Learning the Stress Patterns of the World's Languages

Jeffrey Heinz
heinz@udel.edu

University of Delaware

The 321st Regular Meeting of the Phonetic Society of Japan
at The National Institute of Japanese Languages and Linguistics
June 26, 2010

Introduction

Today I will present a previously unnoticed universal property of stress patterns found in the world's languages.

They are neighborhood-distinct.

(1) This property is a condition on **locality** in phonology.

(2) This property naturally provides an *inductive principle*, allowing learners to generalize correctly from limited experience

- without an a priori set of parameters (Chomsky 1981, Hayes 1995) or Optimality-theoretic constraints (Prince and Smolensky 1993, 2004).

On the Role of Locality in Learning Stress Patterns.
Phonology 26:2 (2009). 303-351.

Outline

Why Look at Stress Patterns?

- Typology is well-understood
- Children appear to learn stress early

Kinds of Stress Patterns

- Over the past thirty years, linguists have developed a picture of the range of variation that exists among the world's languages regarding how stress predictably occurs in words.

(Chomsky and Halle 1968, Liberman 1975, Hyman 1977, Prince 1980, Hayes 1981, Prince 1983, Halle and Vergnaud 1987, Idsardi 1992, Prince 1992, Bailey 1995, Hayes 1995, Hammond 1999, Elenbaas and Kager 1999, Walker 2000, Hyde 2002, Gordon 2002, Baković 2004).

Early Sensitivity to Language Rhythms

- Newborns(!) can discriminate languages based on their rhythmic properties (Mehler et al. 1988, Nazzi and Ramus 2003)
- English babies demonstrate a trochaic bias ($\acute{\sigma}\ \sigma$) at nine months, but not at six months (Jusczyk et al. 1993, Echols et al. 1997).

Kinds of Attested Stress Patterns

- Combining Bailey's 1995 survey and Gordon's 2002 survey yields a survey of 405 languages (423 patterns, 109 are distinct).
  - Quantity-Insensitive: Single, Dual, Binary, Ternary
  - Quantity-Sensitive Bounded: Single, Dual, Binary, Ternary, Multiple
  - Quantity-Sensitive Unbounded: Single, Binary, Multiple
- Dominant pattern for a language.
- Patterns apply within some domain assumed to be known.
- Lexical exceptions and morphological factors are ignored.
- Relevant phonotactics included (minimal word conditions, etc.)

Example 1: Pintupi (Quantity-Insensitive Binary)

a.	**σ́** σ	páɳa	'earth'
b.	**σ́** σ σ	tʲúʈaya	'many'
c.	**σ́** σ **σ̀** σ	máɭawàna	'through from behind'
d.	**σ́** σ **σ̀** σ σ	púɭiŋkàlatʲu	'we (sat) on the hill'
e.	**σ́** σ **σ̀** σ **σ̀** σ	tʲámulìmpatʲùŋku	'our relation'
f.	**σ́** σ **σ̀** σ **σ̀** σ σ	ʈíɭirìŋulàmpatʲu	'the fire for our benefit flared up'
g.	**σ́** σ **σ̀** σ **σ̀** σ **σ̀** σ	kúranʲùlulìmpatʲùɻa	'the first one who is our relation'
h.	**σ́** σ **σ̀** σ **σ̀** σ **σ̀** σ σ	yúmaɻìŋkamàratʲùɻaka	'because of mother-in-law'

- Secondary stress falls on nonfinal odd syllables (counting from left)
- Primary stress falls on the initial syllable

(Hayes 1995:62, Hansen and Hansen 1969:163)

Quantity Insensitive Single and Dual Systems

	#	Name	Main	Secondary	Note
Single	1.	Afrikaans	1L	None	
	2.	Abun	1R	None	
	3.	Diegueño	1R	None	A
	4.	Agul North	2L	None	
	5.	Alawa	2R	None	
	6.	Mohawk	2R	None	A
	7.	Cora	1L (2-), 3R (3+)	None	
	8.	Paamese	3R (3+), 1L (2-)	None	A
	9.	Bhojpuri	3R (4+), 2R (3-)	Not included	
	10.	Icua Tupi	3R (5+), 2R (4-)	None	
	11.	Bulgarian	lexical	None	
Dual	12.	Gugu-Yalanji	1L	2R	
	13.	Sorbian	1L	None (3-), 2R (4+)	
	14.	Walmatjari	1L	2R or 3R (5+), 2R (4), None (3-)	
	15.	Mingrelian	1L	3R (4+), None (3-)	
	16.	Armenian	1R	1L	
	17.	Udihe	1R	None (2-), 1L (3+)	
	18.	Anyula	2R	1L (4+), None (3-)	
	19.	Georgian	3R (3+), 2R (2-)	1L (5+), None (4-)	

- See appendix for Notes and a discussion of the code used to describe stress placement.

Quantity Insensitive Binary and Ternary Systems

	#	Name	Main	Secondary	Note
Binary	20.	Bagandji	1L	i2@1L	
	21.	Maranungku	1L	i2@1L	A
	22.	Asmat	1R	i2@1R	
	23.	Araucanian	2L	i2@2L	
	24.	Anejom	2R	i2@2R	
	25.	Cavineña	2R	i2@2R	A
w/ Lapse	26.	Anguthimri	1L	i2@1L, no 1R	
	27.	Bidyara Gungabula	1L	i2@1L, no 1R	A
	28.	Burum	1L	i2@1L, optional no 1R	
	29.	Garawa	1L	i2@2R, 1L, no 2L	
	30.	Indonesian	2R	i2@2R, 1L, no 2L (4+), None (3-)	
	31.	Piro	2R	i2@1L, 2R, no 3R	
	32.	Malakmalak	12@sL (3+), 1L (3-)	i2@2R (3+), None (3-)	
w/ Clash	33.	Gosiute Shoshone	1L	i2@1L, 1R	
	34.	Tauya	1R	i2@1R, 1L	
	35.	Southern Paiute	2L (3+), 1L (2-)	i2@2L, 2R, no 1R (3+), None (2-)	A
	36.	Biangai	2R	i2@2R, 1L	
	37.	Central Alaskan Yupik	1R	i2@2L	A
Ternary	38.	Cayuvava	1L (2-), 3R (3+)	None (2-), i3@3R (3+)	A
	39.	Ioway-Oto	2L	i3@2L	

Example 2: Latin (Quantity-Sensitive Bounded Single)

- If the penult is heavy (CVC, CVː), stress falls on the penult. Otherwise stress falls on the antepenult
- Initial stress in disyllables

a.	L H **H́** H	d.	L L H **Ĺ** L L	g.	**H́** H
b.	H L **H́** H	e.	**Ĺ** L H	h.	**Ĺ** H
c.	L L **H́** H	f.	L **H́** L H	i.	**Ĺ** L

Hayes (1995:50) citing Mester (1992) and Jacobs (1989)

Quantity Sensitive Single, Dual and Multiple Systems

	#	Name	Main	Secondary	Note
Single	66.	Maidu	12/2L	Not included	
	67.	Hopi	12/2L	None	B
	68.	English verbs	12/2R	Not included	
	69.	Kawaiisu	12/2R	None	B
	70.	Shoshone T umpisa	21/1L	Not included	
	71.	Javanese	21/1R	None	
	72.	Manobo Sarangani (per M & M)	21/1R	None	B
	73.	Awadhi	21/2R	None	B
	74.	Malay (per Lewis)	23/3R (3+), 12/2L (2-)	None	
	75.	Latin Classical	23/3R (3+), 1L (2-)	None	B
	76.	Hebrew Tiberian	12/21/1R	Not included	
	77.	English (nouns per Pater)	1@w3/234@sR	i('H,'LL)R	
	78.	Arabic Cairene	1@w3/23@sR	None	B
	79.	Arabic Damascene	1@w3/23R	None	
	80.	Arabic Cyrenaican Bedouin	1@w3/23@sR(3+)	i(H',LX')L (invs) (3+)	B
			12/1R (2-)	None (2-)	
	81.	Hindi (per Fairbanks)	12/2/34@sR (3+)	i('H,'LL)R (invs) (3+)	
			1L (2-)	None (2-)	
	82.	Pirahã	123/123/ 123/123/1R	None	
Dual	83.	Maithili	213/2R	1L	B
Multiple	84.	Cambodian	1R	H	B, G
	85.	Yapese	12/1R	H	
	86.	Tongan	12/2R	H	B
	87.	Miwok Sierra	12/2L	H	B
	88.	Gurkhali	12/1L	m $<$ H	

