Slide 1

Classifying Stress Patterns by Cognitive Complexity

James Rogers
Dept. of Computer Science
Earlham College
`jrogers@cs.earlham.edu`

`http://cs.earlham.edu/~jrogers/slides/UConn.ho.pdf`

Joint work with Jeff Heinz, U. Delaware,
and a raft of Earlham College undergrads.

Slide 2

Cognitive Complexity from First Principles

What kinds of distinctions does a cognitive mechanism need to be sensitive to in order to classify an event with respect to a pattern?

Descriptive Classes of Formal Languages

- Characterized by the nature of information about the properties of strings that determine membership
- Independent of mechanisms for recognition
- Subsume wide range of types of patterns

Slide 3

Local Classes—Adjacency

Blocks of consecutive syllables

- SL—Strictly Local (Restricted Propositional Logic with Successor)
  - Co-ocurrence of negative atomic constraints
- LT—Locally Testable (Propositional Logic with Successor)
  - Boolean combinations of atomic constraints
- LTT—Locally Threshold Testable (First-Order Logic with Successor)
  - Boolean combinations of constraints on multiplicity of blocks, up to some threshold
- SF—Star-Free (First-Order Logic with Less-Than)
  - Boolean combinations of constraints on order of blocks

Slide 4

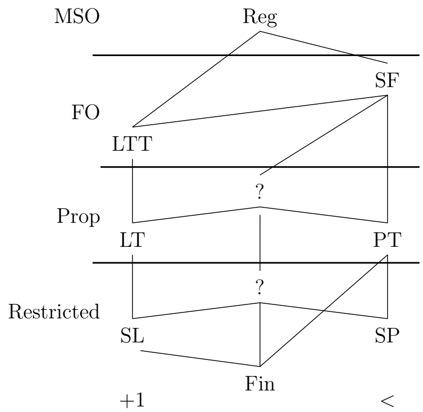
Piecewise Classes—Precedence

Subsequences of syllables, not necessarily consecutive

- SP—Strictly Piecewise (Restricted Propositional Logic with Less-Than)
  - Co-ocurrence of negative atomic constraints
- PT—Piecewise Testable (Propositional Logic with Less-Than)
  - Boolean combinations of atomic constraints
- SF—Star-Free (First-Order Logic with Less-Than)
  - Boolean combinations of constraints on order of blocks
- Reg—Regular (Monadic Second-Order Logic over Strings)
  - Constraints based on grouping events into finitely many categories

Sub-Regular Hierarchies

Slide 5



Yidin

- Primary stress on the leftmost heavy syllable, else the initial syllable
- Secondary stress iteratively on every second syllable in both directions from primary stress
- No light monosyllables

Slide 6

Explicitly:

- Exactly one $\acute{\sigma}$ (One-$\acute{\sigma}$)
- $\acute{L}$ implies no H (No-H-with-$\acute{L}$)
- σ and $\grave{\sigma}/\acute{\sigma}$ alternate (Alt)
- First H gets primary stress (No-H-before-$\acute{H}$)
- $\acute{L}$ only if initial (Nothing-before-$\acute{L}$)
- No $\acute{L}$ monosyllables (No $\rtimes\acute{L}\ltimes$)

Slide 7

k-Expressions

Atomic Propositions (k-factors)

$$w \models \sigma_1\sigma_2 \mathinner{..} \sigma_k \quad \overset{\text{def}}{\Longleftrightarrow} \quad w = \cdots\ \sigma_1\sigma_2 \mathinner{..} \sigma_k\ \cdots$$
$$w \models \rtimes\sigma_1\sigma_2 \mathinner{..} \sigma_{k-1} \quad \overset{\text{def}}{\Longleftrightarrow} \quad w = \sigma_1\sigma_2 \mathinner{..} \sigma_{k-1}\ \cdots$$
$$w \models \sigma_1\sigma_2 \mathinner{..} \sigma_{k-1}\ltimes \quad \overset{\text{def}}{\Longleftrightarrow} \quad w = \cdots\ \sigma_1\sigma_2 \mathinner{..} \sigma_{k-1}$$

Compound Propositions

$$w \models \varphi \wedge \psi \quad \overset{\text{def}}{\Longleftrightarrow} \quad w \models \varphi \text{ and } w \models \psi$$
$$w \models \neg\varphi \quad \overset{\text{def}}{\Longleftrightarrow} \quad w \not\models \varphi$$

