Local inviolable constraints: A new approach to syllable well-formedness in Berber

Kristina Strother-Garcia

University of Delaware

September 24, 2016

Outline

Outline

What makes a good theory of phonology?

1. Sufficiently expressive (doesn't undergenerate)
2. Maximally restrictive (doesn't overgenerate)
3. Efficiently learnable

Big-Picture Questions

- How can formal language theory and logic inform syllable theory?
- How can syllable well-formedness be accounted for with local inviolable constraints?
- What advantages come with representing syllable well-formedness this way?

Specific Objectives of This Talk

- Briefly review motivations for the present work
- Introduce a model-theoretic representation of syllable structure
- Formalize universal and language-specific local inviolable constraints
- Show how these constraints account for surface patterns in Berber

Outline

Rule-based Approaches to Berber

- Dell & Elmedlaoui (D&E) 1985
  - Ordered set of iterative core syllabification rules
  - Each rule identifies nuclei of a certain sonority class, ordered from most to least sonorous
  - Additional rules assign remaining consonants to onsets/codas
- Frampton 2011
  - Simplified version of D&E's rule set
  - Simultaneously identifies all points of application

OT Approach to Berber

Prince & Smolensky (P&S) 1993

- ONS: 'Syllables must have onsets (except phrase-initially).'
- HNUC: 'A higher sonority nucleus is more harmonic than one of lower sonority.'

Note: HNUC cannot be evaluated locally because every segment in a given syllable must be compared to the nucleus, and there is no a priori restriction on syllable size.

Constraint Ranking

ONSET ≫ HNUC

Correctly predicts the surface form [t**X**.z**N**t] 'you (sg.) stored.'[1]

(17) **Parallel Analysis of Complete Syllabification** of /txznt/

Candidates	ONS	HNUC		Comments
☞ .t**X**.z**N**t.		n	x	optimal
.**T**x.z**N**t.		n	t !	|n| = |n|, |t| < |x|
.t**X**z.n**T**.		x !	t	|x| < |n|, t irrelevant
.tx**Z**.**N**t.	* !	n	z	HNUC irrelevant
.**T**.**X**.**Z**.**N**.**T**.	* ! ***	n z x t t		HNUC irrelevant

[1] As in P&S, I use boldface uppercase letters for consonants that are syllabic nuclei.

Problems with These Frameworks

Expressiveness, restrictiveness, & learnability

- Both are adequate for describing syllable well-formedness in Berber, but they also overgenerate (Riggle 2004; Gainor, Lai, & Heinz 2012; Heinz & Lai 2013; Heinz, forthcoming)
- Classic OT also undergenerates due to difficulties with opacity
- Learning results for rule-based approaches are unclear

Example: Majority Rules

Given a language with front-back vowel harmony, consider these constraints (as in Bakovic 2000):

- AGREE[front]: 'Two consecutive vowels must have the same [front] value.'
- IDENT[front]: 'Do not change the value of [front].'

Majority Rules: [–front]

With two underlying [–back] vowels, the optimal candidate is back-harmonizing.

/– – +/	AGREE[front]	IDENT[front]
– – +	*!	
⇒ – – –		*
+ + +		**!

Majority Rules: [+front]

With two underlying [+back] vowels, the optimal candidate is front-harmonizing.

/+ – +/	AGREE[front]	IDENT[front]
+ – +	*!*	
– – –		**!
⇒ + + +		*

How do we rule out Majority Rules?

- Pathologies like Majority Rules are directly related to the degree of computational power that is allowed (Gainor, Lai, & Heinz 2012)
- Global constraint evaluation allows unbounded counting
- Local constraint evaluation does not

Why Use Inviolable Surface Constraints?

Sets of inviolable surface constraints describe established language classes of known computational power, allowing us to:

- Use computational complexity to make principled distinctions between what is possible (attested) and impossible (unattested) in phonology (Gainor, Lai, & Heinz 2012)
- Evaluate under- and over-generation problems and learnability in existing theoretical treatments

Why Focus on Local Constraints?