Quantity Sensitive Binary and Ternary Systems

	#	Name	Main	Secondary	Note
Binary	89.	Aranda Western	12/2L (3+), 1L (2-)	i2@m, no 1R (3+), None (2-)	
	90.	Nyawaygi	12@sL	i('H,'LL)R	
	91.	Wargamay	12@sL	i('H,'LL)R, no 'H'L	B, I
	92.	Romansh Berguener	12/2R	i('H,'LL)L	
	93.	Greek Ancient	12/2R	i('H,'LL)R	
	94.	Fijian	12/2R	i('H,'LL)R	B
	95.	Romanian	12/2R	i2@m	
	96.	Seminole Creek	12@sR	i(H',LX')L	B
	97.	Aklan	21/1R	i('H,'LL)@m < m	
	98.	Malecite / Passamaquoddy	23@sR	i(H',LX')L	
	99.	Munsee	23@sR (3+)	i(H',LX')L, no 1R (3+)	
			12/2L (2-)	None (2-)	
	100.	Cayuga	23@sR (3+)	i(H',LX')L, no 1R (3+)	B
			1/0L (2-)	None (2-)	
	101.	Manam	123/23/3R	i('H,'LL)@m < m	
	102.	Arabic Negev Bedouin	1@w3/23@sR (3+)	i(H',LX')L (invs) (3+)	
			12/1R (2-)	None (2-)	
	103.	Arabic Bani-Hassan	1@w3/ 23@w2/2R	i('H,'LL)@m < m	
	104.	Arabic Palestinian	1/2/34@sR (3+)	i('H,'LL)L < m	B
			1@w3/9R (2-)		
	105.	Asheninca	234/324@s/ 324@sR	i('H,'LL)L < m (w2=H)	B
	106.	Dutch	1@w4/23@sR	i('H,'LL)R	
Ternary	107.	Estonian	1L	i('HX,'XLL, 'LL)L	
	108.	Hungarian	1L	i('HX,'XLL, 'LL)L, no 1R	
	109.	Sentani	12/2R	i('HX, 'XLX)@mL	

Example 3: Kwakwala (Quantity-Sensitive Unbounded Single)

Leftmost Heavy Otherwise Rightmost

- Stress the heavy syllable closest to the left edge. If there is no heavy syllable, stress the rightmost syllable.

a.	**Ḣ́** H H	d.	L **Ĺ**
b.	L L **H́** L L	e.	L L **Ĺ**
c.	L L L **H́**	f.	L L L **Ĺ**

Walker (2000:21) citing Zec (1994)

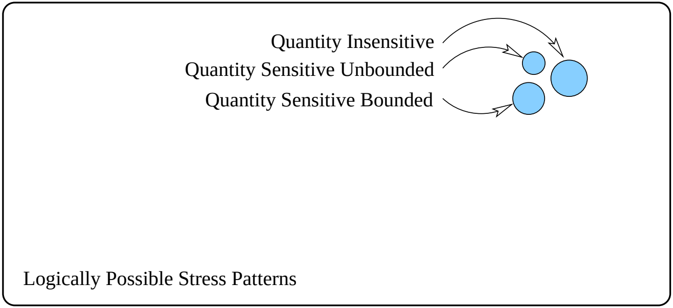
Quantity Sensitive Unbounded Systems

	#	Name	Main	Secondary	Note
LHOL	40.	Murik	12..89/1L	None	C
	41.	Lithuanian	12..89/1L	None	D
	42.	Amele	12..89/1L	None	
	43.	Mongolian Khalkha (per Street)	12..89/1L	H	
	44.	Yidin	12..89/1L	i2@m	B
	45.	Kashmiri	12..78/ 12..78/1L	None	
	46.	Maori	12..89/ 12..89/1L	Not included	
	47.	Mongolian Khalkha (per Stuart)	12..89/2L	None	
LHOR	48.	Komi	12..89/9L	None	
RHOL	49.	Kuuku-Yau	12..89/9R	1L, H	
	50.	Nubian Dongolese	23..89/9R	H	
	51.	Mongolian Khalkha (per Bosson)	23..891/9R	H	
	52.	Buriat	23..891/9R	1L, H	
	53.	Arabic Classical	1/23..89/ 9R	None	
	54.	Cheremis Eastern	23..89/9R	None	
	55.	Chuvash	12..89/9R	None	
RHOR	56.	Golin	12..89/1R	None	
	57.	Cheremis Meadow	1/23..891/ 1R	None	
	58.	Mam	12..89/12/ 2R	None	B
	59.	Klamath	12..89/23/ 3R	if 3R=SH,2R=H then 2R	
	60.	Seneca	see note	i2@m < m	E
	61.	Cheremis Mountain	23..89/2R	None	
	62.	Hindi (per Jones)	23..891/2R	None	
	63.	Sindhi	23..891/2R	H	
	64.	Bhojpuri (per Shukla and Tiwari)	23..891/2R	'Hm'H, m'LL, 1L	
	65.	Hindi (per Kelkar)	23..891/ 23..891/2R	H, i('LL)@m < m m < i(LL')@m	

Stress Patterns *Not* Found in the World's Languages

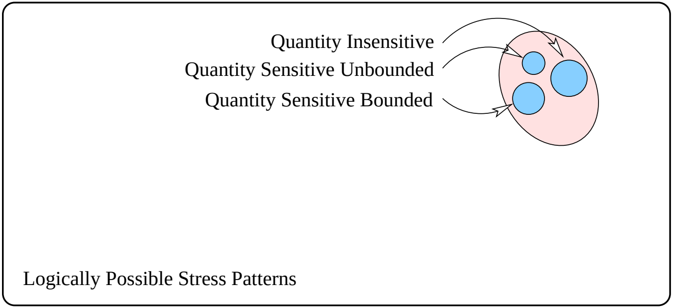
- The middle syllable gets a beat (Single)
- Every fourth syllable gets a beat (Quaternary)
- Every fifth syllable gets a beat (Quinary)
- ...
- The prime-numbered syllables (2,3,5,7,11,...) get a beat
- The prime-numbered syllables minus one (1,2,4,6,10,...) get a beat
- ...

The Space of Logically Possible Stress Systems



- Despite the extensive variation, stress patterns are not arbitrary.
- What property do the attested stress patterns share that separates them from other logically possible stress patterns?

The Space of Logically Possible Stress Systems

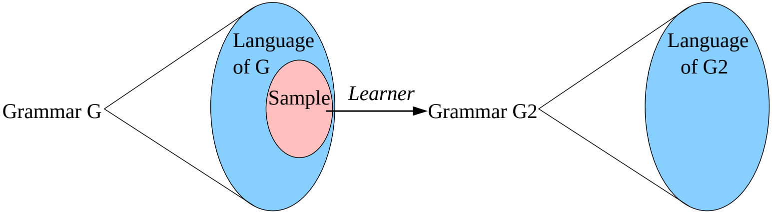


- Despite the extensive variation, stress patterns are not arbitrary.
- What property do the attested stress patterns share that separates them from other logically possible stress patterns?

Learning in Phonology

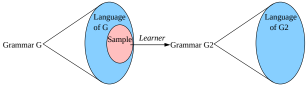
- **Learning in Optimality Theory**
  Tesar (1995), Boersma (1997), Tesar (1998), Tesar and Smolensky (1998), Hayes (1999), Boersma and Hayes (2001), Lin (2002), Pater and Tessier (2003), Pater (2004), Prince and Tesar (2004), Hayes (2004), Riggle (2004), Alderete et al. (2005), Merchant and Tesar (to appear), Wilson (2006), Riggle (2006), Tessier (2006)
- **Learning in Principles and Parameters**
  Wexler and Culicover (1980), Dresher and Kaye (1990), Dresher (1999), Niyogi (2006)
- **Learning Phonological Rules**
  Gildea and Jurafsky (1996), Albright and Hayes (2002, 2003)
- **Learning Phonotactics**
  Ellison (1992), Goldsmith (1994), Frisch (1996), Coleman and Pierrehumbert (1997), Frisch et al. (2004), Albright (2006,2009), Goldsmith (2006), Hayes and Wilson (2008), Heinz (2007, 2009, to appear)

The Learning Framework



- What is *Learner* so that Language of G2 = Language of G?
- How does the learner generalize?
- (Nowak et al. 2002, Niyogi 2006, Jain et al. 1999, Osherson et al. 1986)

Inductive Learning and the Hypothesis Space



- Learning cannot take place unless the hypothesis space is restricted.
- G2 is not drawn from an unrestricted set of possible grammars.

- The hypotheses available to the learner ultimately determine:
  (1) the kinds of generalizations made
  (2) the range of possible natural language patterns
- Under this perspective, Universal Grammar (UG) is the set of available hypotheses.

Different Kinds of Hypothesis Spaces are Learned Differently.

