2013)

Simplicity (in detail)

- In hierarchy of Strictly k-Local formal language classes (SL_k; McNaughton and Papert, 1971), complexity of pattern corresponds to its ***k*-value**
- The k-value is the size of the forbidden piece of string **connected** by adjacency
- **td* is SL_2

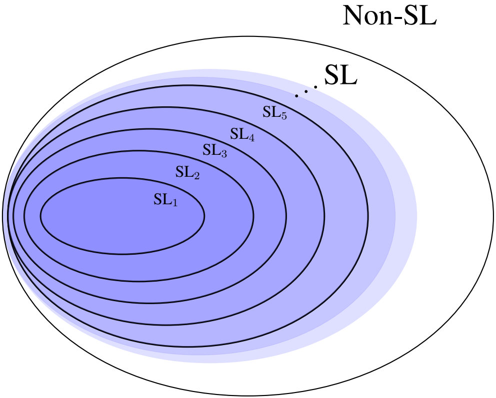
$$\{ata, ada, utu, uda, atta, adda, uttu, ...\} \qquad \text{(no } atda, utda, ...\text{)}$$

- **utd* is SL_3

$$\{ata, ada, utu, uda, atta, adda, uttu, atda...\} \quad (atda, \text{ but no } utda)$$

Simplicity (in detail)

- $\mathrm{SL}_1 \subsetneq \mathrm{SL}_2 \subsetneq \mathrm{SL}_3 \subsetneq \cdots \mathrm{SL}_k \subsetneq \cdots \subsetneq \mathrm{SL}$



- This applies to Strictly Piecewise classes (**s*...*ʃ*) as well (Rogers et al., 2010)

Naturalness distinctions in k-values

- We extend notion of k-value to graphs representing nonlinear phonological structures (Jardine, 2016)
- Goal of representational theory: k for natural constraint is less than k for **un**natural constraint
- For example, *[-voi][+voi] is common, while *[ma] is not
- Both are SL_2:
  - $*\{td, dt, tb, bt, pb, ..., sz\}$
  - $*ma$

Naturalness distinctions in k-values

String representations

$\mathrm{T} \rightarrow \mathrm{D}$

*TD

natural

$k = 2$

$\mathrm{m} \rightarrow \mathrm{a}$

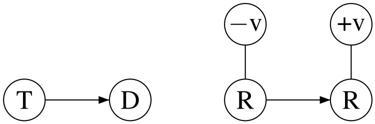
*ma

unnatural

$k = 2$

Naturalness distinctions in *k*-values

String versus featural representations



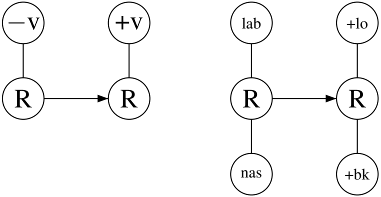
*[-voi][+voi]

string	**features**
$k = 2$	$k = 4$

- Features in a "bottle brush" representation (Hayes, 1990) with no order on tier (Kaye, 1985)

Naturalness distinctions in k-values

Featural representations



*[−voi][+voi]	*ma
natural	**unnatural**
$k = 4$	$k \geq 6$

Naturalness distinctions in *k*-values

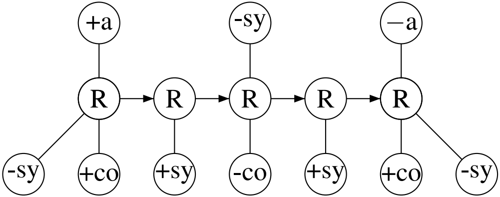
- Formal support for Chomsky and Halle (1968)'s idea of feature-counting
- We can independently support other representational primitives, i.e. autosegmental tiers (order relation between like features)