Slide 8

Strictly Local Constraints

Definition 1 (Strictly Local Sets) *A stringset L over Σ is* Strictly Local *iff there is some k-expression over Σ*

$$\varphi = \neg f_1 \wedge \neg f_2 \wedge \cdots \wedge \neg f_n,$$

a conjunction of negative literals, such that L is the set of all strings that satisfy φ:

$$L = L(\varphi) \overset{def}{=} \{w \in \Sigma^* \mid w \models \varphi\}$$

- Nothing-before-$\acute{L}$ $\quad \neg\sigma\acute{L}$ $\quad$ (SL_2)
- Alt $\quad \neg\sigma\sigma \wedge \neg\acute{\sigma}\acute{\sigma} \wedge \neg\acute{\sigma}\grave{\sigma} \wedge \neg\grave{\sigma}\acute{\sigma} \wedge \neg\grave{\sigma}\grave{\sigma}$ $\quad$ (SL_2)
- No $\rtimes\acute{L}\ltimes$ $\quad \neg\rtimes\acute{L}\ltimes$ $\quad$ (SL_3)

Character of Strictly k-Local Sets

Slide 9

Theorem (Suffix Substitution Closure):
A stringset L is strictly k-local iff whenever there is a string x of length $k-1$ and strings w, y, v, and z, such that

$$\begin{array}{ccccccl} w & \cdot & \overbrace{x}^{k-1} & \cdot & y & \in L \\ v & \cdot & x & \cdot & z & \in L \end{array}$$

then it will also be the case that

$$w \cdot x \cdot z \in L$$

No-H-with-$\acute{L}$ is not SL_k:

Slide 10

$$\begin{array}{ccccccl} \acute{L} & \cdot & \overbrace{L \cdots L}^{k-1} & \cdot & L & \in \text{No-}H\text{-with-}\acute{L} \\ \acute{H} & \cdot & \overbrace{L \cdots L}^{k-1} & \cdot & H & \in \text{No-}H\text{-with-}\acute{L} \\ \hline \acute{L} & \cdot & \overbrace{L \cdots L}^{k-1} & \cdot & H & \notin \text{No-}H\text{-with-}\acute{L} \end{array}$$

Mechanisms that are sensitive only to the fixed length blocks of consecutive syllables in a word cannot distinguish words in which $\acute{L}$ occurs with H from those in which it does not.

Slide 11

Cognitive interpretation of SL

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) SL_k language must be sensitive, at least, to the length k blocks of consecutive events that occur in the presentation of the string.
- If the strings are presented as sequences of events in time, then this corresponds to being sensitive, at each point in the string, to the immediately prior sequence of $k-1$ events.
- Any cognitive mechanism that is sensitive *only* to the length k blocks of consecutive events in the presentation of a string will be able to recognise *only* SL_k languages.

Slide 12

Strictly Local Stress Patterns

Heinz's Stress Pattern Database (ca. 2007)—109 patterns

9 are SL_2	Abun West, Afrikans, ... Cambodian,... Maranungku
44 are SL_3	Alawa, Arabic (Bani-Hassan),...
24 are SL_4	Arabic (Cairene),...
3 are SL_5	Asheninca, Bhojpuri, Hindi (Fairbanks)
1 is SL_6	Icua Tupi
28 are not SL	Amele, Bhojpuri (Shukla Tiwari), Arabic Classical, Hindi (Keldar), Yidin,...

72% are SL, all $k \leq 6$. 49% are SL_3.

Locally definable stringsets

Slide 13

Definition 2 (Locally Testable Sets) *A stringset L over Σ is* Locally Testable *iff (by definition) there is some k-expression φ over Σ (for some k) such that L is the set of all strings that satisfy φ:*

$$L = L(\varphi) \stackrel{def}{=} \{w \in \Sigma^* \mid w \models \varphi\}$$

No-*H*-with-*Ĺ* is $\mathbf{LT_1}$:

$$\neg(H \wedge \acute{L})$$

Character of Locally Testable sets

Slide 14

Theorem 1 (k-Test Invariance) *A stringset L is Locally Testable iff*

there is some k such that, for all strings w and v,

if $\rtimes \cdot w \cdot \ltimes$ and $\rtimes \cdot v \cdot \ltimes$ have exactly the same set of k-factors

then either both w and v are members of L or neither is.

