

MUTUALLY CONSTRAINING STRATEGIES IN LANGUAGE LEARNING

Adrian Brasoveanu Nazarré Merchant
Alan Prince Bruce Tesar

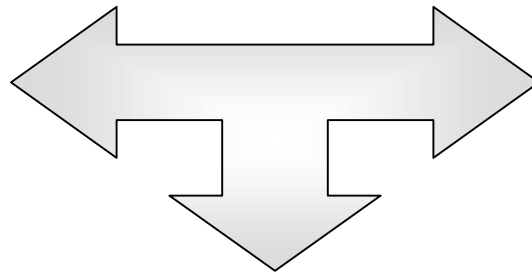
Acknowledgements:

This material is based upon work supported by the National Science Foundation under Grant No. BCS-0083101 to Bruce Tesar and Alan Prince. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Any errors are the sole responsibility of the authors.

THE LEARNING PROBLEM

- no *general* learning mechanism for language acquisition
- Language learning: interaction of universal grammar and partial linguistic structure

universal
grammar



impoverished
stimulus

language specific grammar
full linguistic structure

- Optimality theory: interaction as a major explanatory resource
- grammatical structures:
 - determined by constraint interaction
 - optimal with respect to a constraint ranking
- Language Learning:
 - **Initial state:** partial access to linguistic structure (impoverished stimulus)
 - **Goal:** infer both the full linguistic structure (the lexicon) and the constraint ranking

THE LEARNING HYPOTHESIS: MUTUALLY CONSTRAINING STRATEGIES

STAGE 1

PHONOTACTIC LEARNING

linguistic structure \Rightarrow
partial constraint ranking
information

STAGE 2

LEARNING THE LEXICON

constraint ranking \Rightarrow
partial linguistic structure
information



MUTUAL DEPENDENCE

the interaction converges on complete linguistic
structure and constraint ranking information

Interdisciplinary collaboration:

- psycholinguistics – an explicit theory of language acquisition makes specific psycholinguistic predictions, e.g. different learning stages that interact
- computer science – mutual dependence strategies used in machine learning
- Optimality Theory – connections with neural networks, decision theory, relevance logic

TESTING GROUND

Languages restricted to $[CV]_{\text{root}}[CV]_{\text{suffix}}$

Only 2 interacting features: accent [á] & length [a:]

Inputs	paka	pa:ka:	paka:	pa:ka	Outputs
	páka	pá:ka:	páka:	pá:ka	
	paká	pa:ká:	paká:	pa:ká	
	páká	pá:ká:	páká:	pá:ká	

Only 6 constraints:

WSP Weight-to-Stress Principle

***V:** No Long Vowels

ML Main Stress Left

MR Main Stress Right

F(acc) Faithfulness to Accent

F(len) Faithfulness to Length

**Markedness
Constraints**

**Faithfulness
Constraints**

The constraints generate 14 sets of output forms.

EXAMPLE. The learner has access to the following set of output forms:

{páka, paká, pá:ka, paká:}

The ranking conditions that generate this set of output forms:

F(acc) >> {MR, ML} & F(len) >> *V:

Many constraint rankings are consistent with these conditions; the learner must determine:

- (i) **the input forms** that map to these output forms;
- (ii) **the constraint ranking** for this mapping.

STAGE 1 - PHONOTACTIC LEARNING.

Error driven learning with:

- Biased Constraint Demotion:

rank Markedness high & Faithfulness low

- Identity Map: input form=output form

		F(len)	*V:	WSP	F(acc)	MR	ML
a	/paká/ → [paká]~[páka]	e	e	e	W	W	L
b	/paká:/ → [paká:]~[paká]	W	L	e	e	e	e
c	/pá:ka/ → [pá:ka]~[paká]	W	L	e	W	L	W

Generates the incomplete ranking:

WSP >> F(acc) >> {MR, ML} >> F(len) >> *V:

STAGE 2 – LEARNING THE LEXICON.

- Allow non-identity maps: input form ≠ output form
- Learner becomes morphologically aware.