- The set of syntactic hypotheses available to children is not the same as the set of phonological hypotheses available to children.
- The two domains do not have the same kind of patterns and so we expect them to have different kinds of learners.
- The learner and the class of patterns to be learned are in an intimate, not accidental, relationship.

Factoring the Learning Problem

- We can study how learning or generalization occurs by *isolating* factors which play a role in the learning process.
- What are some of the relevant factors for phonotactic learning?
  (1) Sociolinguistic factors . . .
  (2) Articulatory, perceptual biases . . .
  (3) Similarity, locality, . . .
- We should ask: How can any one particular factor benefit learning (in some domain)?

Locality in Phonology

- “Consider first the role of counting in grammar. How long may a count run? General considerations of locality, . . . suggest that the answer is probably ‘up to two’: a rule may fix on one specified element and examine a structurally adjacent element and no other.” (McCarthy and Prince 1986:1)
- “. . . the well-established generalization that linguistic rules do not count beyond two . . . ” (Kenstowicz 1994:597)
- “. . . it was felt that phonological processes are essentially local and that all cases of nonlocality should derive from universal properties of rule application” (Halle and Vergnaud 1987:ix)
- “Metrical theory forms part of a general research program to define the ways in which phonological rules may apply non-locally by characterizing such rules as local with respect to a particular representation.” (Hayes 1995:34)

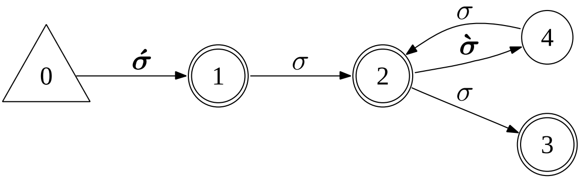
Locality and Learning

- What contribution can this “well-established generalization” make to learning stress patterns?

Representing Stress Patterns with Finite-state Machines

- Stress patterns are *regular*—that is, describable by a finite-state acceptor (Johnson 1972, Kaplan and Kay 1994, Ellison 1992, Eisner 1997, Albro 1998, 2005, Karttunen 1998, Frank and Satta 1998, Riggle 2004, Idsardi 2005, Karttunen 2006)
- Finite-state acceptors
  (1) accept or reject words. So it meets the minimum requirement for a phonotactic grammar– a device that at least answers Yes or No when asked if some word is possible. (Chomsky and Halle 1968, Halle 1978)
  (2) can be related to finite state OT models, which allow us to compute a phonotactic finite-state acceptor (Riggle 2004), which becomes the target grammar for the learner.
  (3) are well-defined and can be manipulated (Hopcroft et. al. 2001).
  (4) Insights in this doman can be extended if more complex grammars are needed (Albro 2005).

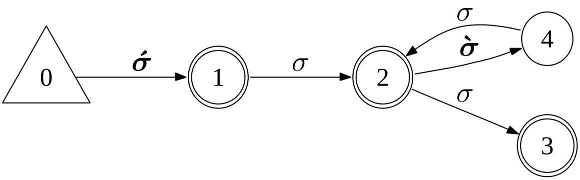
Example 1: Pintupi (Quantity-insensitive Binary)



Accepts	Rejects
$\boldsymbol{\acute{\sigma}}$	
$\boldsymbol{\acute{\sigma}}\ \sigma$	
$\boldsymbol{\acute{\sigma}}\ \sigma\ \sigma$	
$\boldsymbol{\acute{\sigma}}\ \sigma\ \boldsymbol{\grave{\sigma}}\ \sigma$	
$\boldsymbol{\acute{\sigma}}\ \sigma\ \boldsymbol{\grave{\sigma}}\ \sigma\ \sigma$	
...	

- Every word the machine accepts obeys the rule.
- Every word the machine rejects violates it.

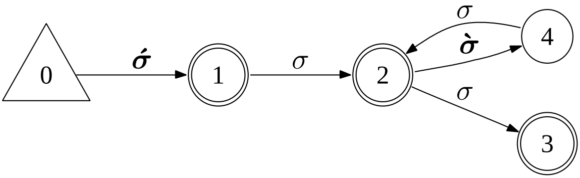
Example 1: Pintupi (Quantity-insensitive Binary)



Accepts	Rejects
	σ́ σ σ̀
	σ́ σ σ σ̀ σ
	σ σ σ σ σ́ σ
	...

- Every word the machine accepts obeys the rule.
- Every word the machine rejects violates it.

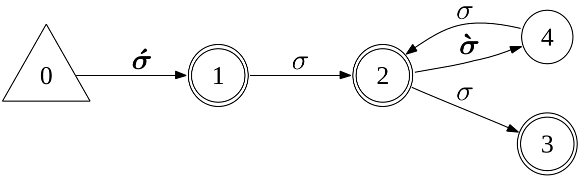
Example 1: Pintupi (Quantity-insensitive Binary)



Accepts	Rejects
σ́	σ́ σ σ̀
σ́ σ	σ́ σ σ σ̀ σ
σ́ σ σ	σ σ σ σ σ́ σ
σ́ σ σ̀ σ	. . .
σ́ σ σ̀ σ σ	
. . .	

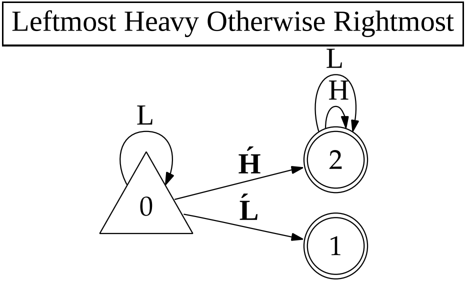
- Every word the machine accepts obeys the rule.
- Every word the machine rejects violates it.

Example 1: Pintupi (Quantity-insensitive Binary)



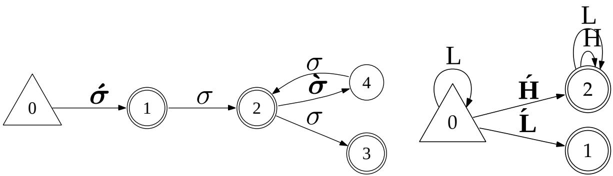
- Also note that if the (different) OT analyses of the Pintupi pattern above given in Gordon (2002) and Tesar (1998) were encoded in finite-state OT, Riggle's (2004) algorithm yields the (same) phonotactic acceptor above.

Example 2: Kwakwala (Quantity-sensitive Unbounded)



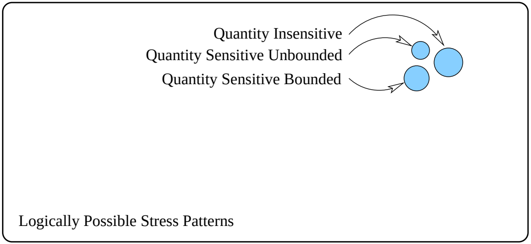
- Likewise, if the (different) OT analyses of the Kwakwala pattern above given in Walker (2000) and Baković (2004) were encoded in finite-state OT, Riggle's (2004) algorithm yields the (same) phonotactic acceptor above.

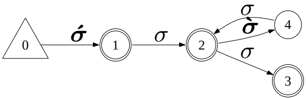
Local Summary



- All stress patterns can be represented with finite state machines, i.e. they are regular patterns.
- These grammars recognize an infinite number of legal words, just like the generative grammars of earlier researchers.

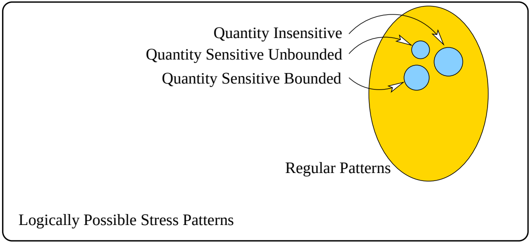
The Learning Question in Context

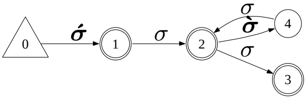




- How can this finite state acceptor be learned from a finite list of Pintupi words?
- This question is not easy. There is no simple 'fix'.

The Learning Question in Context

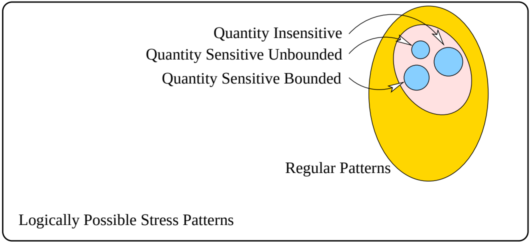


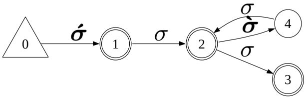


- How can this finite state acceptor be learned from a finite list of Pintupi words?
- This question is not easy. There is no simple 'fix'.