LT_k definitions cannot distiguish between strings that are made up of the same set of k-factors.

One-$\acute{\boldsymbol{\sigma}}$ is not *LT*

$$\rtimes \sigma_1 \overbrace{\sigma_0 \cdots \sigma_0}^{k-1} \acute{\sigma}_1 \overbrace{\sigma_0 \cdots \sigma_0}^{k-1} \ltimes$$

$$\equiv^L_k$$

$$\star \rtimes \sigma_1 \overbrace{\sigma_0 \cdots \sigma_0}^{k-1} \acute{\sigma}_1 \overbrace{\sigma_0 \cdots \sigma_0}^{k-1} \acute{\sigma}_1 \overbrace{\sigma_0 \cdots \sigma_0}^{k-1} \ltimes$$

Slide 15

Mechanisms that are sensitive only to the set of fixed length blocks of syllables in a word cannot, *in general*, distinguish words with a single primary stressed syllable from those with more than one.

Valid stress patterns are either SL or they are not LT.

Cognitive interpretation of LT

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) LT_k language must be sensitive, at least, to the *set* of length k contiguous blocks of events that occur in the presentation of the string—both those that do occur and those that do not.

Slide 16

- If the strings are presented as sequences of events in time, then this corresponds to being sensitive, at each point in the string, to the set of length k blocks of events that occurred at any prior point.
- Any cognitive mechanism that is sensitive *only* to the occurrence or non-occurrence of length k contiguous blocks of events in the presentation of a string will be able to recognise *only* LT_k languages.

Slide 17

FO(+1)

Models: $\langle \mathcal{D}, \triangleleft, P_\sigma \rangle_{\sigma \in \Sigma}$

First-order Quantification (over positions in the strings)

$$
\begin{array}{ccll}
x \triangleleft y & w, [x \mapsto i, y \mapsto j] \models x \triangleleft y & \overset{\text{def}}{\iff} & j = i + 1 \\
P_\sigma(x) & w, [x \mapsto i] \models P_\sigma(x) & \overset{\text{def}}{\iff} & i \in P_\sigma \\
\varphi \wedge \psi & \vdots & & \\
\neg\varphi & \vdots & & \\
(\exists x)[\varphi(x)] & w, s \models (\exists x)[\varphi(x)] & \overset{\text{def}}{\iff} & w, s[x \mapsto i] \models \varphi(x) \\
 & & & \text{for some } i \in \mathcal{D}
\end{array}
$$

FO(+1)-Definable Stringsets: $L(\varphi) \overset{\text{def}}{=} \{w \mid w \models \varphi\}$.

Slide 18

One-$\acute{\sigma}$ is FO(+1) definable

$$(\exists x)[\acute{\sigma}(x) \wedge (\forall y)[\acute{\sigma}(y) \rightarrow x \approx y]\,]$$

Character of the FO(+1) Definable Stringsets

Slide 19

Definition 3 (Locally Threshold Testable) *A set L is* Locally Threshold Testable *(LTT) iff there is some k and t such that, if two strings either contain the same number of occurrences of each block of k consecutive symbols or both contain at least t occurrences, then either both are in the set or neither is.*

Theorem 2 (Thomas) *A set of strings is First-order definable over $\langle \mathcal{D}, \triangleleft, P_\sigma \rangle_{\sigma \in \Sigma}$ iff it is* Locally Threshold Testable.

FO(+1) definitions cannot distinguish between strings that have the same multiplicity of the k-factors, counting up to some fixed finite threshold.

No H before $\acute{H}$ is not FO(+1)

Slide 20

Primary stress on leftmost heavy syllable

$$\star \;\; H \ldots \acute{H}$$

$$\rtimes\, \overbrace{\grave{L}L \cdots \grave{L}L}^{2kt}\; \acute{H}H\; \overbrace{\grave{L}L \cdots \grave{L}L}^{2kt}\; \grave{H}H\; \overbrace{\grave{L}L \cdots \grave{L}L}^{2kt} \ltimes$$

$$\equiv^{L}_{k,t}$$

$$\star \rtimes\, \underbrace{\grave{L}L \cdots \grave{L}L}_{2kt}\; \grave{H}H\; \underbrace{\grave{L}L \cdots \grave{L}L}_{2kt}\; \acute{H}H\; \underbrace{\grave{L}L \cdots \grave{L}L}_{2kt} \ltimes$$

Mechanisms that are sensitive only to the multiplicity, up to some fixed threshold, of fixed length blocks of syllables in a word cannot distinguish words in which some heavy syllable occurs prior to one with primary stress from those in which the first heavy syllable has primary stress.

