Ling 151/Psych 156A: Acquisition of Language II

Lecture 23
Structure IV
Announcements

Be working on HW8 and the structure review questions

Final review this Friday 3/16/18

Final exam next Friday 3/23/18 between 1:30 and 3:30pm (taken online through Canvas EEE).

Consider taking more language science classes in the future!
Learning with parameters
vs. constraints

metrical phonology
Learning with parameters vs. constraints

- Kitty
- ki TTY

metrical phonology
Learning with parameters vs. constraints

- ✓ a DO ra ble
- × A do RA ble
- × a DO ra BLE
- ✓ Kl tty
- × ki TTY

metrical phonology
Learning with parameters vs. constraints

- ✓ a DO ra ble
- × A do RA ble
- × a DO ra BLE
- ✓ KI tty
- × ki TTY
- ✓ OC to pus
- × oc TO PUS
- × oc to PUS

metrical phonology
Our underlying knowledge representation of the metrical phonology system allows us to generate these metrical stress preferences.
Our underlying knowledge representation of the metrical phonology system allows us to generate these metrical stress preferences.

https://www.youtube.com/watch?v=MdId9wnMNg8&feature=youtu.be
2:03 - 3:00: Metrical stress
Underlying knowledge representations

Account for word-level stress patterns

Observable data: stress contour

Underlying representation?

Points of agreement:

Use **metrical feet**: Units ≥ syllables but (often) smaller than words

Look only at syllable **rimes**

Divide word into **syllables**

(...)(........)

VC V VC

ak əʊs

ak tə pʊs

oc tə pʊs

(Or "rhyming" part, after the onset)

**Pearl** 2016, **Pearl, Ho, & Detrano** 2016
Underlying knowledge representations

Account for word-level stress patterns

Observable data: stress contour

OCtopus

Underlying representation?

Points of cross-linguistic variation:

How to classify syllables

What metrical feet are allowed

How stress interacts with metrical feet

(H L) H (H L L) (S S) S (S S S)

VC V VC
ak ə us
ak tə pus
oc to pus

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Account for word-level stress patterns

Observable data: stress contour

Points of cross-linguistic variation:

https://www.youtube.com/watch?v=MdId9wnMNng8&feature=youtu.be
5:20 - 6:04: Points of variation

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Account for word-level stress patterns

Observable data: stress contour

Points of disagreement:

Underlying grammar = ....?
Underlying knowledge representations

Knowledge representation options

Parameters whose values must be set

English

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Knowledge representation options

Parameters whose values must be set

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Knowledge representation options

Violable constraints that must be ranked

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Knowledge representation options

Violable constraints that must be ranked

Pearl 2016, Pearl, Ho, & Detrano 2016
Parameters

**HV:** Halle & Vergnaud 1987, Dresher 1999, Pearl 2011

5 parameters & 3 sub-parameters

Hypothesis space: 156 grammars

Correct grammar builds compatible contour

Grammar = Set of parameter & sub-parameter values

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Parameters

HV: Halle & Vergnaud 1987, Dresher 1999, Pearl 2011
5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

Correct grammar builds compatible contour

This grammar, comprised of particular parameter values, generates the correct stress contour.

( H  L )  H
OC to pus

Parameter values used:
QS-VC-H, Em-Rt, FtDir-Rt, B-2-Syl, FtHd-Left

...which are the values of the English grammar.

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

**Parameters**

**Hayes**: Hayes 1995

8 parameters

Hypothesis space: 768 grammars

Correct grammar builds compatible contour

OCtopus

- Stress analysis direction
- Extrametricality
- Syllable weight
- Foot directionality
- Word layer end rule
- Degenerate feet
- Foot inventory
- Parsing locality

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Parameters

**Hayes**: Hayes 1995

8 parameters

Hypothesis space: 768 grammars

Correct grammar builds compatible contour

This grammar, comprised of particular parameter values, generates an incorrect stress contour.

H L L

OC TÓ pus

Parameter values used:

Bot, Em-RtCons, VC-H, FtDir-Rt,

PL-Strong, MorTro, DF-Strong, WLER-Rt

...which are the values of the English grammar.

*Pearl 2016, Pearl, Ho, & Detrano 2016*
**Underlying knowledge representations**

**Constraints**

**OT:** Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Premise: Many different candidates for a word’s stress representation and contour are generated and then ranked according to which constraints are violated. Violating higher-ranked constraints is worse than violating lower-ranked constraints.

Best candidate for the correct grammar has a compatible contour

**OCtopus**

**Constraint:** Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Premise: Many different candidates for a word’s stress representation and contour are generated and then ranked according to which constraints are violated. Violating higher-ranked constraints is worse than violating lower-ranked constraints.