In fact, it is known that no learner can learn the class of patterns describable by finite state machines (Gold 1967, Osherson et al. 1986, Jain et al. 1999).

The Learning Question in Context

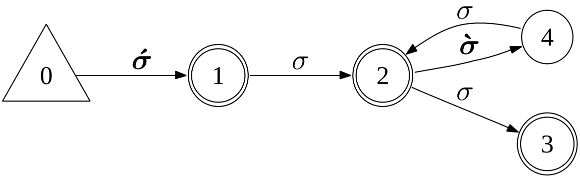




- How can this finite state acceptor be learned from a finite list of Pintupi words?
- This question is not easy. There is no simple 'fix'.

In fact, it is known that no learner can learn the class of patterns describable by finite state machines (Gold 1967, Osherson et al. 1986, Jain et al. 1999).

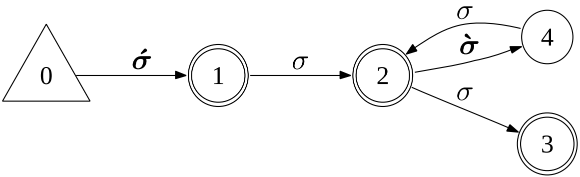
The Answer to the Learning Question



Q: How can this finite state acceptor be learned from the finite list of Pintupi words?

A:
- Generalize by writing smaller and smaller descriptions of the observed forms
- guided by some universal property of the target class…

The Answer to the Learning Question



Q: How can this finite state acceptor be learned from the finite list of Pintupi words?

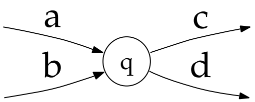
A:
- Generalize by writing smaller and smaller descriptions of the observed forms
- guided by some universal property of the target class. . . locality!

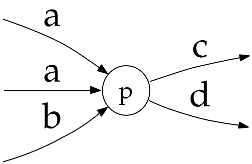
Neighborhoods

- 107 of the 109 stress patterns have have acceptors whose states which are uniquely identified by their local environment.
- The two exceptions are Içuã Tupi (Abrahamson 1968) and Hindi (per Kelkar 1968).
- I call this environment the *neighborhood*. It is:
  1. the set of incoming symbols to the state
  2. the set of outgoing symbols to the state
  3. whether it is a final state or not
  4. whether it is a start state or not

Example of Neighborhoods

- States *p* and *q* have the same neighborhood.





Neighborhood-distinctness

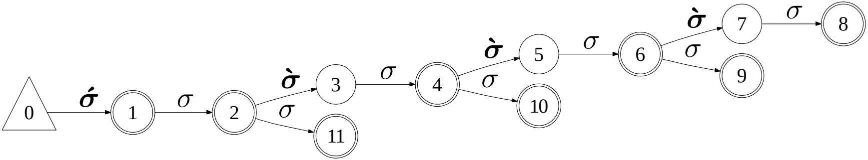
- A language (regular set) is *neighborhood-distinct* iff there is an acceptor for the language such that each state has its own unique neighborhood.
- 107 of the 109 stress patterns are neighborhood-distinct.
- This makes *neighborhood-distinctness* a (near) universal of attested stress patterns.
    - Quantity-insensitive stress patterns
    - Quantity-sensitive bounded stress patterns
    - Quantity-sensitive unbounded stress patterns

Overview of the Neighborhood Learner

- I will describe a simpler version of the learner first, and then describe the actual learner used in this study.
- The learner works in two stages:
  1. Builds a structured representation of the input list of words
  2. Generalizes by merging states which are redundant: i.e. those that have the same local environment—the neighborhood

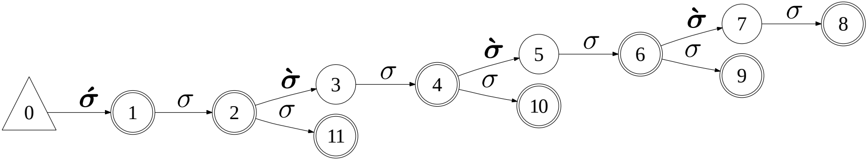
(cf. (Angluin 1982))

The Prefix Tree for Pintupi Stress



- Accepts the words: σ́, σ́ σ, σ́ σ σ, σ́ σ σ̀ σ, σ́ σ σ̀ σ σ, σ́ σ σ̀ σ σ̀ σ, σ́ σ σ̀ σ σ̀ σ σ, σ́ σ σ̀ σ σ̀ σ σ̀ σ
- A structured representation of the input (Angluin 1982).
- It accepts only the forms that have been observed.
- Note that environments are repeated in the tree!

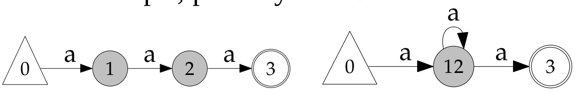
The Prefix Tree for Pintupi Stress



- Accepts the words: σ́, σ́ σ, σ́ σ σ, σ́ σ σ̀ σ, σ́ σ σ̀ σ σ, σ́ σ σ̀ σ σ̀ σ, σ́ σ σ̀ σ σ̀ σ σ, σ́ σ σ̀ σ σ̀ σ σ̀ σ
- A structured representation of the input (Angluin 1982).
- It accepts only the forms that have been observed.
- Note that environments are repeated in the tree!

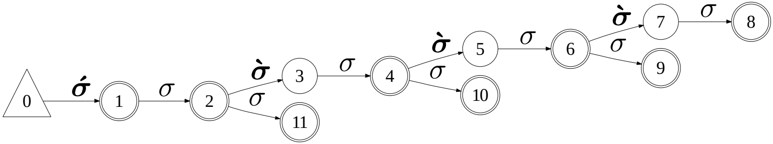
Generalizing by State-merging

- Eliminate redundant environments by *state-merging*.
- This is a process where two states are identified as equivalent and then *merged* (i.e. combined).
- A key concept behind state merging is that transitions are preserved (Hopcroft et al. 2001, Angluin 1982).
- This is one way in which generalizations may occur—because the post-merged machine accepts everything the pre-merged machine accepts, possibly more.



Machine A Machine B

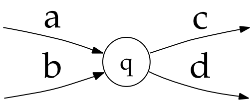
The Learner's State Merging Criteria

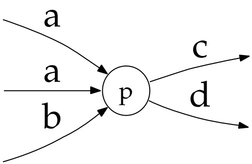


- How does the learner decide whether two states are equivalent in the prefix tree?
- Merge states with the same neighborhood.

Example of Neighborhoods

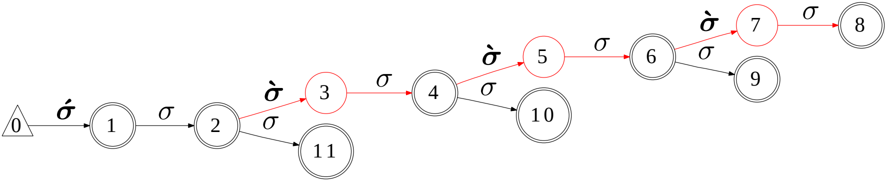
- States *p* and *q* have the same neighborhood.





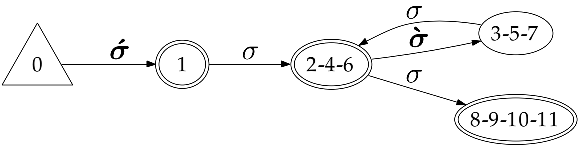
- The learner merges states in the prefix tree with the same neighborhood.

The Prefix Tree for Pintupi Stress



- States 3, 5, and 7 have the same neighborhood.
- So these states are merged.

The Result of Merging Same-Neighborhood States



- The machine above accepts
  $\boxed{\acute{\sigma}}$, $\boxed{\acute{\sigma}\ \sigma}$, $\boxed{\acute{\sigma}\ \sigma\ \sigma}$, $\boxed{\acute{\sigma}\ \sigma\ \grave{\sigma}\ \sigma}$, $\boxed{\acute{\sigma}\ \sigma\ \grave{\sigma}\ \sigma\ \sigma}$, $\boxed{\acute{\sigma}\ \sigma\ \grave{\sigma}\ \sigma\ \grave{\sigma}\ \sigma}$, ...
- The learner has acquired the stress pattern of Pintupi, i.e. it has generalized exactly as desired.
- Each state in the acceptor above has a distinct neighborhood.

Summary of the Forward Neighborhood Learner

(1) Builds a prefix tree of the observed words.

(2) Generalize by merging states which have the same neighborhood (local environment).

(3) The acceptor returned by the algorithm is neighborhood-distinct—every state has a distinct neighborhood.