Naturalness distinctions in k-values

- Navajo (Cook 1978): *[+ant]...[−ant]   (Strictly Piecewise)
- Constraints against arbitrary features, e.g., *[+ant]...[−voi] are unattested
- Constraint against different values of same feature is **natural**, against different features is **unnatural**

Naturalness distinctions in *k*-values

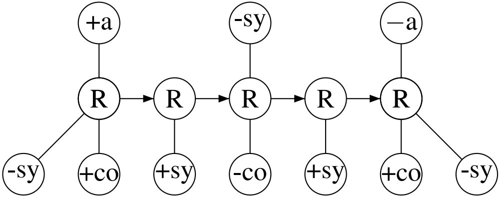
- Current assumption: no order between like features



[s o b o ʃ]

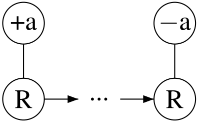
Naturalness distinctions in *k*-values

- Current assumption: no order between like features



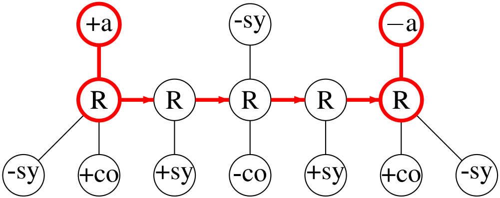
[s o b o ʃ]

- *[+ant]...[−ant]



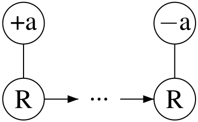
Naturalness distinctions in k-values

- Current assumption: no order between like features



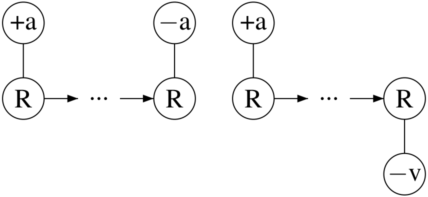
[s o b o ʃ]

- *[+ant]...[−ant]



Naturalness distinctions in *k*-values

Order only on root tier

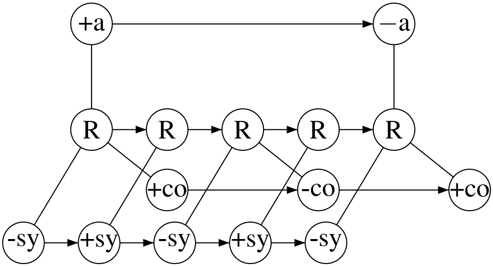


*[+ant]...[−ant]
natural
$k = 4$

*[+ant]...[−voi]
unnatural
$k = 4$

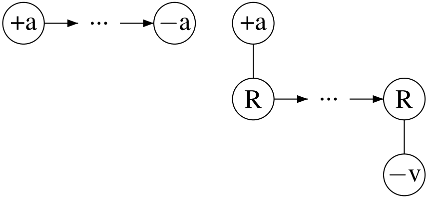
Naturalness distinctions in *k*-values

Adding order between like features



Naturalness distinctions in k-values

Order on feature tiers



*[+ant]...[−ant]	*[+ant]...[−voi]
natural	**unnatural**
$k = 2$	$k = 4$

Discussion

- Independent motivation for phonological tiers
- More general idea: relations in autosegmental structure **connect** natural classes of features
- How does feature geometry (Sagey, 1986; Clements, 1991; Clements and Hume, 1995) accomplish this?
- How to apply same metric for mappings? (Chandlee, 2014; Chandlee and Lindell, prep)
- How can this reduce search space of constraints for a learner (c.f. Hayes and Wilson, 2008)?

Conclusion

- ‘Simple’ constraints refer to **small, connected** pieces of structures
- This prefers representations organized such that
  - natural classes of features are closely connected
  - unnatural classes require traversing many points in representation
- Natural constraints are less cognitively complex than unnatural constraints
- Step towards unifying formal complexity and phonological substance

Thank You

We also thank the attendees of the Rutgers/Delaware/Haverford Computational Phonology Workshop for their questions and comments.