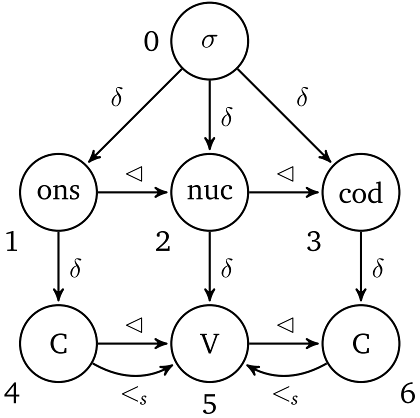
- Reduces hypothetical phonological phenomena to a highly restricted class of patterns (Heinz 2010; Rogers & Pullum 2011; Rogers et al. 2013)
- Rules out certain unattested patterns (Heinz & Lai 2013)
- Previous work shows that local substructure constraints can characterize:
  - Local and long-distance phonotactics (Heinz 2007, 2009, 2010)
  - Tone well-formedness patterns (Jardine 2016)
  - Mappings from URs to SRs (Chandlee 2014)

Why Focus on Syllables?

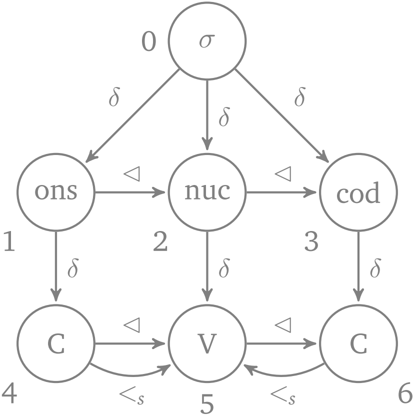
- One of the most referenced phonological domains
- Central to economical accounts of many processes and patterns
- Syllable structure is hierarchical, requiring at least three tiers with dominance relations between them – structures of this complexity have not yet been investigated in this framework

Outline

Elements of the Word Model



Elements of the Word Model: Alphabet

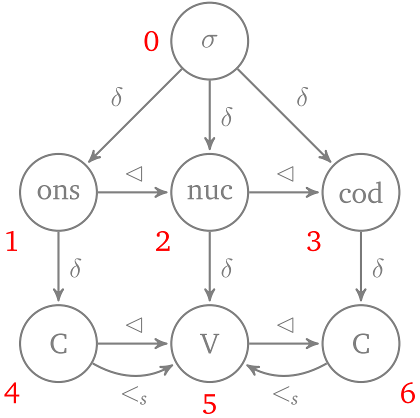


Alphabet, Σ

A set of node labels

$\Sigma = \{\mathrm{C}, \mathrm{V}, \mathrm{ons}, \mathrm{nuc}, \mathrm{cod}, \sigma\}$

Elements of the Word Model: Domain

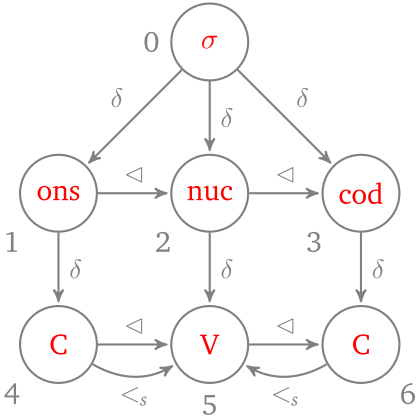


<u>Domain, $\mathcal{D}$</u>

A set of node positions

$\mathcal{D} = \{0, 1, 2, 3, 4, 5, 6\}$

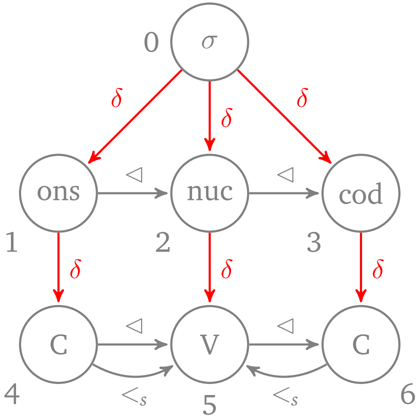
Elements of the Word Model: Labeling Relations



