

# A Concatenation Operation to Derive Autosegmental Graphs

Adam Jardine and Jeffrey Heinz

University of Delaware

July 26 2015

# Introduction



- ▶ What kind of structures are autosegmental phonological representations (APRs; Goldsmith, 1976)?
- ▶ Current conception: top-down constraints on graph structures
- ▶ Our result: these fundamental properties emerge from concatenation of an alphabet of primitives (Engelfriet and Vereijken, 1997; Courcelle et al., 2012)
- ▶ This makes them directly comparable to strings

# Introduction

## Overview

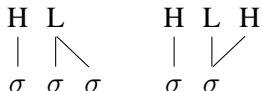
- ▶ Why APRs & what are APRs
- ▶ Representation of APRs as labeled mixed graphs
- ▶ Top-down, axiomatic definition of APRs
- ▶ Constructive derivation of axioms from concatenation & primitives
- ▶ Empirical consequences
- ▶ Discussion

# APRs

- ▶ What are APRs?
- ▶ Mende (Leben, 1973; Goldsmith, 1976): fewer attested tone combinations than possible  
F, HL, HLL, HLLL, ...  
\*HF, \*HLHL, \*HHLL, \*HLLH, ...
- ▶ Answer: tone melodies independent from syllables

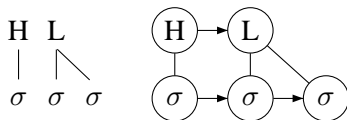


# APRs



- ▶ Fundamentally, APRs are:
  - ▶ Ordered tiers of like autosegments
  - ▶ Association lines connecting items on distinct tiers
- ▶ Set of valid APRs usu. defined *axiomatically* (Goldsmith, 1976; Sagey, 1986; Bird and Klein, 1990; Kornai, 1995)

# APGs

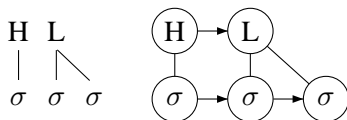


- ▶ To represent APRs, we can use (simple) *labeled mixed graphs*

$$G = \langle V, E, A, \ell \rangle$$

- ▶  $GR(\Sigma)$ 
  - ▶  $V = \{0, 1, \dots, n\}$
  - ▶  $E \subseteq \mathcal{P}_{=2}(V)$
  - ▶  $A \subseteq V \times V$
  - ▶  $\ell : V \rightarrow \Sigma$

# APGs

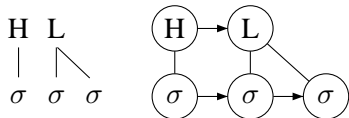


- ▶ To represent APRs, we can use (simple) *labeled mixed graphs*

$$G = \langle V, E, A, \ell \rangle$$

- ▶ Why graphs?
  - ▶ Graphs explicitly encode all information in APRs
  - ▶ Deep literature on pattern matching, characterizing & learning graph sets, graph transductions, etc. (Courcelle et al., 2012; Engelfriet and Hoogeboom, 2001; López et al., 2012; Ferrera, 2013, inter alia)

# APGs



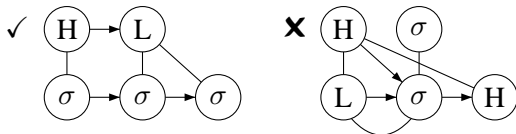
- ▶ To represent APRs, we can use (simple) *labeled mixed graphs*

$$G = \langle V, E, A, \ell \rangle$$

- ▶ To select *well-formed* autosegmental phonological graphs (APGs), there are two options
  - ▶ Axiomatically (top-down approach)
  - ▶ Constructively (bottom-up approach)

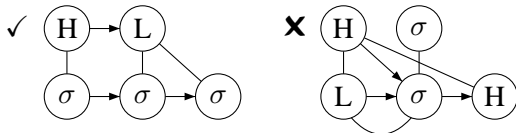


# Axiomatic approach: $APG(\Sigma, T) \subset GR(\Sigma)$



- ▶ We first define the basic autosegmental structure in  $APG(\Sigma, T)$
- ▶ We consider two *tiers*

Axiomatic approach:  $APG(\Sigma, T) \subset GR(\Sigma)$

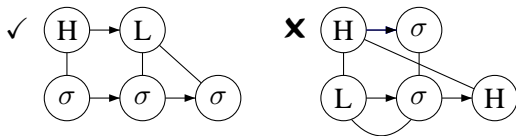


- ▶ Let  $\preceq$  be the reflexive, transitive closure of  $A$

Axiom (Tier structure)

$V$  is partitioned into two sets  $V_0, V_1$  which are each totally ordered by  $\preceq$ .  $V_0$  and  $V_1$  are the **tiers** of  $G$ .