**Table:**

<table>
<thead>
<tr>
<th></th>
<th>Higher</th>
<th>Lower</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C1</td>
<td>C2</td>
</tr>
<tr>
<td>(OC to) pus</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>oc (TO pus)</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>(oc TO) pus</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Constraints

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Best candidate for the correct grammar has a compatible contour

**OCtopus**

[YouTube Video](https://www.youtube.com/watch?v=MdId9wnMNg8&feature=youtu.be)

1:29 - 2:40: Intro to constraint ranking

Pearl 2016, Pearl, Ho, & Detrano 2016
**Underlying knowledge representations**

**Constraints**


9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Grammar = ranked ordering of all constraints

Best candidate for the correct grammar has a compatible contour

**metrical phonology**

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Constraints


9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Best candidate for the correct grammar has a compatible contour

OCtopus

Official grammars for languages are often described as partial orderings of constraints.

English grammar

Pearl 2016, Pearl, Ho, & Detrano 2016
This means the “grammar” for a language is often a set of the possible rankings (grammars) that obey those orderings.

Ex: The English “grammar” is compatible with 26 rankings.

Constraints

9 violable constraints
Hypothesis space: 9! rankings = 362,880 grammars

Best candidate for the correct grammar has a compatible contour

OCtopus

Constraint

9 violable constraints
Hypothesis space: 9! rankings = 362,880 grammars

This means the “grammar” for a language is often a set of the possible rankings (grammars) that obey those orderings.

Ex: The English “grammar” is compatible with 26 rankings.

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Constraints

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

Best candidate for the correct grammar has a compatible contour

9 violable constraints

Hypothesis space: $9!$ rankings = 362,880 grammars

Constraints:

- Nonfinality, Parse-$\sigma$
- Foot binarity
- Trochaic
- Weight-to-Stress
- Align left, Align right
- *Sonorant nucleus

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Constraints

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Best candidate for the correct grammar has a compatible contour

Sample candidates

A sample grammar that is a version of the English “grammar”:

(OC to) (PUS)  
(oc TO (PUS)  

(OC to) pus  
occ (TO pus)

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

Constraints


9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

A sample grammar that is a version of the English “grammar”:

Sample candidates

Best candidate for the correct grammar has a compatible contour

Only one candidate left, and it has a compatible contour.

(A sample grammar)

(OC to) (PUS)

(OC to) pus

(oc TO) (PUS)

oc (TO pus)

Pearl 2016, Pearl, Ho, & Detrano 2016
Underlying knowledge representations

**Constraints**

**OT:** Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Best candidate for the correct grammar has a compatible contour

A sample grammar that is a version of the English “grammar”:

English “grammar”

Pearl 2016, Pearl, Ho, & Detrano 2016
Knowledge representation comparison

**HV**: 5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

**Hayes**: 8 parameters
Hypothesis space: 768 grammars

**OT**: 9 violable constraints
Hypothesis space: 362,880 grammars

Pearl 2016, Pearl, Ho, & Detrano 2016
These representations have some similarities, but aren’t obviously using identical variables.

How do we choose among these representations and their English versions?
Knowledge representation comparison

Answer: Let’s see how learnable they are from the English data children typically encounter!

Pearl 2016, Pearl, Ho, & Detrano 2016
Knowledge representation comparison

Acquisition goal: Identify the grammar that can account for the word-level stress patterns in the language

Pearl 2016, Pearl, Ho, & Detrano 2016
Knowledge representation comparison

Observable data: stress contour

All representations: use metrical feet based on syllable rimes

\( \text{OCtopus} \)

\( \text{ak e os} \)

\( \text{VC V VC} \)

Pearl 2016, Pearl, Ho, & Detrano 2016
Knowledge representation comparison

Observable data: stress contour

All representations: use metrical feet based on syllable rimes

Parametric inference: Does this set any values?

Pearl 2016, Pearl, Ho, & Detrano 2016
Knowledge representation comparison

Observable data: stress contour

All representations: use metrical feet based on syllable rimes

OT inference:
Does this implicate any constraint rankings?

Pearl 2016, Pearl, Ho, & Detrano 2016
English metrical phonology is non-trivial to learn because there are many data that are ambiguous for which parameter value or constraint ranking they implicate.

This is generally a problem for acquisition (poverty of the stimulus = the data are compatible with many hypotheses).
Non-trivial because there are many irregularities. This is less common for acquisition – usually there aren’t a lot of exceptions to the system being acquired.
Learning English

irregularities

Interactions with morphology (Chomsky & Halle 1968, Hayes 1982, Kiparsky 1979)

Example: Adding productive morphology doesn’t change the stress pattern, even though all grammars base their stress patterns on the syllables present in the word.