Labeling Relations (unary)

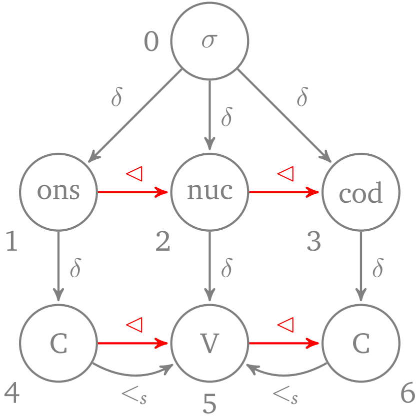
- $\sigma(x)$: node x is labeled σ
- $\text{ons}(x)$: node x is labeled ons
- ...etc.

Elements of the Word Model: Dominance Relation



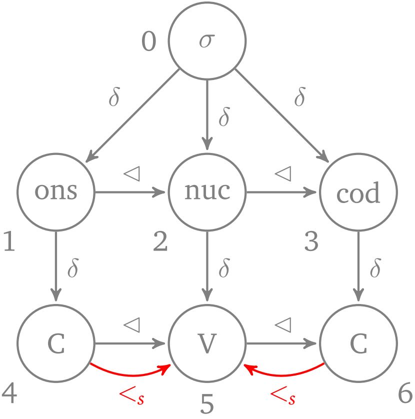
Immediate Dominance Relation
(binary)
$\delta(x,y)$: x immediately dominates y.

Elements of the Word Model: Immediate Precedence Relation



Immediate Precedence Relation
(binary)
$\lhd(x,y)$: x immediately precedes y.

Elements of the Word Model: Sonority Relation

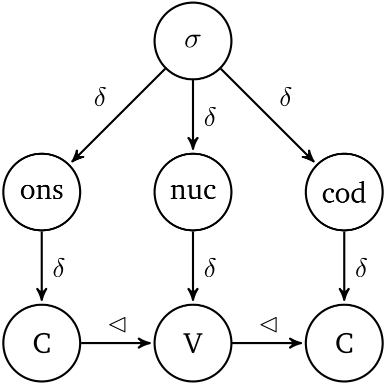


Less Sonorous
(binary)
$<_s(x, y)$: x is less sonorous than y.

Simplifying the Visual Representation

For clarity in the remaining figures, I will sometimes omit:

- Position numbers
- Sonority relations
- Immediate precedence edges between ons, nuc, and cod



Outline

Outline

Universal Constraints

Sticking to canonical syllable types for now (e.g., no ambisyllabicity, extrasyllabicity, etc.), we can establish some universal constraints on syllable structure.

- Every syllable has exactly one nucleus
- An onset must not immediately precede a coda
- ...and so on

Universal Constraints

Sticking to canonical syllable types for now (e.g., no ambisyllabicity, extrasyllabicity, etc.), we can establish some universal constraints on syllable structure.

- Every syllable has exactly one nucleus
- An onset must not immediately precede a coda
- ...and so on

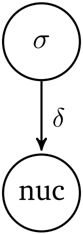
Exactly One Nucleus

This breaks down into two constraints:

1. NUCLEUS REQUIRED
2. NUCLEUS UNIQUE

Nucleus Required

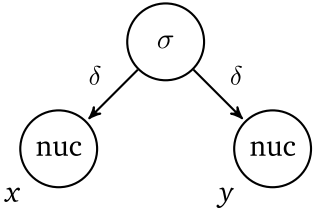
Every σ node must dominate a nuc node. Thus every syllable must contain the following substructure:



Note: This is a **positive** constraint that refers to a connected sub-graph of size 2.

Nucleus Unique

A σ node may not dominate two unique nuc nodes. Thus the following substructure is banned:



Note: This is a **negative** constraint that refers to a connected sub-graph of size 3.

