Logical Characterizations of Local vs. Long-distance Phonology

Jane Chandlee (jchandlee@haverford.edu)
Haverford College, Haverford, Pennsylvania, U.S.A.

Central Contribution

Phonological maps from underlying to surface forms can be represented using logical formulae over graph transductions. Doing so provides a very natural extension from local to long-distance phonology, allowing us to build on what is know about the computational nature of local phonology (Heinz, 2007, 2009; Chandlee, 2014) to establish a more complete understanding of what is phonologically possible.

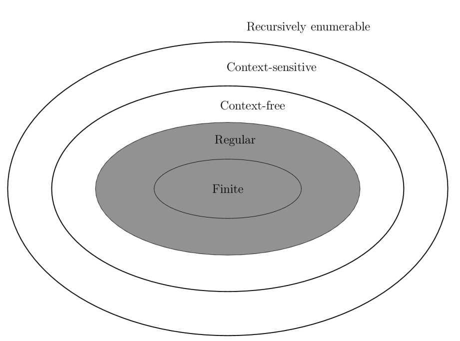
Computational Approach

- Both rule- and constraint-based theories of generative phonology concur on the existence of a map from input (underlying) to output (surface) forms.
- We model these maps with *functions* with the goal of identifying computational properties that are independent of any grammatical formalism (rules or constraints).

(1) Let f be a function which voices obstruents after nasals. Example (Quechua, Pater (2004)) f(kampa) = [kamba], 'yours'

Q: What class of functions does f belong to?

- Identifying the most *restrictive* set of functions needed for phonological maps leads to a better characterization of the components of phonological grammars.
- Previous work showed that phonological maps are regular (Johnson, 1972; Kaplan and Kay, 1994), with substantial evidence indicating they are in fact subregular (Mohri (1997); Heinz and Lai (2013); Chandlee (2014); Jardine (to appear)).

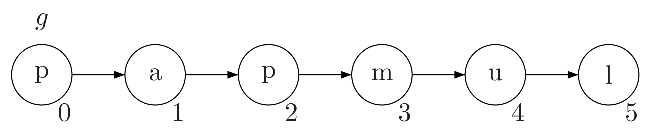


(Chomsky, 1956)

- A variety of formalisms can be used to define the exact same class of functions (i.e., finite-state automatic, logical, algebraic, formal language-theoretic).

Local/Bounded Maps

(2) Korean (Lee and Pater 2008)
/pa**pm**ul/ ↦ [pa**mm**ul] 'rice water'

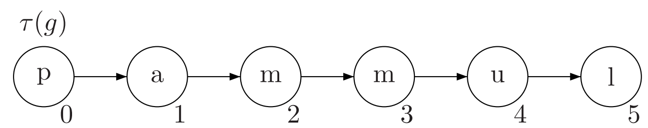


- Set of positions: {0, 1, 2, 3, 4, 5}
- Successor relation: {(0,1), (1,2), (2,3), (3,4), (4,5)}
- Labeling function/segment predicates: $p(x) = \{0, 2\}$, $a(x) = \{1\}$, $m(x) = \{3\}$, $u(x) = \{4\}$, $l(x) = \{5\}$
- First define natural class predicates (e.g., $N(x) \equiv n(x) \vee m(x) \vee$ ŋ(x)...), and then define 'substring predicates': $CN(x) \equiv C(x) \wedge (\exists y)[N(y) \wedge x \triangleleft y]$
- Formulae define the output graph in terms of the input graph (Engelfriet and Hoogeboom, 2001):

$$\varphi^0_N(x) \equiv N(x) \vee CN(x)$$

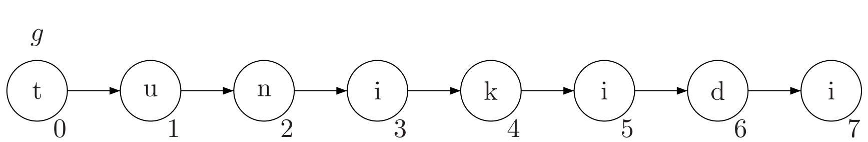
$$\varphi^0_C(x) \equiv C(x) \wedge \neg CN(x)$$