EARly
EARlier
PREtty
PREttiest
senSAtion
senSAtional
senSAtionally
irregularities


Stress contours may be different across syntactic categories, even though the syllabic word form doesn’t change.

<table>
<thead>
<tr>
<th>NOUNS</th>
<th>VERBS</th>
<th>Syllabic word form</th>
</tr>
</thead>
<tbody>
<tr>
<td>CONduct</td>
<td>conDUCT</td>
<td>VC VCC</td>
</tr>
<tr>
<td>DEsert</td>
<td>deSERT</td>
<td>V VCC</td>
</tr>
<tr>
<td>SUspect</td>
<td>suSPECT</td>
<td>V VCC</td>
</tr>
</tbody>
</table>
These *irregularities* can cause *multiple stress contours* to be associated with a syllabic word form. This is problematic for the grammars in these knowledge representations...

Syllabic word form: $V \quad VV$

- KItty
- a WAY
- Vvv
- v VV
- v VV

*Pearl 2016, Pearl, Ho, & Detrano 2016*
These **irregularities** can cause **multiple stress contours** to be associated with a syllabic word form. This is problematic for the grammars in these knowledge representations, since a grammar can only **generate a single stress contour** per syllabic word form...

Syllabic word form:  V  VV

Generate one of these...

- KI tty
- V vv
- a WAY
- v VV
- UH OH
- V VV
These irregularities can cause multiple stress contours to be associated with a syllabic word form. This is problematic for the grammars in these knowledge representations, since a grammar can only generate a single stress contour per syllabic word form or select a single stressed syllabic word form as the best candidate.

Syllabic word form: V VV

Select one of these...

KI tty a WAY UH OH
V vv v VV V VV

Nonfinality, Parse-σ
Foot binarity
Trochaic
Weight-to-Stress
Align left, Align right
*Sonorant nucleus

Pearl 2016, Pearl, Ho, & Detrano 2016
Learning English

Upshot of **multiple stress contours**: No one grammar can account for all the stressed words in the input.

Syllabic word form:  V  VV

KI tty  a WAY  UH OH
V vv   v  VV

But how big of a problem is this in English child-directed speech?

*Pearl 2016, Pearl, Ho, & Detrano 2016*
**Learning English**

**multiple stress contours** = pretty big problem

Analysis of Brent corpus (CHILDES database): 4780 word types (99,968 tokens) of American English speech directed at children between the ages of 6 and 12 months

<table>
<thead>
<tr>
<th>Syllabic word form: V VV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kl tty</td>
</tr>
<tr>
<td>a WAY</td>
</tr>
<tr>
<td>UH OH</td>
</tr>
<tr>
<td>V vv</td>
</tr>
<tr>
<td>v VV</td>
</tr>
<tr>
<td>V VV</td>
</tr>
</tbody>
</table>

**Multiple stress contours**

- HV: 73 of 123 syllabic word forms
- Hayes: 86 of 149 syllabic word forms
- OT: 166 of 452 syllabic word forms

This occurs a lot!

*Pearl 2016, Pearl, Ho, & Detrano 2016*
multiple stress contours = pretty big problem

Syllabic word form: V VV

KI tty       a WAY       UH OH
V vv         v VV        V VV

Acquisition success: Identify the grammar that can account for the word-level stress patterns in the language

This isn’t unreasonable: A grammar is useful because it provides a compact representation of some aspect of the data. Even if it doesn’t cover all the data, covering some is helpful.

Pearl 2016, Pearl, Ho, & Detrano 2016
Knowledge representation comparison

parameters

constraints

English

Computational-level analysis

Pearl 2016, Pearl, Ho, & Detrano 2016
Knowledge representation comparison

Working premise: Rational learners

A learner trying to learn which grammar is the right one for the language will choose the grammar perceived to be the best.

able to account for the most data in the acquisitional intake = most useful to have

Pearl 2016, Pearl, Ho, & Detrano 2016
Once we define the **acquisitional intake**, we can then ask which grammar in the hypothesis space defined by the knowledge representation is **best**, assuming a rational learner that will choose the grammar **compatible with** the most data.
Knowledge representation comparison

It turns out that all three English grammars are compatible with 49-59% of the data in English child-directed speech.

Not too bad!

Pearl 2016, Pearl, Ho, & Detrano 2016
It turns out that all three English grammars are compatible with 49-59% of the data in English child-directed speech.