Cognitive interpretation of FO(+1)

Slide 21

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) FO(+1) stringset must be sensitive, at least, to the multiplicity of the length k blocks of events, for some fixed k, that occur in the presentation of the string, distinguishing multiplicities only up to some fixed threshold t.
- If the strings are presented as sequences of events in time, then this corresponds to being able count up to some fixed threshold.
- Any cognitive mechanism that is sensitive *only* to the multiplicity, up to some fixed threshold, (and, in particular, not to the order) of the length k blocks of events in the presentation of a string will be able to recognize *only* FO(+1) stringsets.

First-Order(<) definable stringsets

Slide 22

$\langle \mathcal{D}, \triangleleft^+, P_\sigma \rangle_{\sigma \in \Sigma}$

First-order Quantification over positions in the strings

$x \triangleleft^+ y$	$w, [x \mapsto i, y \mapsto j] \models x \triangleleft^+ y$	$\overset{\text{def}}{\iff}$	$i < j$
$P_\sigma(x)$	$w, [x \mapsto i] \models P_\sigma(x)$	$\overset{\text{def}}{\iff}$	$i \in P_\sigma$
$\varphi \wedge \psi$	$\vdots$		
$\neg\varphi$	$\vdots$		
$(\exists x)[\varphi(x)]$	$w, s \models (\exists x)[\varphi(x)]$	$\overset{\text{def}}{\iff}$	$w, s[x \mapsto i] \models \varphi(x)$ for some $i \in \mathcal{D}$

Slide 23

No-H-before-$\acute{H}$ is First-Order($<$) definable

$$\neg(\exists x, y)[x \triangleleft^{+} y \wedge H(x) \wedge \acute{H}(y)]$$

Slide 24

Star-Free stringsets

Definition 4 (Star-Free Set) *The class of* Star-Free Sets *(SF) is the smallest class of languages satisfying:*

- *Fin* $\subseteq$ *SF.*
- *If* $L_1, L_2 \in$ *SF then:* $L_1 \cdot L_2 \in$ *SF,*
  $L_1 \cup L_2 \in$ *SF,*
  $\overline{L_1} \in$ *SF.*

Theorem 3 (McNauthton and Papert) *A set of strings is First-order definable over* $\langle \mathcal{D}, \triangleleft^{+}, P_\sigma \rangle_{\sigma \in \Sigma}$ *iff it is* Star-Free*.*

Slide 25

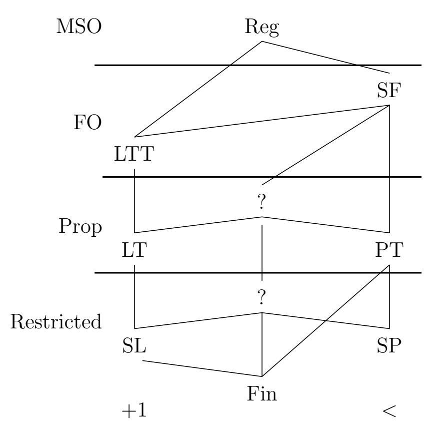
Classifying Conjunctive Constraints

- One-$\acute{\sigma}$ $(\exists! x)[\acute{\sigma}(x)]$ ($LTT_{1,2}$)
- No-H-before-$\acute{H}$ $\neg(\exists x, y)[x \triangleleft^{+} y \wedge H(x) \wedge \acute{H}(y)]$ (SF)
- No-H-with-$\acute{L}$ $\neg(H \wedge \acute{L})$ (LT_1)
- Nothing-before-$\acute{L}$ $\neg\sigma\acute{L}$ (SL_2)
- Alt $\neg\sigma\sigma \wedge \neg\acute{\sigma}\acute{\sigma} \wedge \neg\acute{\sigma}\grave{\sigma} \wedge \neg\grave{\sigma}\acute{\sigma} \wedge \neg\grave{\sigma}\grave{\sigma}$ (SL_2)
- No $\rtimes\acute{L}\ltimes$ $\neg\rtimes\acute{L}\ltimes$ (SL_3)