Quantity Insensitive Single and Dual Systems FL Results

	#	Name	Main	Secondary	Note	FL
Single	1.	Afrikaans	1L	None		✓(4)
	2.	Abun	1R	None		✓(4)
	3.	Diegueño	1R	None	A	✓(4)
	4.	Agul North	2L	None		✓(5)
	5.	Alawa	2R	None		✓(5)
	6.	Mohawk	2R	None	A	✓(5)
	7.	Cora	1L (2-), 3R (3+)	None		✓(6)
	8.	Paamese	3R (3+), 1L (2-)	None	A	×
	9.	Bhojpuri	3R (4+), 2R (3-)	Not included		×
	10.	Icua Tupi	3R (5+), 2R (4-)	None		×
	11.	Bulgarian	lexical	None		✓(4)
Dual	12.	Gugu-Yalanji	1L	2R		✓(6)
	13.	Sorbian	1L	None (3-), 2R (4+)		×
	14.	Walmatjari	1L	2R or 3R (5+), 2R (4), None (3-)		×
	15.	Mingrelian	1L	3R (4+), None (3-)		×
	16.	Armenian	1R	1L		✓(5)
	17.	Udihe	1R	None (2-), 1L (3+)		✓(5)
	18.	Anyula	2R	1L (4+), None (3-)		✓(6)
	19.	Georgian	3R (3+), 2R (2-)	1L (5+), None (4-)		✓(7)

Quantity Insensitive Binary and Ternary Systems FL Results

	#	Name	Main	Secondary	Note	FL
Binary	20.	Bagandji	1L	i2@1L		✓(5)
	21.	Maranungku	1L	i2@1L	A	✓(5)
	22.	Asmat	1R	i2@1R		✓(5)
	23.	Araucanian	2L	i2@2L		✓(6)
	24.	Anejom	2R	i2@2R		✓(6)
	25.	Cavineña	2R	i2@2R	A	✓(6)
w/ Lapse	26.	Anguthimri	1L	i2@1L, no 1R		✓(6)
	27.	Bidyara Gungabula	1L	i2@1L, no 1R	A	✓(6)
	28.	Burum	1L	i2@1L, optional no 1R		✓(5)
	29.	Garawa	1L	i2@2R, 1L, no 2L		✓(6)
	30.	Indonesian	2R	i2@2R, 1L, no 2L (4+), None (3-)		×
	31.	Piro	2R	i2@1L, 2R, no 3R		✓(6)
	32.	Malakmalak	12@sL (3+), 1L (3-)	i2@2R (3+), None (3-)		✓(6)
w/ Clash	33.	Gosiute Shoshone	1L	i2@1L, 1R		✓(5)
	34.	Tauya	1R	i2@1R, 1L		✓(6)
	35.	Southern Paiute	2L (3+), 1L (2-)	i2@2L, 2R, no 1R (3+), None (2-)	A	✓(7)
	36.	Biangai	2R	i2@2R, 1L		✓(7)
	37.	Central Alaskan Yupik	1R	i2@2L	A	✓(6)
Ternary	38.	Cayuvava	1L (2-), 3R (3+)	None (2-), i3@3R (3+)	A	×
	39.	Ioway-Oto	2L	i3@2L		✓(7)

Quantity Sensitive Unbounded Systems FL Results

	#	Name	Main	Secondary	Note	FL
LHOL	40.	Murik	12..89/1L	None	C	✓(4)
	41.	Lithuanian	12..89/1L	None	D	✓(4)
	42.	Amele	12..89/1L	None		✓(4)
	43.	Mongolian Khalkha (per Street)	12..89/1L	H		✓(4)
	44.	Yidin	12..89/1L	i2@m	B	✓(5)
	45.	Kashmiri	12..78/ 12..78/1L	None		×
	46.	Maori	12..89/ 12..89/1L	Not included		✓(5)
	47.	Mongolian Khalkha (per Stuart)	12..89/2L	None		✓(5)
LHOR	48.	Komi	12..89/9L	None		✓(4)
RHOL	49.	Kuuku-Yau	12..89/9R	1L, H		✓(5)
	50.	Nubian Dongolese	23..89/9R	H		✓(5)
	51.	Mongolian Khalkha (per Bosson)	23..891/9R	H		×
	52.	Buriat	23..891/9R	1L, H		×
	53.	Arabic Classical	1/23..89/ 9R	None		✓(4)
	54.	Cheremis Eastern	23..89/9R	None		✓(5)
	55.	Chuvash	12..89/9R	None		✓(4)
RHOR	56.	Golin	12..89/1R	None		✓(5)
	57.	Cheremis Meadow	1/23..891/ 1R	None		✓(5)
	58.	Mam	12..89/12/ 2R	None	B	✓(5)
	59.	Klamath	12..89/23/ 3R	if 3R=SH,2R=H then 2R		×
	60.	Seneca	see note	i2@m < m	E	✓(7)
	61.	Cheremis Mountain	23..89/2R	None		✓(6)
	62.	Hindi (per Jones)	23..891/2R	None		×
	63.	Sindhi	23..891/2R	H		×
	64.	Bhojpuri (per Shukla and Tiwari)	23..891/2R	'Hm'H, m'LL, 1L		×
	65.	Hindi (per Kelkar)	23..891/ 23..891/2R	H, i('LL)@m < m m < i(LL')@m		×

QS Single, Dual and Multiple Systems FL Results

	#	Name	Main	Secondary	Note	FL
Single	66.	Maidu	12/2L	Not included		✓(4)
	67.	Hopi	12/2L	None	B	✓(4)
	68.	English verbs	12/2R	Not included		✓(4)
	69.	Kawaiisu	12/2R	None	B	✓(4)
	70.	Shoshone T umpisa	21/1L	Not included		✓(5)
	71.	Javanese	21/1R	None		✓(5)
	72.	Manobo Sarangani (per M & M)	21/1R	None	B	✓(5)
	73.	Awadhi	21/2R	None	B	✓(5)
	74.	Malay (per Lewis)	23/3R (3+), 12/2L (2-)	None		×
	75.	Latin Classical	23/3R (3+), 1L (2-)	None	B	✓(5)
	76.	Hebrew Tiberian	12/21/1R	Not included		✓(4)
	77.	English (nouns per Pater)	1@w3/234@sR	i('H,'LL)R		✓(5)
	78.	Arabic Cairene	1@w3/23@sR	None	B	✓(4)
	79.	Arabic Damascene	1@w3/23R	None		✓(5)
	80.	Arabic Cyrenaican Bedouin	1@w3/23@sR (3+)	i(H',LX')L (invs) (3+)	B	×
			12/1R (2-)	None (2-)		
	81.	Hindi (per Fairbanks)	12/2/34@sR (3+)	i('H,'LL)R (invs) (3+)		×
			1L (2-)	None (2-)		
	82.	Pirahã	123/123/123/123/1R	None		×
Dual	83.	Maithili	213/2R	1L	B	✓(6)
Multiple	84.	Cambodian	1R	H	B, G	✓(5)
	85.	Yapese	12/1R	H		✓(4)
	86.	Tongan	12/2R	H	B	✓(4)
	87.	Miwok Sierra	12/2L	H	B	✓(4)
	88.	Gurkhali	12/1L	$m < H$		✓(4)

Quantity Sensitive Binary and Ternary Systems FL Results

	#	Name	Main	Secondary	Note	FL
Binary	89.	Aranda Western	12/2L (3+), 1L (2-)	i2@m, no 1R (3+), None (2-)		✓(6)
	90.	Nyawaygi	12@sL	i('H,'LL)R		✓(5)
	91.	Wargamay	12@sL	i('H,'LL)R, no 'H'L	B, I	✓(6)
	92.	Romansh Berguener	12/2R	i('H,'LL)L		✓(6)
	93.	Greek Ancient	12/2R	i('H,'LL)R		✓(5)
	94.	Fijian	12/2R	i('H,'LL)R	B	✓(5)
	95.	Romanian	12/2R	i2@m		✓(5)
	96.	Seminole Creek	12@sR	i(H',LX')L	B	✓(5)
	97.	Aklan	21/1R	i('H,'LL)@m $<$ m		✓(5)
	98.	Malecite / Passamaquoddy	23@sR	i(H',LX')L		✓(6)
	99.	Munsee	23@sR (3+) 12/2L (2-)	i(H',LX')L, no 1R (3+) None (2-)		×
	100.	Cayuga	23@sR (3+) 1/0L (2-)	i(H',LX')L, no 1R (3+) None (2-)	B	✓(6)
	101.	Manam	123/23/3R	i('H,'LL)@m $<$ m		✓(5)
	102.	Arabic Negev Bedouin	1@w3/23@sR (3+) 12/1R (2-)	i(H',LX')L (invs) (3+) None (2-)		×
	103.	Arabic Bani-Hassan	1@w3/ 23@w2/2R	i('H,'LL)@m $<$ m		✓(5)
	104.	Arabic Palestinian	1/2/34@sR (3+) 1@w3/9R (2-)	i('H,'LL)L $<$ m	B	×
	105.	Asheninca	234/324@s/ 324@sR	i('H,'LL)L $<$ m (w2=H)	B	×
	106.	Dutch	1@w4/23@sR	i('H,'LL)R		✓(5)
Ternary	107.	Estonian	1L	i('HX,'XLL, 'LL)L		✓(6)
	108.	Hungarian	1L	i('HX,'XLL, 'LL)L, no 1R		✓(6)
	109.	Sentani	12/2R	i('HX, 'XLX)@mL		✓(7)

Summary of the Forward Neighborhood Learner Results

- The learner succeeds for 31 of the 39 quantity-insensitive patterns.
- The learner succeeds for 36 of the 44 quantity-sensitive bounded patterns.
- The learner succeeds for 18 of the 26 quantity-sensitive unbounded patterns.
- But we can do better!