Not too bad...but can we do better?
Knowledge representation comparison

It turns out that all three English grammars are compatible with 49-59% of the data in English child-directed speech.

Let’s look more closely at the acquisitional intake.

Pearl 2016, Pearl, Ho, & Detrano 2016
Previous working assumption: The learner will try to learn a grammar that can account for all the data encountered.

But we know that’s impossible, because of the multiple stress contours!

KI tty a WAY UH OH
V vv v VV V VV

*Pearl 2016, Pearl, Ho, & Detrano 2016*
Updated working assumption: The learner will try to learn a grammar that can account for all the productive data encountered (Legate & Yang 2012).

**Productive** = the one you use when producing a novel word form

**Acquisitional intake** = only productive data because those are the predictable, rule-based data.

*Pearl 2016, Pearl, Ho, & Detrano 2016*
Updated working assumption: The learner will try to learn a grammar that can account for all the **productive** data encountered (Legate & Yang 2012).

**Productive data filter**

- COO kie
- DA ddy
- FU nny

**Syllabic word form:**

V VV

**Principled way to implement this = Tolerance Principle**

Pearl 2016, Pearl, Ho, & Detrano 2016
A formal way for identifying if there is a dominant rule for a set of items is the Tolerance Principle (Yang 2005, Legate & Yang 2012). This is used to estimate how many exceptions a rule can tolerate in a set before it’s no longer useful for the learner to have the rule.

If there are too many exceptions, it’s better not to have a rule and learn patterns on an individual item basis instead of having a rule that keeps getting violated.
The number of exceptions a rule can tolerate for a set of N items is

\[ \frac{N}{\ln(N)} \]

(Yang 2005, Legate & Yang 2012)
For every syllable word form with multiple stress contours, the learner could assess whether any of those contours is the dominant one (the “rule” for that syllable word form), using the Tolerance Principle.

\[
\frac{N}{\ln(N)}
\]

Tolerance Principle

**Pearl 2016, Pearl, Ho, & Detrano 2016**
If one contour is dominant, the learner should focus on accounting for that pattern, since it’s regular and productive. The grammar should be able to generate it. The other contours can be ignored for purposes of learning the grammar.
If no contour is dominant, the learner should **ignore this syllable word form** for the purposes of learning the grammar since there is no obvious regularity to account for.

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

Productive data filter in action

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

\[ \frac{N}{\ln(N)} \]

English

parameters

HV

Hayes

<table>
<thead>
<tr>
<th>a WAY</th>
<th>KI tty</th>
<th>UH OH</th>
</tr>
</thead>
<tbody>
<tr>
<td>162 types</td>
<td>325 types</td>
<td>19 types</td>
</tr>
<tr>
<td>3713 tokens</td>
<td>12709 tokens</td>
<td>1509 tokens</td>
</tr>
</tbody>
</table>

Hayes, 2016,
These items are good for the HV English grammar.

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

\[ \frac{N}{\ln(N)} \]

English

parameters

HV

Hayes

metrical phonology

criteria

Pearl, Ho, & Detrano 2016

Tolerance Principle

162 types
3713 tokens

325 types
12709 tokens

✗

✓

19 types
1509 tokens

but bad for the Hayes English grammar

Pearl 2016, Pearl, Ho, & Detrano 2016
These items are bad for the HV English grammar.

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

\( \frac{N}{\ln(N)} \)

English

Hayes

HV

parameters

metrical phonology

constraints

162 types
3713 tokens

✓

✗

19 types
1509 tokens

✗

but good for the Hayes grammar.

325 types
12709 tokens

Hayes grammar.

a WAY

KI tty

UH OH

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

These are bad for both grammars.

Pearl 2016, Pearl, Ho, & Detrano 2016
The Tolerance Principle looks at the **word types** for each stress pattern. Each type represents an individual item that might follow the regular stress pattern rule (if there is one).

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

\[ \frac{N}{\ln(N)} \]

English

parameters

HV

Hayes

It doesn’t matter how frequently a type appears (which is what “tokens” indicates).

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

How many items should the stress “rule” apply to?

\[ N = 162 + 325 + 19 = 506 \]

\[ \frac{N}{\ln(N)} \]

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

English

V VV

V

a WAY

KI tty

UH OH

162 types

325 types

19 types

How many exceptions are allowed?

\[
\frac{N}{\ln(N)} = 81
\]

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

English parameters

Hayes

V VV

a WAY

KI tty

UH OH

162 types

325 types

19 types

How many exceptions are allowed?

\[
\frac{N}{\ln(N)} = 81
\]

If this is the dominant pattern, too many exceptions:

\[325 + 19 > 81\]

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

English

parameters

constraints

metrical phonology

HV

Hayes

V  VV

V  VV

a WAY

KI tty

UH OH

162 types

325 types

19 types

If this is the dominant pattern, too many exceptions: 162 + 19 > 81

How many exceptions are allowed?