...etc.

Other structural well-formedness constraints can be formalized in a similar way

- Certain substructures are required
- Certain substructures are banned
- These types of constraints all refer to connected sub-graphs of a finite size

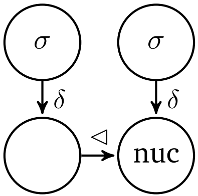
Outline

Language-specific Constraints

- Every language will have some language-specific constraints
- Examples: onset required, coda forbidden
- As with universals, these are local substructure constraints

Internal Onsets Required

In Berber, all non-initial syllables must have an onset. That is, a nuc node may not immediately follow a node dominated by a different σ node. Thus the following substructure is banned:



Outline

The Sonority Hierarchy in Berber

While there may be some universal sonority relations, I assume for now that every language has its own sonority hierarchy. D&E give the following the sonority hierarchy for Berber:

voiceless stops $<_s$ voiced stops $<_s$ voiceless fricatives
$<_s$ voiced fricatives $<_s$ nasals $<_s$ liquids $<_s$ high vowels $<_s$ [a]

Sonority Relations

- If segment x is less sonorous than segment y, we write $<_s (x, y)$ or, equivalently, $x <_s y$.
- As with the traditional notion of lesser sonority, I assume that the binary relation $<_s$ is irreflexive, asymmetric, and transitive.

Sonority Relations

- A binary relation $R(x,y)$ is *irreflexive* iff for all x, $\neg R(x,x)$.
  Example: [t] is not less sonorous than itself.
- A binary relation $R(x,y)$ is *asymmetric* iff for all x,y, if $R(x,y)$ then $\neg R(y,z)$.
  Example: If [t] is less sonorous than [m], then [m] cannot be less sonorous than [t].
- A binary relation $R(x,y)$ is *transitive* iff for all x,y,z, if $R(x,y)$ and $R(y,z)$ then $R(x,z)$.
  Example: If [t] $<_s$ [m] and [m] $<_s$ [a], then [t] $<_s$ [a].

Sonority Relations

- A binary relation $R(x, y)$ is *irreflexive* iff for all x, $\neg R(x, x)$.
  Example: [t] is not less sonorous than itself.
- A binary relation $R(x, y)$ is *asymmetric* iff for all x, y, if $R(x, y)$ then $\neg R(y, z)$.
  Example: If [t] is less sonorous than [m], then [m] cannot be less sonorous than [t].
- A binary relation $R(x, y)$ is *transitive* iff for all x, y, z, if $R(x, y)$ and $R(y, z)$ then $R(x, z)$.
  Example: If [t] $<_s$ [m] and [m] $<_s$ [a], then [t] $<_s$ [a].

Sonority Relations

- A binary relation $R(x,y)$ is *irreflexive* iff for all x, $\neg R(x,x)$.
  Example: [t] is not less sonorous than itself.
- A binary relation $R(x,y)$ is *asymmetric* iff for all x, y, if $R(x,y)$ then $\neg R(y,z)$.
  Example: If [t] is less sonorous than [m], then [m] cannot be less sonorous than [t].
- A binary relation $R(x,y)$ is *transitive* iff for all x, y, z, if $R(x,y)$ and $R(y,z)$ then $R(x,z)$.
  Example: If [t] $<_s$ [m] and [m] $<_s$ [a], then [t] $<_s$ [a].

Sonority Relations

Given these properties of $<_s$, it is simple to define a relation $=_s$ to represent equal sonority and a relation $\leq_s$ to represent equal or lesser sonority.

- $=_s (x,y) \stackrel{def}{=} \neg <_s (x,y) \wedge \neg <_s (y,x)$

  Interpretation: x and y are equally sonorous iff x is not less sonorous than y and y is not less sonorous than x.

- $\leq_s (x,y) \stackrel{def}{=} <_s (x,y) \vee =_s (y,x)$

  Interpretation: x is equally or less sonorous than y iff x is less sonorous than y or x and y are equally sonorous.