References I

Chandlee, J. (2014). *Strictly Local Phonological Processes*. PhD thesis, University of Delaware.

Chandlee, J. and Lindell, S. (in prep.). A logical characterization of strictly local functions. In Heinz, J., editor, *Doing Computational Phonology*. OUP.

Chomsky, N. and Halle, M. (1968). *The Sound Pattern of English*. Harper & Row.

Clements, G. N. (1991). Place of articulation in consonants and vowels: a unified theory. *Working Papers of the Cornell Phonetics Laboratory*, 5:77–123.

Clements, G. N. and Hume, E. V. (1995). The internal organization of speech sounds. In Goldsmith, J., editor, *The handbook of phonological theory*, pages 245–306. Oxford: Blackwell.

Hayes, B. (1990). Diphthongisation and Coindexing. *Phonology*, 7(1):31–71.

References II

Hayes, B. and Wilson, C. (2008). A maximum entropy model of phonotactics and phonotactic learning. *Linguistic Inquiry*, 39:379–440.

Hyman, L. (1975). *Phonology: Theory and Analysis*. New York: Holt, Rinehart and Winston.

Jardine, A. (2016). *Locality and non-linear representations in tonal phonology*. PhD thesis, University of Delaware.

Kaye, J. D. (1985). On the syllable structure of certain west african languages. In Goyvaerts, D. L., editor, *African Linguistics: Essays in Memory of M. W. K. Semikenke*. Amsterdam: John Benjamins.

Kornai, A. (1995). *Formal Phonology*. Garland Publication.

McNaughton, R. and Papert, S. (1971). *Counter-Free Automata*. MIT Press.

Rogers, J., Heinz, J., Bailey, G., Edlefsen, M., Visscher, M., Wellcome, D., and Wibel, S. (2010). On languages piecewise testable in the strict sense. In Ebert, C., Jäger, G., and Michaelis, J., editors, *The Mathematics of Language*, volume 6149 of *Lecture Notes in Artifical Intelligence*, pages 255–265. Springer.

References III

Rogers, J., Heinz, J., Fero, M., Hurst, J., Lambert, D., and Wibel, S. (2013). Cognitive and sub-regular complexity. In *Formal Grammar*, volume 8036 of *Lecture Notes in Computer Science*, pages 90–108. Springer.

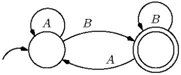
Sagey, E. (1986). *The Representation of Features and Relations in Non-Linear Phonology*. PhD thesis, Massachusetts Institute of Technology.

Appendix: Encoding and description length

Sequences of 'A's and 'B's which end in 'B' (EndB)

Regular Grammar: $S_0 \longrightarrow AS_0,\ S_0 \longrightarrow BS_0,\ S_0 \longrightarrow B$

DFA:

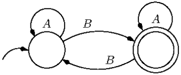


Regular Expression: $(A+B)^*B$

Sequences of 'A's and 'B's which contain an odd number of 'B's (OddB)

Regular Grammar: $S_0 \longrightarrow AS_0,\ S_0 \longrightarrow BS_1,$
$S_1 \longrightarrow AS_1,\ S_1 \longrightarrow BS_0,\ S_1 \longrightarrow \varepsilon$

DFA:



Regular Expression: $(A^*BA^*BA^*)^*A^*BA^*$

Fig. 1. Minimal descriptions: strings which end in 'B' vs. strings with an odd number of 'B's.

(Rogers et al., 2013, p. 93)

Appendix: ‘Simplicity’ in Chomsky and Halle (1968)

> *“It should be observed in this connection that although definition (9) has commonly been referred to as the “simplicity” or “economy condition,” it has never been proposed or intended that the condition defines “simplicity” or “economy” in the very general (and still very poorly understood) sense in which these terms usually appear in writings on the philosophy of science.”*
>
> *(Chomsky and Halle, 1968, pp. 334–335)*