506 / ln(506) = 81

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

English

parameters

HV

Hayes

How many exceptions are allowed?

\[ \frac{N}{\ln(N)} = 81 \]

If this is the dominant pattern, way too many exceptions:

162 + 325 > 81

Pearl 2016, Pearl, Ho, & Detrano 2016
Learner conclusion: No dominant stress pattern, so none of these syllable word form data should be used to learn the English grammar.

Pearl 2016, Pearl, Ho, & Detrano 2016
This will end up helping both grammars, since they won’t be penalized for the patterns they can’t account for.
Tolerance Principle

\[ \frac{N}{\ln(N)} \]

parameters

constraints

English

metrical phonology

Pearl 2016, Pearl, Ho, & Detrano 2016

V  VV

a  WAY

KI  tty

UH  OH

25 types
976 tokens

316 types
12664 tokens

14 types
1480 tokens
Tolerance Principle

\[ \frac{N}{\ln(N)} \]

parameters

English

constraints

metrical phonology

OT

These items are bad for all English grammars.

25 types 976 tokens

316 types 12664 tokens

14 types 1480 tokens

Pearl 2016, Pearl, Ho, & Detrano 2016
These items are good for most English grammars (21/26).

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

\[ \frac{N}{\ln(N)} \]

English parameters

25 types
976 tokens
✓

316 types
12664 tokens
✓

14 types
1480 tokens
✗

These items are good for a few English grammars (5/26).

Pearl 2016, Pearl, Ho, & Detrano 2016
How many items should the stress “rule” apply to? 
\[ N = 25 + 316 + 14 = 355 \]
Tolerance Principle

How many exceptions are allowed?
\[ \frac{N}{\ln(N)} = 60 \]

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

How many exceptions are allowed?
355 / ln(355) = 60

If this is the dominant pattern, too many exceptions:
316 + 14 > 60

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

How many exceptions are allowed?
\[ \frac{N}{\ln(N)} \]
355 / ln(355) = 60

If this is the dominant pattern, NOT too many exceptions:
25 + 14 < 60

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

How many exceptions are allowed?

\[
\frac{N}{\ln(N)} = \frac{355}{\ln(355)} = 60
\]

If this is the dominant pattern, too many exceptions:

\[25 + 316 > 60\]

Pearl 2016, Pearl, Ho, & Detrano 2016
Tolerance Principle

How many exceptions are allowed?

\[ \frac{N}{\ln(N)} = 60 \]

Pearl 2016, Pearl, Ho, & Detrano 2016
Under the OT syllable representation, there is a dominant stress pattern for this word form. Therefore, this pattern should be accounted for by the English grammar and included in the acquisitional intake.
Unfortunately, for the OT English constraint ranking, this is the only pattern the English grammars can’t account for....this means a learner using the productivity filter would have even more trouble learning.
Knowledge representation comparison

Updated working assumption: The learner will try to learn a grammar that can account for all the productive data encountered (Legate & Yang 2012).

**Principled way to implement this = Tolerance Principle**

<table>
<thead>
<tr>
<th>Syllabic word form:</th>
<th>V  VV</th>
</tr>
</thead>
</table>

**Productive data filter**

- COO kie
- DA ddy
- FU nny

- a WAY
- be LOW
- to DAY

**English**

**Constraints**

**Parameters**
Knowledge representation comparison

Now it turns out that all three English grammars are compatible with 63-87% of the data in productive English child-directed speech.

Not too bad!
And definitely an improvement over 49-59%!
Knowledge representation comparison

Now it turns out that all three English grammars are compatible with 63-87% of the data in productive English child-directed speech.

But how does this compare to other possible grammars in the hypothesis space?

Pearl 2016, Pearl, Ho, & Detrano 2016
Now it turns out that all three English grammars are compatible with 63-87% of the data in productive English child-directed speech.

It turns out that this is worse compatibility than tens (HV), hundreds (Hayes), or tens of thousands (OT) of other possible grammars.

Uh oh!!
Knowledge representation comparison

Now it turns out that all three English grammars are compatible with 63-87% of the data in productive English child-directed speech.

This means the best grammar options for English data aren’t the ones currently proposed for English.

Which ones do better?

Pearl 2016, Pearl, Ho, & Detrano 2016
Other options (differing very slightly) are much more easily learnable - these grammars have much higher English child-directed speech data coverage when a productive data filter is in place: **84-93%**.
Knowledge representation comparison

And two of these other options are better than 96-99% of all the other grammars available! This makes them much more easily learnable.