Sonority Relations

Given these properties of $<_s$, it is simple to define a relation $=_s$ to represent equal sonority and a relation $\leq_s$ to represent equal or lesser sonority.

- $=_s (x,y) \overset{def}{=} \neg <_s (x,y) \wedge \neg <_s (y,x)$

  Interpretation: x and y are equally sonorous iff x is not less sonorous than y and y is not less sonorous than x.

- $\leq_s (x,y) \overset{def}{=} <_s (x,y) \vee =_s (y,x)$

  Interpretation: x is equally or less sonorous than y iff x is less sonorous than y or x and y are equally sonorous.

Sonority Relations

Given these properties of $<_s$, it is simple to define a relation $=_s$ to represent equal sonority and a relation $\leq_s$ to represent equal or lesser sonority.

- $=_s (x,y) \stackrel{def}{=} \neg <_s (x,y) \wedge \neg <_s (y,x)$

  Interpretation: x and y are equally sonorous iff x is not less sonorous than y and y is not less sonorous than x.

- $\leq_s (x,y) \stackrel{def}{=} <_s (x,y) \vee =_s (y,x)$

  Interpretation: x is equally or less sonorous than y iff x is less sonorous than y or x and y are equally sonorous.

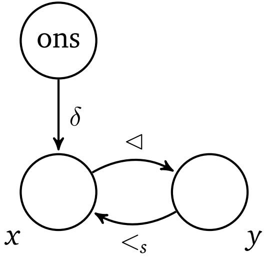
Sonority Constraints

Using these binary sonority relations as a starting point, the SSP can be formulated in two parts:

1. RIGHT OF ONS: Sonority must not fall rightward from the onset
2. LEFT OF COD: Sonority must not fall leftward from the coda

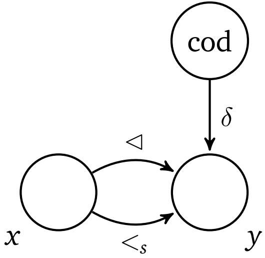
Right of Ons

A node dominated by an ons node may not immediately precede a node of lesser sonority. Thus the following substructure is banned:



LEFT OF CODA

A node dominated by a cod node may not immediately follow a node of lesser sonority. Thus the following substructure is banned:



Outline

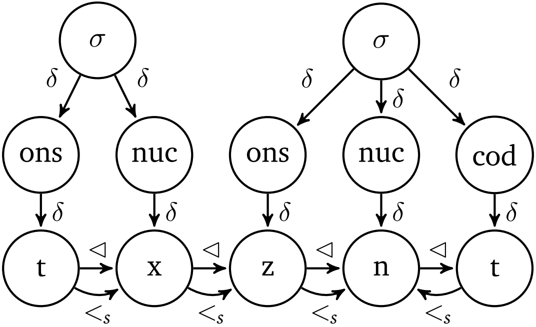
Refresher: P&S

(17) **Parallel Analysis of Complete Syllabification** of /txznt/

Candidates	ONS	HNUC		Comments
☞ .tX.zNt.		n	x	optimal
.Tx.zNt.		n	t !	|n| = |n|, |t| < |x|
.tXz.nT.		x !	t	|x| < |n|, t irrelevant
.txZ.Nt.	* !	n	z	HNUC irrelevant
.T.X.Z.N.T.	* ! ***	n z x t t		HNUC irrelevant

Again, the adequacy of this result is not in question; the goal here is to show that the same result is obtained by evaluating only local inviolable substructure constraints.