Possible lexical forms (roots and suffixes)

	[-acc]&[-len]	[-acc]&[+len]	[+acc]&[-len]	[+acc]&[+len]
roots	pa (r ₁)	pa: (r ₂)	pá (r ₃)	pá: (r ₄)
suffixes	ka (s ₁)	ka: (s ₂)	ká (s ₃)	ká: (s ₄)

Target language:

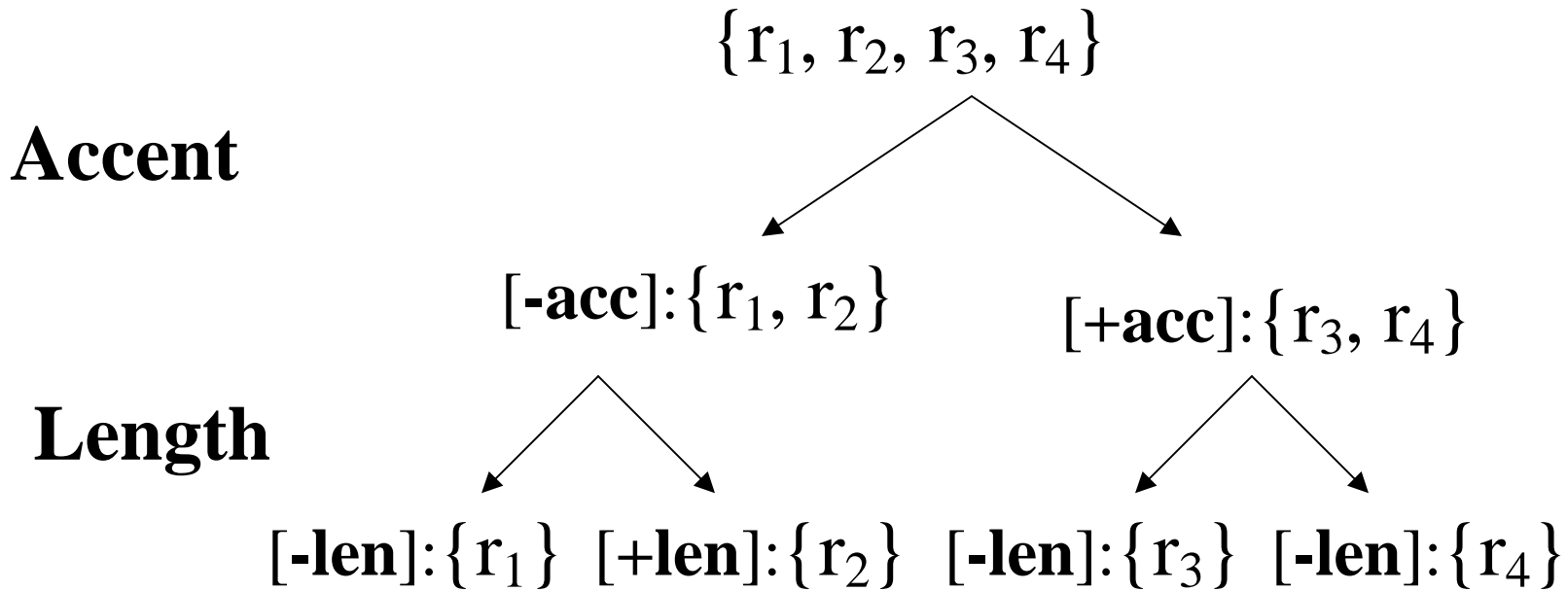
WSP >> F(acc) >> ML >> MR >> F(len) >> *V:

Step 1. Select highest ranking feature: Accent.

Step 2. Group morphemes based on accent & environment, e.g.

		roots			
		r₁	r₂	r₃	r₄
Environments	s₁	[+acc]	[+acc]	[+acc]	[+acc]
	s₂	[-acc]	[+acc]	[+acc]	[+acc]
	s₃	[-acc]	[-acc]	[+acc]	[+acc]
	s₄	[-acc]	[-acc]	[+acc]	[+acc]

Step 3. Underlying form assignment



Step 4. Reapply Biased Constraint Demotion.

Final complete ranking:

WSP>>**F(acc)**>>**ML**>>**MR**>>**F(len)**>>***V**:

Step 5. Correct lexicon and input/output mapping

		Root inputs				
		r₁: pa	r₂: pa:	r₃: pá	r₄: pá:	
Suffix inputs	S₁: ka	páka	pá:ka	páka	pá:ka	Outputs
	S₂: ka:	paká:	pá:ka	páka	pá:ka	
	S₃: ká	paká	paká	páka	pá:ka	
	S₄: ká:	paká:	paká:	páka	pá:ka	