Yidin is SF

Slide 26

Sub-Regular Hierarchies



Slide 27

PT$_k$-expressions

Atomic Propositions (k-sequences)

$$w \models \sigma_1 \mathinner{..} \sigma_k \quad \stackrel{\text{def}}{\Longleftrightarrow} \quad w = \cdots\ \sigma_1\ \cdots\ \sigma_k\ \cdots$$

Compound Propositions

$$w \models \varphi \wedge \psi \quad \stackrel{\text{def}}{\Longleftrightarrow} \quad w \models \varphi \text{ and } w \models \psi$$

$$w \models \neg\varphi \quad \stackrel{\text{def}}{\Longleftrightarrow} \quad w \not\models \varphi$$

Slide 28

Strictly Piecewise Constraints

Definition 5 (Strictly Piecewise Sets) *A stringset L over Σ is* Strictly Piecewise *iff there is some k-expression over Σ*

$$\varphi = \neg f_1 \wedge \neg f_2 \wedge \cdots \wedge \neg f_n,$$

a conjunction of negative literals, such that L is the set of all strings that satisfy φ:

$$L = L(\varphi) \stackrel{def}{=} \{w \in \Sigma^* \mid w \models \varphi\}$$

- No-H-before-$\acute{H}$ $\quad \neg H\acute{H}$ $\quad$ (SP$_2$)
- No-H-with-$\acute{L}$ $\quad \neg H\acute{L} \wedge \neg\acute{L}H$ $\quad$ (SP$_2$)
- Nothing-before-$\acute{L}$ $\quad \neg\sigma\acute{L}$ $\quad$ (SP$_2$)

Character of the Strictly k-Piecewise Sets

Theorem 4 *A stringset L is Strictly k-Piecewise Testable iff it is closed under subsequence:*

$$w\sigma v \in L \Rightarrow wv \in L$$

Slide 29

One-$\acute{\sigma}$ is not SP

$$\sigma\acute{\sigma}\sigma \in L \text{ but } \sigma\sigma \notin L$$

Mechanisms that are sensitive only to subsequences (insensitive to intevening symbols) cannot distinguish words in which some primary stress occurs from those in which none does.

But SP can forbid multiple primary stress: $\neg\acute{\sigma}\acute{\sigma}$ (At-Most-One-$\acute{\sigma}$)

Slide 30

Cognitive interpretation of SP

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) SP_k stringset must be sensitive, at least, to the length k (not necessarily consecutive) sequences of events that occur in the presentation of the string.
- If the strings are presented as sequences of events in time, then this corresponds to being sensitive, at each point in the string, to up to $k - 1$ events distributed arbitrarily among the prior events.
- Any cognitive mechanism that is sensitive *only* to the length k sequences of events in the presentation of a string will be able to recognize *only* SP_k stringsets.

Piecewise definable stringsets

Slide 31

Definition 6 (Piecewise Testable Sets) *A stringset L over Σ is* Piecewise Testable *iff (by definition) there is some PT_k-expression φ over Σ (for some k) such that L is the set of all strings that satisfy φ:*

$$L = L(\varphi) \stackrel{def}{=} \{w \in \Sigma^* \mid w \models \varphi\}$$

- No-H-with-$\acute{L}$ $\qquad \neg(H \wedge \acute{L}) \qquad (\mathrm{PT}_2)$
- No $\rtimes\acute{L}\ltimes$ $\qquad \acute{L} \rightarrow (\sigma\acute{L} \vee \acute{L}\sigma) \wedge \neg\acute{L}\acute{L} \qquad (\mathrm{PT}_2)$

Character of Piecewise Testable sets

Slide 32

Theorem 5 (k-Test Invariance) *A stringset L is Piecewise Testable iff*

there is some k such that, for all strings w and v,

if w and v have exactly the same set of k-sequences

then either both w and v are members of L or neither is.

PT_k definitions cannot distiguish between strings that are made up of the same set of k-sequences.

Slide 33

Alt is not *PT*

$$\overbrace{\sigma\grave{\sigma}\cdots\sigma\grave{\sigma}}^{2k} \equiv^{P}_{k} \star\ \overbrace{\sigma\grave{\sigma}\cdots\sigma\grave{\sigma}}^{2k}\grave{\sigma}$$

Mechanisms that are sensitive only to the set of fixed length subsequences of syllables in a word (insensitive to intevening syllables) cannot distinguish words in which stressed and unstressed syllables alternate from those in which adjacent pairs occur.