Implication: Maybe these are a better description of the knowledge representation for the English metrical phonology grammar.

Pearl 2016, Pearl, Ho, & Detrano 2016
Knowledge representation comparison

By modeling acquisition, we provide support for these two theories of English representation in metrical phonology.

Pearl 2016, Pearl, Ho, & Detrano 2016
Linguistic knowledge can be represented by different types of abstract knowledge, such as linguistic parameters or linguistic constraints.

There are many cases where multiple options have been proposed for a knowledge representation, and acquisition modeling can be used to evaluate the options.

For learning English metrical phonology, one important aspect is a productive data filter, because there are so many exceptions.

One principled way to implement a productive data filter is with the Tolerance Principle.
Questions?

You should be able to do all the questions on the structure review questions and all of HW8.
Three knowledge representations

Parametric systems

HV: Halle & Vergnaud 1987, Dresher 1999, Pearl 2011
5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

Correct grammar builds compatible contour

OCtopus

Grammar = Set of parameter & sub-parameter values

Foot headedness

Foot directionality

Extrametricality

Quantity sensitivity

Boundedness
Three knowledge representations

Parametric systems

HV: Halle & Vergnaud 1987, Dresher 1999, Pearl 2011
5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

Correct grammar builds compatible contour
OCtopus

oc to pus
Three knowledge representations

Parametric systems

**HV**: Halle & Vergnaud 1987, Dresher 1999, Pearl 2011

5 parameters & 3 sub-parameters

Hypothesis space: 156 grammars

---

**Quantity sensitivity**
Are syllables all identical, or are they differentiated by syllable weight (into Heavy and Light syllables)?

---

**OCtopus**
Correct grammar builds compatible contour

---

Extrametricality
Foot headedness
Boundedness
Foot directionality

---
Three knowledge representations

Parametric systems

**HV**: Halle & Vergnaud 1987, Dresher 1999, Pearl 2011
5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

-extrametricality
-are all syllables included in the larger units of metrical feet, or are some excluded?

**Extrametricality**

Correct grammar builds compatible contour

\[ \text{OCtopus} \]

\[ \text{HV} \quad \text{L} \quad \text{H} \quad \text{oc} \quad \text{to} \quad \text{pus} \]
Three knowledge representations

Parametric systems

**HV**: Halle & Vergnaud 1987, Dresher 1999, Pearl 2011
5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

---

Correct grammar builds compatible contour

**OCtopus**

**Foot directionality**
Are feet constructed from the left or from the right?
Three knowledge representations

Parametric systems

**HV**: Halle & Vergnaud 1987, Dresher 1999, Pearl 2011
5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

Correct grammar builds compatible contour

(H L) OCtopus

(Boundedness)
How big are metrical feet?
Three knowledge representations

Parametric systems

HV: Halle & Vergnaud 1987, Dresher 1999, Pearl 2011
5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

Correct grammar builds compatible contour

Foot headedness
Which syllable in a foot is stressed?
Three knowledge representations

Parametric systems

**HV**: Halle & Vergnaud 1987, Dresher 1999, Pearl 2011

5 parameters & 3 sub-parameters

Hypothesis space: 156 grammars

Correct grammar builds compatible contour

This grammar, comprised of particular parameter values, generates the correct stress contour.

\[(H \ L) H\]

**OCtopus**

Parameter values used:

- **Quantity sensitive**, VC syllables = Heavy, **Extrametricality** on rightmost syllable, Feet built from the right, Foot = 2 syllables, Leftmost syllable in foot stressed
Three knowledge representations

Parametric systems

5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

Quantity sensitivity
Foot headedness
Boundedness
Extrametricality
Foot directionality

Correct grammar builds compatible contour

This grammar, comprised of particular parameter values, generates the correct stress contour.

(H L) H
OC to pus

Parameter values used:
QS-VC-H, Em-Rt, FtDir-Rt, B-2-Syl, FtHd-Left

...which are the values of the English grammar.
Three knowledge representations

Parametric systems

Hayes: Hayes 1995
8 parameters
Hypothesis space: 768 grammars

Correct grammar builds compatible contour

OCtopus

Stress analysis direction
Extrametricality
Syllable weight
Foot directionality
Word layer end rule
Degenerate feet
Foot inventory
Parsing locality
Three knowledge representations

Hayes: Hayes 1995
8 parameters
Hypothesis space: 768 grammars

Correct grammar builds compatible contour

OCtopus
Three knowledge representations

Parametric systems

**Hayes**: Hayes 1995

8 parameters

Hypothesis space: 768 grammars

- Stress analysis direction
- Extrametricality
- Syllable weight
- Foot directionality
- Parsing locality
- Word layer end rule
- Degenerate feet
- Foot inventory

Stress analysis direction
Are metrical feet created before word-level stress is assigned to the edge syllables or after?