# Axiomatic approach: $APG(\Sigma, T) \subset GR(\Sigma)$



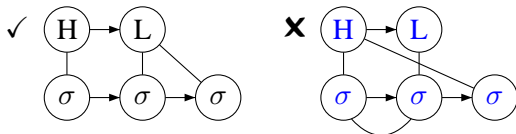
- ▶ Let  $T$  be a partition  $\{T_t, T_m\}$  on  $\Sigma$

Axiom (Like tier labels)

*Partition of  $V$  into tiers respects partition  $T$  on  $\Sigma$*

- ▶ Let  $V_m$  correspond to  $T_m$ , likewise  $V_t, T_t$
- ▶ Let's set  $T_m = \{H, L\}$ ,  $T_t = \{\sigma\}$

Axiomatic approach:  $APG(\Sigma, T) \subset GR(\Sigma)$

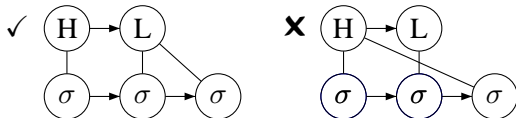


- Associations can only be between tiers

Axiom (Associations between tiers)

For all  $\{x, y\} \in E$ ,  $x \not\approx y$  and  $y \not\approx y$ .

Axiomatic approach:  $APG(\Sigma, T) \subset GR(\Sigma)$

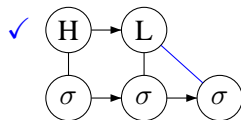
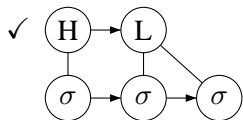


- ▶ No line crossing (Goldsmith, 1976; Hammond, 1988; Coleman and Local, 1991)

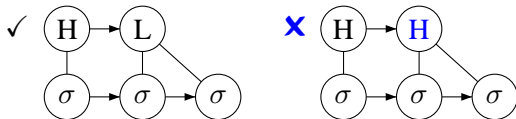
Axiom (NCC)

For all  $u, v, x, y \in V$ , if  $\{u, x\}, \{v, y\} \in E$  and  $u \preceq v$ , then  $x \preceq y$ .

Axiomatic approach:  $APG(\Sigma, T) \subset GR(\Sigma)$



Axiomatic approach:  $APG(\Sigma, T) \subset GR(\Sigma)$



- ▶ The Obligatory Contour Principle Leben (1973); Goldsmith (1976); McCarthy (1986); Odden (1986)

Axiom (OCP)

For tier  $V_m$ , for all  $x, y \in V_m$ ,  $(x, y) \in A$  implies  $\ell(x) \neq \ell(y)$ .

# Axiomatic approach: $APG(\Sigma, T) \subset GR(\Sigma)$

- ▶  $APG(\Sigma, T)$ : graphs in  $GR(\Sigma)$  obeying these axioms

Axioms

1) Tier structure	4) NCC
2) Like tier labels	5) OCP
3) Associations between tiers	

- ▶ These properties can all be derived from concatenation



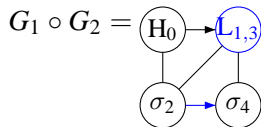
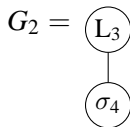
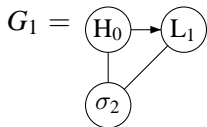
## Constructive approach: $APG(\Gamma) \subset GR(\Sigma)$

- ▶ Engelfriet and Vereijken (1997); Courcelle et al. (2012): two methods of graph concatenation:

*One way is to “glue” them together, by identifying some of their vertices. The other way is to “bridge” them (or rather, “bridge the gap between them”), by adding edges between their vertices. (Courcelle et al., 2012, p. 6)*

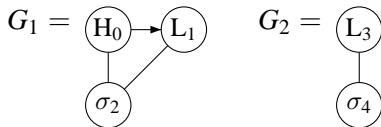
- ▶ We define a specialized version of this machinery (using both “gluing” and “bridging”) to generate the set of APGs