Summary of the Forward Neighborhood Learner Results

- The learner succeeds for 31 of the 39 quantity-insensitive patterns.
- The learner succeeds for 36 of the 44 quantity-sensitive bounded patterns.
- The learner succeeds for 18 of the 26 quantity-sensitive unbounded patterns.
- But we can do better!

Directionality and the Prefix Tree

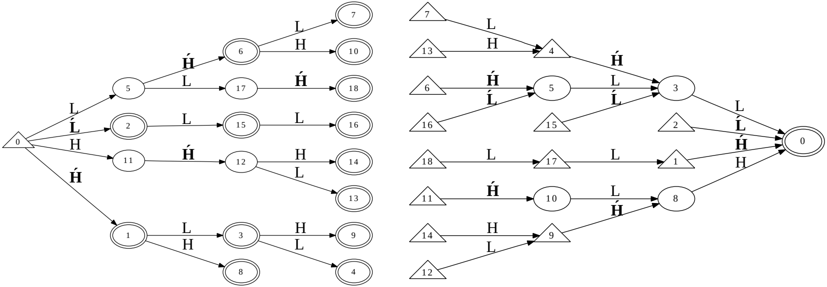
- Many patterns the Forward Learner fails to learn are typically analyzed as having a metrical unit at the right word edge.
  - Içuã Tupi, Lower Sorbian, Walmatjari, Indonesian, Cayuvava, Pirahã, Kashmiri, Buriat, ...
- In each case the learner overgeneralized (i.e. accepted a language strictly larger than the target language).

Elaborating the Forward Learner

- The learner in this study is more elaborate than the Forward Learner.
  - The generalization strategy is the same.
  - But it addresses the inherent left-to-right bias in prefix trees.
  - This must be addressed since stress patterns can be sensitive to *either* word edge (or both).
- Thus the few failures of the Forward Learner are attributed *not* to the generalization strategy but rather to an inherent bias of the (independent) choice of how the input is represented.

Suffix Trees

- If the input were represented with a suffix tree, then the structure obtained has the reverse bias, a right-to-left bias.



Prefix Tree for Buriat Stress Suffix Tree for Buriat Stress
(all words three syllables or less)

- These two representations are not mirror images of each other. They have different structures, though both accept exactly the same (finite) set of words.

The Forward Backward Neighborhood Learner

- The Forward Backward Neighborhood Learner
  (1) Builds a forward prefix tree and merge states with the same neighborhood.
  (2) Builds a suffix tree and merge states with the same neighborhood.
  (3) Intersects these two machines to get the final grammar.

- Intersection of two acceptors A and B results in an acceptor which only accepts words accepted by both A and B (Hopcroft et al. 2001).

Quantity Insensitive Single and Dual Systems FBL Results

	#	Name	Main	Secondary	Note	FBL
Single	1.	Afrikaans	1L	None		✓(4)
	2.	Abun	1R	None		✓(4)
	3.	Diegueño	1R	None	A	✓(4)
	4.	Agul North	2L	None		✓(5)
	5.	Alawa	2R	None		✓(5)
	6.	Mohawk	2R	None	A	✓(5)
	7.	Cora	1L (2-), 3R (3+)	None		✓(6)
	8.	Paamese	3R (3+), 1L (2-)	None	A	✓(6)
	9.	Bhojpuri	3R (4+), 2R (3-)	Not included		✓(6)
	10.	Icua Tupi	3R (5+), 2R (4-)	None		×
	11.	Bulgarian	lexical	None		✓(4)
Dual	12.	Gugu-Yalanji	1L	2R		✓(6)
	13.	Sorbian	1L	None (3-), 2R (4+)		✓(6)
	14.	Walmatjari	1L	2R or 3R (5+), 2R (4), None (3-)		✓(6)
	15.	Mingrelian	1L	3R (4+), None (3-)		×
	16.	Armenian	1R	1L		✓(5)
	17.	Udihe	1R	None (2-), 1L (3+)		✓(6)
	18.	Anyula	2R	1L (4+), None (3-)		✓(7)
	19.	Georgian	3R (3+), 2R (2-)	1L (5+), None (4-)		✓(8)

Quantity Insensitive Binary and Ternary Systems FBL Results

	#	Name	Main	Secondary	Note	FBL
Binary	20.	Bagandji	1L	i2@1L		✓(5)
	21.	Maranungku	1L	i2@1L	A	✓(5)
	22.	Asmat	1R	i2@1R		✓(5)
	23.	Araucanian	2L	i2@2L		✓(6)
	24.	Anejom	2R	i2@2R		✓(6)
	25.	Cavineña	2R	i2@2R	A	✓(6)
w/ Lapse	26.	Anguthimri	1L	i2@1L, no 1R		✓(6)
	27.	Bidyara Gungabula	1L	i2@1L, no 1R	A	✓(6)
	28.	Burum	1L	i2@1L, optional no 1R		✓(5)
	29.	Garawa	1L	i2@2R, 1L, no 2L		✓(6)
	30.	Indonesian	2R	i2@2R, 1L, no 2L (4+), None (3-)		✓(8)
	31.	Piro	2R	i2@1L, 2R, no 3R		✓(7)
	32.	Malakmalak	12@sL (3+), 1L (3-)	i2@2R (3+), None (3-)		✓(6)
w/ Clash	33.	Gosiute Shoshone	1L	i2@1L, 1R		✓(6)
	34.	Tauya	1R	i2@1R, 1L		✓(6)
	35.	Southern Paiute	2L (3+), 1L (2-)	i2@2L, 2R, no 1R (3+), None (2-)	A	✓(8)
	36.	Biangai	2R	i2@2R, 1L		✓(7)
	37.	Central Alaskan Yupik	1R	i2@2L	A	✓(6)
Ternary	38.	Cayuvava	1L (2-), 3R (3+)	None (2-), i3@3R (3+)	A	✓(9)
	39.	Ioway-Oto	2L	i3@2L		✓(8)

Quantity Sensitive Unbounded Systems FBL Results

	#	Name	Main	Secondary	Note	FBL
LHOL	40.	Murik	12..89/1L	None	C	✓(4)
	41.	Lithuanian	12..89/1L	None	D	✓(4)
	42.	Amele	12..89/1L	None		✓(5)
	43.	Mongolian Khalkha (per Street)	12..89/1L	H		✓(5)
	44.	Yidin	12..89/1L	i2@m	B	✓(5)
	45.	Kashmiri	12..78/ 12..78/1L	None		✓(6)
	46.	Maori	12..89/ 12..89/1L	Not included		✓(5)
	47.	Mongolian Khalkha (per Stuart)	12..89/2L	None		✓(5)
LHOR	48.	Komi	12..89/9L	None		✓(4)
RHOL	49.	Kuuku-Yau	12..89/9R	1L, H		✓(5)
	50.	Nubian Dongolese	23..89/9R	H		✓(5)
	51.	Mongolian Khalkha (per Bosson)	23..891/9R	H		✓(5)
	52.	Buriat	23..891/9R	1L, H		✓(6)
	53.	Arabic Classical	1/23..89/ 9R	None		✓(4)
	54.	Cheremis Eastern	23..89/9R	None		✓(5)
	55.	Chuvash	12..89/9R	None		✓(4)
RHOR	56.	Golin	12..89/1R	None		✓(5)
	57.	Cheremis Meadow	1/23..891/ 1R	None		✓(5)
	58.	Mam	12..89/12/ 2R	None	B	✓(5)
	59.	Klamath	12..89/23/ 3R	if 3R=SH,2R=H then 2R		✓(6)
	60.	Seneca	see note	i2@m $<$ m	E	✓(7)
	61.	Cheremis Mountain	23..89/2R	None		✓(6)
	62.	Hindi (per Jones)	23..891/2R	None		✓(6)
	63.	Sindhi	23..891/2R	H		✓(6)
	64.	Bhojpuri (per Shukla and Tiwari)	23..891/2R	’Hm’H, m’LL, 1L		✓(6)
	65.	Hindi (per Kelkar)	23..891/ 23..891/2R	H, i(’LL)@m $<$ m m $<$ i(LL’)@m		×