Slide 34

Cognitive interpretation of PT

- Any cognitive mechanism that can distinguish member strings from non-members of a (properly) PT_k stringset must be sensitive, at least, to the set of length k subsequences of events that occur in the presentation of the string—both those that do occur and those that do not.
- If the strings are presented as sequences of events in time, then this corresponds to being sensitive, at each point in the string, to the set of all length k subsequences of the sequence of prior events.
- Any cognitive mechanism that is sensitive *only* to the set of length k subsequences of events in the presentation of a string will be able to recognize *only* PT_k stringsets.

Slide 35

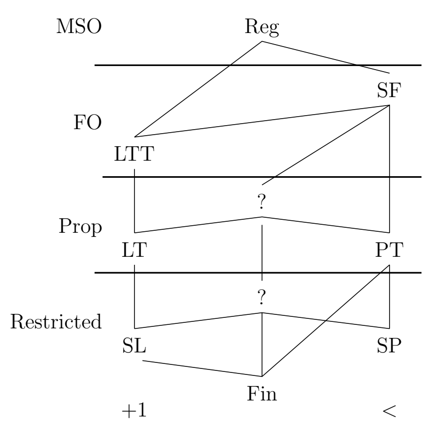
Yidin wrt Local and Piecewise Constraints

One-$\acute{\sigma}$	$\text{LTT}_{1,2}$	PT_2
Some-$\acute{\sigma}$	LT_1	PT_1
At-Most-One-$\acute{\sigma}$	$\text{LTT}_{1,2}$	SP_2
No-H-before-$\acute{H}$	SF	SP_2
No-H-with-$\acute{L}$	LT_1	SP_2
Nothing-before-$\acute{L}$	SL_2	SP_2
Alt	SL_2	SF
No $\rtimes\acute{L}\ltimes$	SL_3	PT_2

Yidin is co-occurence of SL and PT constraints or of LT and SP constraints

Slide 36

Sub-Regular Hierarchies



Slide 37

Some Additional Preliminary Results

Stress Patterns wrt Local Constraints

- SL
  89 of 109 patterns
- LT
  None
- LTT
  Alawa, Bulgarian, Murik
- SF
  Amele, Arabic (Classical), Buriat, Cheremis (East), Cheremis (Meadow), Chuvash, Golin, Komi, Kuuku Yau, Lithuanian, Mam, Maori, K. Mongolian (Street), K. Mongolian (Stuart), K. Mongolian (Bosson), Nubian, Yidin

Slide 38

Some Additional Preliminary Results

Stress Patterns wrt Piecewise Constraints

- SP
  None
- PT
  Amele, Bulgarian, Chuvash, Golin, Lithuanian, Maori K. Mongolian (Street), Murik,
- SF
  Alawa, Arabic (Classical), Buriat, Cheremis (East), Cheremis (Meadow), Komi, Kuuku Lau, Mam, K. Mongolian (Bosson), K. Mongolian (Stuart), Nubian, Yidin

Some Additional Preliminary Results

Stress Patterns wrt Co-occurrence of Local and Piecewise Constraints

Slide 39

- SL + SP
  89 of 109 patterns
- SL + PT
  Komi, Kuuku Lau, Yidin
- LT + SP
  Alawa Amele, Arabic (Classical), Bulgarian, Buriat, Cheremis (East), Cheremis (Meadow), Chuvash, Golin, Komi, Kuuku Lau, Lithuanian, Mam, Maori K. Mongolian (Bosson), K. Mongolian (Street), K. Mongolian (Stuart), Murik, Nubian, Yidin
- SF
  None

Some Constraints

Slide 40

- Forbidden syllables (SL_1, SP_1)
  - No heavy syllables
- Required syllables (LT, SP_1)
  - Some primary stress
- Forbidden initial/final syllables (SL_2, SF)
  - Cannot start with unstressed light
  - Cannot start with unstressed heavy
  - Cannot end with stressed light
- Forbidden adjacent pairs (SL_2, SF)
  - No adjacent unstressed
  - No adhacent secondary stress
  - No heavy immediately following a stressed light

...