Constructive approach:  $APG(\Gamma) \subset GR(\Sigma)$

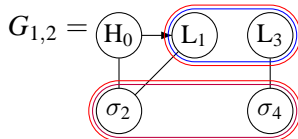


# Constructive approach: $APG(\Gamma) \subset GR(\Sigma)$

## “Gluing” and “bridging”

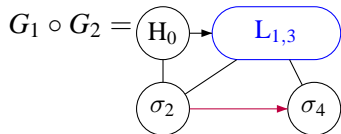
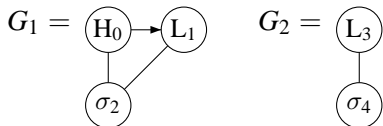


- ▶ Identify the ‘ends’ of the tiers
- ▶ Merge identical end nodes on  $V_m$
- ▶ Draw arcs between any other pair of end nodes



Constructive approach:  $APG(\Gamma) \subset GR(\Sigma)$

“Gluing” and “bridging”



Constructive approach:  $APG(\Gamma) \subset GR(\Sigma)$

### The identity graph

- ▶ Let  $G_\lambda = \langle \emptyset, \emptyset, \emptyset, \emptyset \rangle$

Theorem ( $G_\lambda$  is identity)

For any  $G \in GR(\Sigma)$ ,  $G \circ G_\lambda = G_\lambda \circ G = G$ .

# Constructive approach: $APG(\Gamma) \subset GR(\Sigma)$

## Graph primitives

- ▶ Engelfriet and Vereijken (1997): graph sets can interpret string sets
- ▶ Concatenation of symbols = concatenation of graph primitives

## Definition

An *alphabet of graph primitives* over  $GR(\Sigma)$  is a finite set  $\Gamma$  of symbols and a naming function  $g : \Gamma \rightarrow GR(\Sigma)$ .

## Definition

For  $w \in \Gamma^*$ ,  $g(w) \stackrel{\text{def}}{=}$

- ▶  $G_\lambda$  if  $w = \lambda$
- ▶  $g(u) \circ g(\gamma)$  if  $w = u\gamma, u \in \Gamma^*, \gamma \in \Gamma$

# Constructive approach: $APG(\Gamma) \subset GR(\Sigma)$

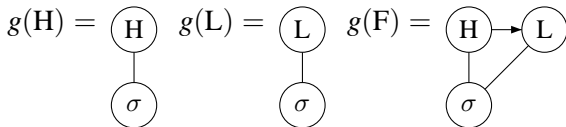
## APG graph primitives

Definition (APG graph primitive)

For  $\Sigma$  and  $T = \{T_t, T_m\}$ , an *APG graph primitive* is a graph  $G \in GR(\Sigma)$  for which

- $V_t$  is a singleton set  $\{v_t\}$
- $V_m$  is totally ordered by  $\preceq$
- All  $e \in E$  are of the form  $\{v_m, v_t\}$ ,  $v_m \in V_m$

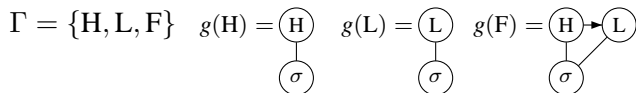
$$\Gamma = \{H, L, F\}$$



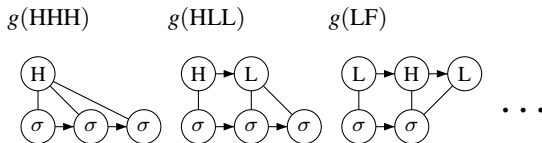
# Constructive approach: $APG(\Gamma) \subset GR(\Sigma)$

## Graph primitives

- Def. of APG prim.
- $V_t = \{v_t\}$
  - $\preceq$  totally orders  $V_m$
  - All  $e \in E = \{v_m, v_t\}$   
for  $v_m \in V_m$



►  $APG(\Gamma) = \{g(w) \mid w \in \Gamma^*\}$





# Constructive approach: $APG(\Gamma) \subset GR(\Sigma)$

APG prim.	a. $V_t = \{v_t\}$	Ax.	1) Tier struct.	4) NCC
	b. $\preceq$ totally orders $V_m$		2) Tier labels	5) OCP
	c. All $e \in E = \{v_m, v_t\}$		3) Assoc.	
	for $v_m \in V_m$			

## Theorem

Every  $G \in APG(\Gamma)$  follows Ax. 1 – 5

- ▶ *APG graph primitives have two tiers  $V_t$  and  $V_m$  and concatenation preserves this*
- ▶ *NCC from singleton  $V_t$*
- ▶ *OCP from merging of melody nodes*