QS Single, Dual and Multiple Systems FBL Results

	#	Name	Main	Secondary	Note	FBL
Single	66.	Maidu	12/2L	Not included		✓(4)
	67.	Hopi	12/2L	None	B	✓(4)
	68.	English verbs	12/2R	Not included		✓(4)
	69.	Kawaiisu	12/2R	None	B	✓(4)
	70.	Shoshone T umpisa	21/1L	Not included		✓(5)
	71.	Javanese	21/1R	None		✓(5)
	72.	Manobo Sarangani (per M & M)	21/1R	None	B	✓(5)
	73.	Awadhi	21/2R	None	B	✓(5)
	74.	Malay (per Lewis)	23/3R (3+), 12/2L (2-)	None		✓(5)
	75.	Latin Classical	23/3R (3+), 1L (2-)	None	B	✓(5)
	76.	Hebrew Tiberian	12/21/1R	Not included		✓(4)
	77.	English (nouns per Pater)	1@w3/234@sR	i('H,'LL)R		✓(5)
	78.	Arabic Cairene	1@w3/23@sR	None	B	✓(4)
	79.	Arabic Damascene	1@w3/23R	None		✓(5)
	80.	Arabic Cyrenaican Bedouin	1@w3/23@sR (3+)	i(H',LX')L (invs) (3+)	B	×
			12/1R (2-)	None (2-)		
	81.	Hindi (per Fairbanks)	12/2/34@sR (3+)	i('H,'LL)R (invs) (3+)		×
			1L (2-)	None (2-)		
	82.	Pirahã	123/123/123/123/1R	None		×
Dual	83.	Maithili	213/2R	1L	B	✓(6)
Multiple	84.	Cambodian	1R	H	B, G	✓(5)
	85.	Yapese	12/1R	H		✓(4)
	86.	Tongan	12/2R	H	B	✓(4)
	87.	Miwok Sierra	12/2L	H	B	✓(4)
	88.	Gurkhali	12/1L	m < H		✓(4)

Quantity Sensitive Binary and Ternary Systems FBL Results

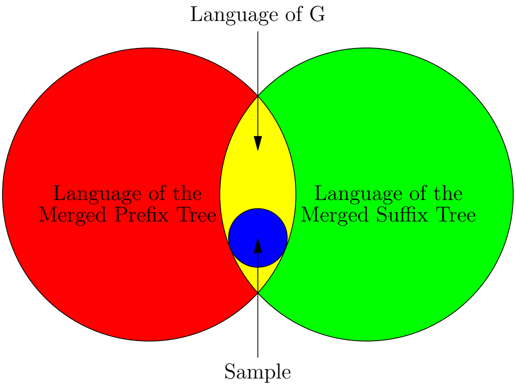
	#	Name	Main	Secondary	Note	FBL
Binary	89.	Aranda Western	12/2L (3+), 1L (2-)	i2@m, no 1R (3+), None (2-)		✓(6)
	90.	Nyawaygi	12@sL	i('H,'LL)R		✓(5)
	91.	Wargamay	12@sL	i('H,'LL)R, no 'H'L	B, I	✓(6)
	92.	Romansh Berguener	12/2R	i('H,'LL)L		✓(6)
	93.	Greek Ancient	12/2R	i('H,'LL)R		✓(5)
	94.	Fijian	12/2R	i('H,'LL)R	B	✓(5)
	95.	Romanian	12/2R	i2@m		✓(5)
	96.	Seminole Creek	12@sR	i(H',LX')L	B	✓(5)
	97.	Aklan	21/1R	i('H,'LL)@m $<$ m		✓(5)
	98.	Malecite / Passamaquoddy	23@sR	i(H',LX')L		✓(6)
	99.	Munsee	23@sR (3+)	i(H',LX')L, no 1R (3+)		✓(6)
			12/2L (2-)	None (2-)		
	100.	Cayuga	23@sR (3+)	i(H',LX')L, no 1R (3+)	B	✓(6)
			1/0L (2-)	None (2-)		
	101.	Manam	123/23/3R	i('H,'LL)@m $<$ m		✓(5)
	102.	Arabic Negev Bedouin	1@w3/23@sR (3+)	i(H',LX')L (invs) (3+)		×
			12/1R (2-)	None (2-)		
	103.	Arabic Bani-Hassan	1@w3/ 23@w2/2R	i('H,'LL)@m $<$ m		✓(5)
	104.	Arabic Palestinian	1/2/34@sR (3+)	i('H,'LL)L $<$ m	B	×
			1@w3/9R (2-)			
	105.	Asheninca	234/324@s/ 324@sR	i('H,'LL)L $<$ m (w2=H)	B	×
	106.	Dutch	1@w4/23@sR	i('H,'LL)R		✓(5)
Ternary	107.	Estonian	1L	i('HX,'XLL, 'LL)L		✓(6)
	108.	Hungarian	1L	i('HX,'XLL, 'LL)L, no 1R		✓(6)
	109.	Sentani	12/2R	i('HX, 'XLX)@mL		✓(7)

Summary of Results of the Forward Backward Learner

- The learner succeeds for 100 of the 109 distinct patterns (414 of 423 language patterns).
  - 37 of the 39 quantity-insensitive patterns
  - 38 of the 44 quantity-sensitive bounded patterns
  - 25 of the 26 quantity-sensitive unbounded patterns
- The patterns it misses will be discussed later.

Why It Works: Intersection keeps robust generalizations

- In only a prefix (suffix) tree is used then sometimes the state merging procedure *overgeneralizes*.
- The Forward-Backward Learner works because it is conservative—it keeps only the robust generalizations—those made in both the prefix and suffix trees (see appendix).



Why It Works: Neighborhood-distinctness

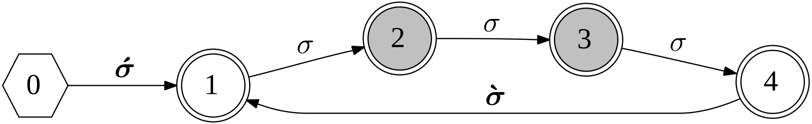
- Every attested stress pattern is neighborhood-distinct except Icua Tupi and Hindi (per Kelkar) (this can be verified upon inspection).
- Because the learner merges states with the same neighborhood, it learns neighborhood-distinct patterns (except Pirahã, Asheninca, ...).
- Thus, the learner is really taking advantage of this previously unnoticed universal property of these grammars: neighborhood-distinctness

Patterns not Learnable by the Neighborhood Learner

- The middle syllable gets a beat (Single)
- Every fourth syllable gets a beat (Quaternary)
- Every fifth syllable gets a beat (Quinary)
- ...
- The prime-numbered syllables (2,3,5,7,11,...) get a beat
- The prime-numbered syllables minus one (1,2,4,6,10,...) get a beat
- ...

These patterns are not neighborhood-distinct.

Example: Quaternary Stress



- States 2 and 3 have the same neighborhood.
- It is not possible to write any finite state machine for this stress pattern where no two states have the same neighborhood.
- The learner fails because it does not distinguish in some sense "exactly three" from "more than two."
- It cannot count past two.

N-gram Models

- In this respect, this learner compares favorably to n-gram models (see Jurafsky and Martin 2008):
  (1) N-gram models cannot learn quantity-sensitive unbounded stress patterns unless they operate on tiers distinguished by syllable weight (e.g. a heavy syllable tier) (Hayes and Wilson 2008).
  (2) A 4-gram model is needed to learn antepenultimate stress patterns, but 4-gram models also admit unattested patterns describable with 4-syllable sized feet.