Correct grammar builds compatible contour

OCtopus

(...feet first...)

oc to pus
Three knowledge representations

Parametric systems

Hayes: Hayes 1995
8 parameters
Hypothesis space: 768 grammars

Extrametricality
Are syllables on the edge (or parts of syllables) excluded from metrical feet?

Correct grammar builds compatible contour

OCtopus
Three knowledge representations

Parametric systems

**Hayes**: Hayes 1995

8 parameters

Hypothesis space: 768 grammars

---

**Syllable weight**

Syllables are distinguished into Heavy and Light. Are syllables ending in VC (like *oc*) Heavy or Light?

---

Correct grammar **builds** compatible contour

OCtopus
Three knowledge representations

Parametric systems

Hayes: Hayes 1995
8 parameters
Hypothesis space: 768 grammars

Correct grammar builds compatible contour

Foot directionality
Are metrical feet constructed from the left or the right?
Three knowledge representations

Parametric systems

Hayes: Hayes 1995
8 parameters
Hypothesis space: 768 grammars

Parsing locality
Is one Light syllable skipped between metrical feet?
Three knowledge representations

Parametric systems

**Hayes**: Hayes 1995
8 parameters
Hypothesis space: 768 grammars

Foot inventory
How big are metrical feet?
Where does the stress fall within them?
Three knowledge representations

**Parametric systems**

**Hayes**: Hayes 1995

8 parameters

Hypothesis space: 768 grammars

---

**Degenerate feet**

What do you do with leftover Light syllables if you have any?
Three knowledge representations

Parametric systems

Hayes: Hayes 1995
8 parameters
Hypothesis space: 768 grammars

Correct grammar builds compatible contour

Word layer end rule
Where does word-level stress go if there are multiple stressed syllables? Can leftover Light syllables have word-level stress?
Three knowledge representations

Parametric systems

**Hayes**: Hayes 1995

8 parameters

Hypothesis space: 768 grammars

This grammar, comprised of particular parameter values, generates an incorrect stress contour.

Parameter values used:

- **Bottom-up**, **Extrametricality** on rightmost consonant, **VC** syllables = Heavy, Feet built from the right, **Light** syllables not skipped in between feet, **Foot** = Moraic trochee (2 moras with stress on leftmost), **Single Light edge** syllables not allowed to have stress, **Rightmost syllable** gets main stress
Three knowledge representations

Parametric systems

Hayes: Hayes 1995
8 parameters
Hypothesis space: 768 grammars

Correct grammar builds compatible contour

This grammar, comprised of particular parameter values, generates an incorrect stress contour.

( H ) ( L L )
OC TÓ pus

Parameter values used:
Bot, Em-RtCons, VC-H, FtDir-Rt,
PL-Strong, MorTro, DF-Strong, WLER-Rt

...which are the values of the English grammar.
Three knowledge representations

Constraint-ranking systems


9 violable constraints
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Premise: Many different candidates for a word’s stress representation and contour are generated and then ranked according to which constraints are violated. Violating higher-ranked constraints is worse than violating lower-ranked constraints.

Best candidate for the correct grammar has a compatible contour

<table>
<thead>
<tr>
<th>C1</th>
<th>C2</th>
<th>C3</th>
<th>C4</th>
</tr>
</thead>
<tbody>
<tr>
<td>(OC to) pus</td>
<td></td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>oc (TO pus)</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>(oc TO) pus</td>
<td>*</td>
<td>*</td>
<td></td>
</tr>
</tbody>
</table>
Three knowledge representations

**Constraint-ranking systems**

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Grammar = ranked ordering of all constraints

Best candidate for the correct grammar has a compatible contour

OCtopus
Three knowledge representations

**Constraint-ranking systems**

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

**Best candidate** for the correct grammar has a compatible contour

---

Official grammars for languages are often described as partial orderings of constraints.

---

English grammar
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

This means the “grammar” for a language is often a set of the possible rankings (grammars) that obey those orderings.

Ex: The English “grammar” is compatible with 26 rankings.

Best candidate for the correct grammar has a compatible contour
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Best candidate for the correct grammar has a compatible contour

OCtopus
Three knowledge representations

**Constraint-ranking systems**

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Best candidate for the correct grammar has a compatible contour

**OCtopus**

Nonfinality, Parse-σ

Foot binarity

Trochaic

Weight-to-Stress

Align left, Align right

*Sonorant nucleus

**Nonfinality**

Should the final syllable not be in a metrical foot?