$$\varphi^0_V(x) \equiv V(x)$$



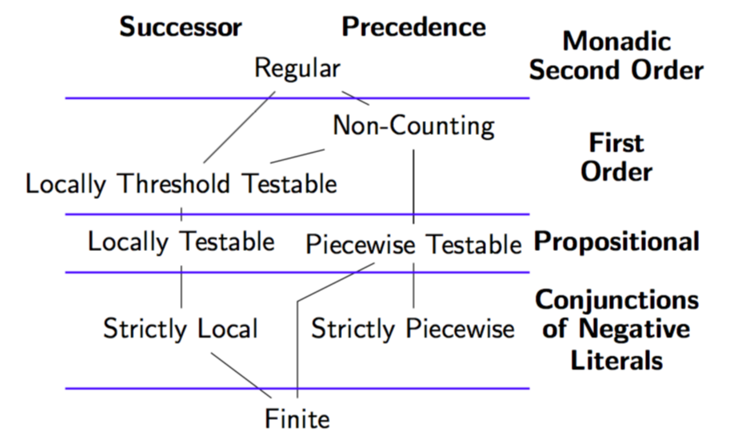
Long-distance/Unbounded Maps

(3) Kikongo (Meinhof 1932, Odden 1994, Rose and Walker 2004)
/tu**n**iki**d**i/ ↦ [tu**n**iki**n**i] 'we ground'



- Following Heinz (2010), the distinction between local and long-distance phonology is a shift from successor to precedence.
- Precedence relation: {(0,1), (0,2)...(1,3)...**(2,6)**}
- Subsequence predicate: $ND(x) \equiv D(x) \wedge (\exists y)[N(y) \wedge y < x]$
- $\varphi^0_N(x) \equiv N(x) \vee ND(x)$

Subregular Hierarchy



(Rogers & Pullum 2011, Rogers et al. 2013)

Conclusions and Open Questions

- Logical characterizations of phonological maps indicate that the difference between local and long-distance is one of representation, not computational complexity.
- Previous findings for phonotactics (Heinz, 2009, 2010) suggest that phonology is restricted to the lowest regions of the subregular hierarchy - what are the equivalent regions for subregular phonological maps?
- Learnability results exist for local maps (Chandlee et al., 2014; Jardine et al., 2014) based on an FST characterization; a corresponding FST for LD maps will lead to a more complete computational learning model.

Selected References

Chandlee, J. (2014). *Strictly Local Phonological Processes*. PhD thesis, U. of Delaware.
Chandlee, J., Eyraud, R., and Heinz, J. (2014). Learning strictly local subsequential transducers. *TACL 2*, pages 491–503.
Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, pages 113–124.
Engelfriet, J. and Hoogeboom, H. J. (2001). Mso definable string transductions and two-way finite state transducers. *ACM Transactions on Computational Logic*, 1(4):1–38.
Heinz, J. (2007). *The Inductive Learning of Phonotactic Patterns.* PhD thesis.
Heinz, J. (2009). On the role of locality in learning stress patterns. *Phonology*, 26:303–351.
Heinz, J. (2010). Learning long-distance phonotactics. *Linguistic Inquiry*, 41:623–661.
Heinz, J. and Lai, R. (2013). Vowel harmony and subsequentiality. In Kornai, A. and Kuhlmann, M., editors, *Proceedings of the 13th Meeting on Mathematics of Language*, Sofia, Bulgaria.
Jardine, A. (to appear). Computationally, tone is different. *Phonology*.
Jardine, A., Chandlee, J., Eryaud, R., and Heinz, J. (2014). Very efficient learning of structured classes of subsequential functions from positive data. In *Proceedings of ICGI 2014*, pages 94–108.
Johnson, C. D. (1972). *Formal aspects of phonological description*. Mouton.
Kaplan, R. and Kay, M. (1994). Regular models of phonological rule systems. *CL*, 20:331–78.
Mohri, M. (1997). Finite-state transducers in language and speech processing. *CL*, 23(2):269–311.