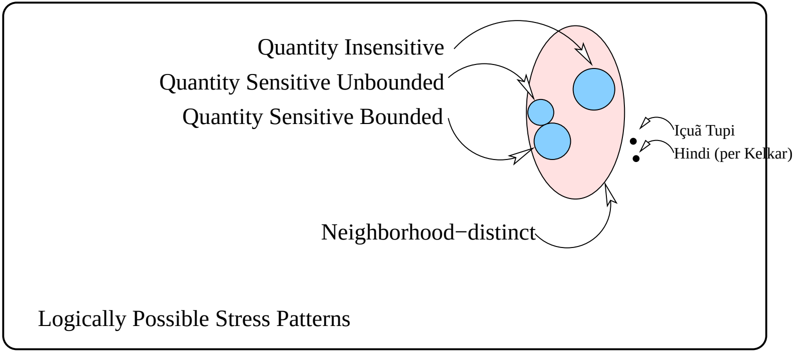
Comparisons to Other Theories

- Neighborhood-distinctness offers a unified analysis of the disparate types of stress patterns
- Consider the explanations other theories offer for ternary patterns:
  1. Only one 'stray' syllable may occur between binary feet (Hayes 1995)
  2. *EXTENDEDLAPSE (Gordon 2002).
- Why binary feet and 'stray' syllables, or why just one 'stray' syllable? And why not *EVENMOREEXTENDEDLAPSE?
  - The answers to these questions fall out from neighborhood-distinctness.

Neighborhood-distinctness

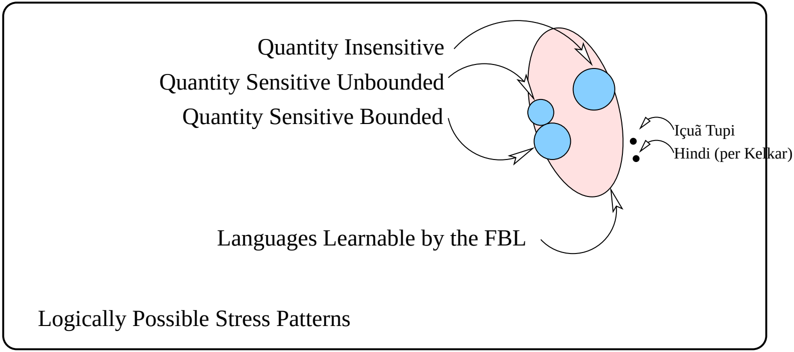
- It is an abstract notion of locality.
- It provides an *inductive principle* by limiting the kinds of generalizations that can be made (e.g. cannot distinguish ‘three’ from ‘more than two’)
- It places real limits on typology: only finitely many languages are neighborhood-distinct (since there are only finitely many neighborhoods given some alphabet).

Local Summary



- Almost all attested stress patterns are neighborhood-distinct.
- The learner uses this property to generalize correctly from finitely many observations.

Local Summary



- Almost all attested stress patterns are neighborhood-distinct.
- The learner uses this property to generalize correctly from finitely many observations.

The Missed Languages

Içuã Tupi, Hindi (per Kelkar), Mingrelian, Palestinian Arabic, Cyranaican Bedouin Arabic, Negev Bedouin Arabic, and Hindi (per Fairbanks), Ashéninca, and Pirahã

- A success was counted only if the learner acquired the exact target pattern.
- How far off were the misses?

For the most part, the differences are slight

Içuã Tupi (not ND) (Abrahamson 1968): Stress falls on the penult in words with four or fewer syllables and on the antepenult in words with five or more syllables.

The FBL acceptor predicts that secondary stress may fall optionally on the penult instead of the antepenult in words five syllables or longer.

For the most part, the differences are slight

Ashéninca (ND) (Payne 1990): The stress pattern is complicated and involves foot extrametricality at the right word edge, among other things.

The FBL's predicts that words ending with a long vowel followed by three syllables with the high front vowel like attested [má:kiriti] 'type of bee' could have two pronunciations: [máːkiriti] and [màːkiríti]

For the most part, the differences are slight

Hindi, per Fairbanks (ND): Stress falls on the initial syllable in disyllabic words.

The only overgeneralization made by the FBL is in disyllabic words ending with a superheavy syllable: the initial syllable may be stressed or the superheavy syllable may be stressed (but not both).

Perhaps actual pattern differs slightly from reported pattern

1. In certain words, there will be optionality
2. In languages currently described as lacking secondary stress, there may in fact be secondary stress

- It is plausible that these might be overlooked (or in the case of secondary stress, difficult to detect) in earlier hypothesis formation.

Interaction with Minimal Word Conditions

- For every quantity-insensitive language considered in the study, the learner succeeds whether or not a (stressed) monosyllable is included in the input sample with one exception.
- Hopi assigns stress to the peninitial syllable but disyllabic words place stress initially.
- If there is a quantity-insensitive language with this stress pattern and a disyllabic minimal word condition (McCarthy and Prince 1986), the Forward Backward Learner fails– this pattern is not neighborhood-distinct.
- To my knowledge, no such language exists.

More Unlearnable Unattested Stress Patterns

- It was discovered that if secondary stress is excluded from the grammars of Klamath (Barker 1963, 1964, Hammond 1986, Hayes 1995) and Seneca (Chafe 1977, Stowell 1979, Prince 1983, Hayes 1995), then the Forward Backward Neighborhood Learner fails to learn these grammars.
- It fails because, in the actual grammars of Klamath and Seneca, the presence of secondary stress distinguishes the neighborhoods of certain states.
- Removing secondary stress causes the patterns to no longer be neighborhood-distinct and hence unlearnable.

Open Questions

- Is this humanlike?

Open Questions

- Is this humanlike?
- This question can be explored in a labratory setting (Buckley and Seidl 2005, Gerken 2006, Cristia and Seidl 2009)

Learnable Unnatural Patterns

- There are stress patterns that can be learned by neighborhood learning which are not considered natural.
  (1) Leftmost Light otherwise Rightmost.
  (2) A stress pattern requiring both lapses and clashes.
  (3) A stress pattern where all syllables have primary stress.
- If these patterns are harder to learn, do we expect the explanation for those facts to follow from considerations of locality?

Learnable Unnatural Patterns

- There are stress patterns that can be learned by neighborhood learning which are not considered natural.
  (1) Leftmost Light otherwise Rightmost.
  (2) A stress pattern requiring both lapses and clashes.
  (3) A stress pattern where all syllables have primary stress.
- If these patterns are harder to learn, do we expect the explanation for those facts to follow from considerations of locality?

Locality is But One Factor in Learning

- This work belongs to a larger research program which is to identify and isolate properties of natural language which are helpful to learning.
- We should ask: What other properties exist
  - which better approximate the class of possible stress systems?
  - and which might assist learning?
- Neighborhood-distinctness is restrictive: most logically possible patterns do not have this property.

Neighborhood-distinctness as a Necessary Condition on Possible Phonotactic Patterns

- It can be shown that other significant classes of phonotactic patterns are also neighborhood-distinct.
  - long distance agreement patterns
  - patterns describable with trigram grammars

Hypothesis: All phonotactic patterns are neighborhood-distinct.

Goals

- The study of learnabilty in phonology is just beginning!
- With any success, we must examine our assumptions and roll them back:
  - Learner is given noiseless sample
  - Output grammar is categorical, not gradient
  - Learner is given word-sized units as input
  - Word-sized units are sequences of symbols (syllables)

The Miracle of Language Learning

- Children somehow internalize productive rules of their language based on limited, incomplete experience.
- How they do this is, how any device can do this, is a nontrivial question that goes to the heart of linguistics.

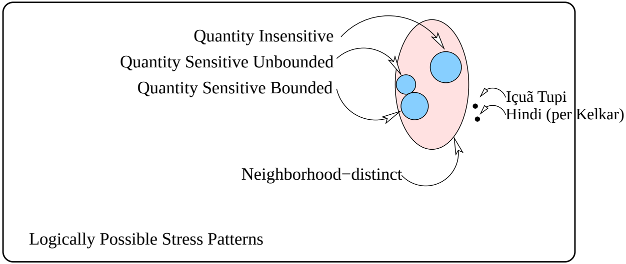
Last Words

- Languages productively assign stress in different ways.
- But not so different! Most logically possible stress assignment rules are unattested (and considered bizarre by linguists).
- Attested patterns are neighborhood-distinct—most unattested ones are not.
- This particular notion of locality, when embedded in the learning model shown here, explains significant aspects of the typology of stress patterns in the world's languages.

Last Words

- Languages productively assign stress in different ways.
- But not so different! Most logically possible stress assignment rules are unattested (and considered bizarre by linguists).
- Attested patterns are neighborhood-distinct—most unattested ones are not.
- This particular notion of locality, when embedded in the learning model shown here, explains significant aspects of the typology of stress patterns in the world's languages.

Acknowledgements



Thank You*

*I thank Bruce Hayes, Ed Stabler, Colin Wilson, and Kie Zuraw for valuable discussion related to this material. I also thank Dieter Gunkel, Greg Kobele, Andy Martin, Katya Pertsova, Rachel Schwarz, Shabnam Shademan, Stephen Tran, and Sarah VanWagenen.