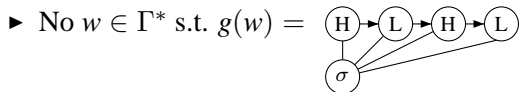
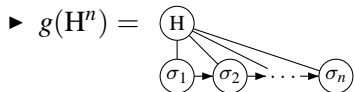
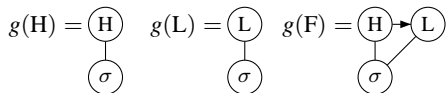
## Theorem

○ *is associative over  $APG(\Gamma)$*

- ▶ *Concatenation preserves tier structure; proof similar of associativity over  $\Gamma^*$*

# Concatenation and empirical phenomena

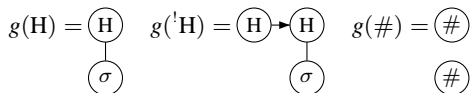
- ▶  $APG(\Gamma)$  also naturally captures facts of phonology
- ▶ Unbounded spreading, but no ‘unbounded contouring’



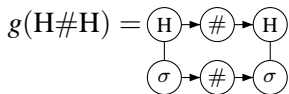
- ▶ (unless added to  $\Gamma$ )

# Concatenation and empirical phenomena

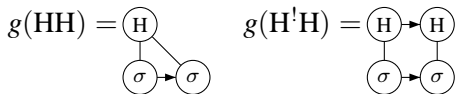
- ▶ Most OCP violations at morpheme boundaries or signaled by downstep (Hyman, 2014)



- ▶ Adjacent Hs across boundaries (Aghem; Hyman, 2014)



- ▶ Adjacent Hs tautomorphemically (Kishambaa; Odden, 1986)



# Discussion

- ▶  $g$  is direct way of relating strings to APRs
- ▶ Cognitive interpretation: how humans relate linear speech stream to APRs
- ▶  $(APG(\Gamma), \circ)$  as free monoid
- ▶ Multi-tier representations:  $n$ -ary partitions on  $\Sigma$
- ▶ Full feature geometry may require more structure on partition
- ▶ Less restrictive version w.r.t. OCP could include second concatenation operation without ‘gluing’

# Conclusion

- ▶ We defined APRs as  $APG(\Gamma)$  parallel to strings
- ▶ Maintains key properties of APRs
- ▶ Positive empirical implications (e.g. no unbounded contouring)
- ▶ *Restrictive* theory of APRs (w.r.t OCP)

# Acknowledgements

We thank three reviewers for their insightful comments and suggestions. Adam Jardine acknowledges support from a University of Delaware Graduate Research Fellowship.

# References I

- Bird, S. and Klein, E. (1990). Phonological events. *Journal of Linguistics*, 26:33–56.
- Coleman, J. and Local, J. (1991). The “No Crossing Constraint” in autosegmental phonology. *Linguistics and Philosophy*, 14:295–338.
- Courcelle, B., Engelfriet, J., and Nivat, M. (2012). *Graph structure and monadic second-order logic: A language-theoretic approach*. Cambridge University Press.
- Engelfriet, J. and Hoogetboom, H. J. (2001). MSO definable string transductions and two-way finite-state transducers. *ACM Transactions on Computational Logic*, 2:216–254.
- Engelfriet, J. and Vereijken, J. J. (1997). Context-free graph grammars and concatenation of graphs. *Acta Informatica*, 34:773–803.
- Ferrera, R. (2013). *Efficiently Listing Combinatorial Patterns in Graphs*. PhD thesis, Università degli Studi di Pisa.
- Goldsmith, J. (1976). *Autosegmental Phonology*. PhD thesis, Massachusetts Institute of Technology.

## References II

- Hammond, M. (1988). On deriving the Well-Formedness Condition. *Linguistic Inquiry*, 19(2):319–325.
- Hyman, L. (2014). How autosegmental is phonology? *The Linguistic Review*, 31:363–400.
- Kornai, A. (1995). *Formal Phonology*. Garland Publication.
- Leben, W. R. (1973). *Suprasegmental phonology*. PhD thesis, Massachusetts Institute of Technology.
- López, D., Calera-Rubio, J., and Gallego-Sánchez, A.-J. (2012). Inference of  $k$ -testable directed acyclic graph languages. In *Proceedings of the 11th International Conference on Grammatical Inference (ICGI 2012)*, JMLR Workshop Proceedings.
- McCarthy, J. J. (1986). OCP effects: gemination and antigemination. *Linguistic Inquiry*, 17:207–263.
- Odden, D. (1986). On the role of the Obligatory Contour Principle in phonological theory. *Language*, 62(2):353–383.
- Sagey, E. (1986). *The Representation of Features and Relations in Non-Linear Phonology*. PhD thesis, Massachusetts Institute of Technology.