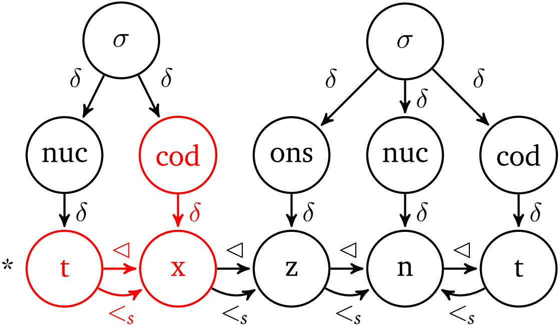
'Winner'



[t**X**.z**N**t]

✓Right of Ons
✓Left of Coda
✓Internal Onsets Required

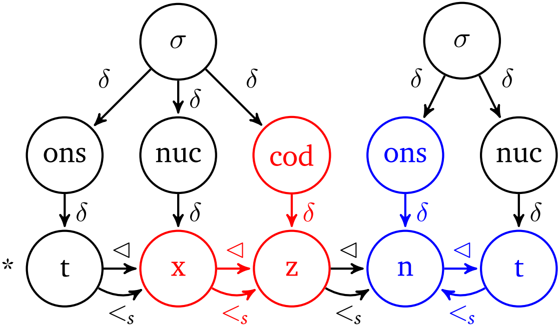
Loser 1



*[**T**x.z**N**t]

✗LEFT OF CODA

Loser 2

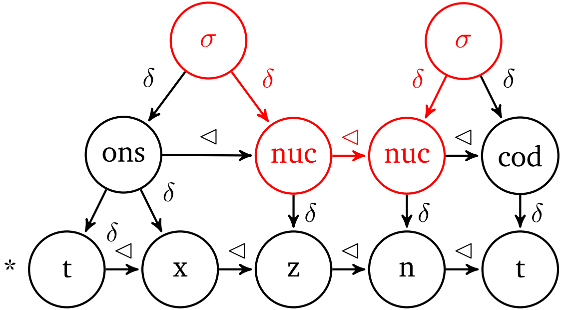


*[t**X**z.n**T**]

✗Left of Coda

✗Right of Onset

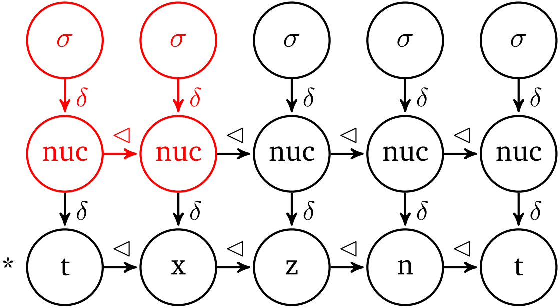
Loser 3



*[txZ.Nt]

✗INTERNAL ONSETS REQUIRED

Loser 4



*[T.X.Z.N.T]

✗INTERNAL ONSETS REQUIRED

Outline

Putting It All Together

Universal structural well-formedness constraints

$+$

Language-specific constraints

$+$

Language-specific sonority relations

$=$

Language-specific syllable well-formedness

Take-home Points

- Hierarchical word models provide a maximally explicit representation of syllable structure
- Syllable well-formedness can be characterized by local inviolable constraints, both universal and language-specific
- The posited constraints describe a restricted class of graph sets because they all refer to sub-graphs of size 4 or smaller – much less expressive than SPE-style and OT frameworks

OT Comparison

OT Constraints	Proposed Constraints
Violable	Inviolable
Global	Local
Solely universal	A combination of universals and language-specifics

Additional Considerations

- The exact processes that repair ill-formed syllable structures (e.g., epenthesis, deletion, etc.) must be guided by additional language-specific principles
- Regardless of the nature of the repair processes, the **necessity** of such repairs can be determined by evaluating surface forms with respect to local inviolable constraints – no optimization

Future Work

- Conduct more case studies to account for complex margins and non-canonical syllable structures (e.g., ambisyllabicity)
- Write a program to generate possible syllabifications of a string and evaluate them with respect to the proposed constraints – as in OT, need to ensure that all the crucial 'candidates' are considered
- Develop graph transductions to characterize the mapping from URs to SRs

Thanks!

Special thanks to Jeff Heinz, Adam Jardine, Taylor Miller, Eric Bakovic, Kevin McMullin, and the entire NAPhC 2016 audience for their insightful feedback.

Contact Info
kmsg@udel.edu
sites.udel.edu/kmsg