(OC to) (PUS)

(OC to) pus

(oc TO) (PUS)

oc (TO pus)
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

**Parse-σ**

Should all syllables be in metrical feet?

- (OC to) (PUS)
- (oc TO) (PUS)
- (OC to) pus
- oc (TO pus)

Best candidate for the correct grammar has a compatible contour

OCtopus
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: $9!$ rankings = 362,880 grammars

Principle (Rooting): All words must have stress

**Foot binarity**

Should all metrical feet consist of two units?

(OC to) (PUS)  
(oc TO) (PUS)  
(OC to) pus  
oc (TO pus)

**Best candidate** for the correct grammar has a compatible contour

OCtopus
Three knowledge representations

Constraint-ranking systems


9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Best candidate for the correct grammar has a compatible contour

OCtopus

Trochaic

Should metrical feet have stress on the leftmost syllable?

(OC to) (PUS)✓

(oc TO) (PUS)

(OC to) pus✓

oc (TO pus)✓
Three knowledge representations

Constraint-ranking systems


9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Best candidate for the correct grammar has a compatible contour

Nonfinality, Parse-σ
Foot binarity
Trochaic
Weight-to-Stress
Align left, Align right
*Sonorant nucleus

Weight-to-Stress (VV)
Should all VV syllables be stressed?

(ba BY)✔

(BA) by

(BA by)
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Best candidate for the correct grammar has a compatible contour

**OCtopus**

Principle (Rooting):

- Nonfinality, Parse-σ
- Foot binarity
- Trochaic
- Weight-to-Stress
- Align left, Align right
- *Sonorant nucleus

**Weight-to-Stress (VC)**

Should all VC syllables be stressed?

- (OC to) (PUS)
- (oc TO) (PUS)
- (OC to) pus
- oc (TO pus)
Three knowledge representations

Constraint-ranking systems

9 violable constraints
Hypothesis space: 9! rankings = 362,880 grammars
Principle (Rooting): All words must have stress

Best candidate for the correct grammar has a compatible contour

OCtopus

Align left
≈ Should metrical feet include the leftmost syllable?

(OC to) (PUS)
(oc TO) (PUS)
(OC to) pus
oc (TO pus)
Three knowledge representations

**Constraint-ranking systems**

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Nonfinality, Parse-σ
Foot binarity
Trochaic
Weight-to-Stress
Align left, Align right
*Sonorant nucleus

Best candidate for the correct grammar has a compatible contour

**Align right**

≈ Should metrical feet include the rightmost syllable?

- (OC to) (PUS)
- (oc TO) (PUS)
- (OC to) pus
- oc (TO pus)
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: $9!$ rankings = 362,880 grammars

Principle (Rooting): All words must have stress

*Sonorant nucleus*

Should syllables not have sonorants ($m$, $n$, $\eta$, $l$, $r$) as the nucleus?

- your (**SELF**)  
- *(yr** **SELF)**
- *(YOUR) (**SELF)**
- *(YOUR slf)**

Best candidate for the correct grammar has a compatible contour

your **SELF**
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Sample candidates

A sample grammar that is a version of the English “grammar”:

- Nonfinality, Parse-\( \sigma \)
- Foot binarity
- Trochaic
- Weight-to-Stress
- Align left, Align right
- *Sonorant nucleus

Best candidate for the correct grammar has a compatible contour

**OCtopus**

Sample candidates:

- (OC to) (PUS)
- (OC to) pus
- (oc TO) (PUS)
- oc (TO pus)
Three knowledge representations

**Constraint-ranking systems**

**OT:** Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Best candidate for the correct grammar has a compatible contour

OCtopus

**Sample candidates**

A sample grammar that is a version of the English “grammar”:

**Most important:** Metrical feet have stress on the leftmost syllable.

(OC to) (PUS)

(OC to) pus

(oc TO) (PUS)

oc (TO pus)
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Next important: VV syllables are stressed.

Sample candidates

A sample grammar that is a version of the English “grammar”:

Best candidate for the correct grammar has a compatible contour

Sample candidates

(OC to) (PUS)  (OC to) pus

(oc TO) (PUS)  oc (TO pus)
Three knowledge representations

Constraint-ranking systems

**OT**: Hammond 1999, Pater 2000, Tesar & Smolensky 2000

9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Next important: The final syllable is not included in a foot.

Sample candidates

A sample grammar that is a version of the English “grammar”:

Best candidate for the correct grammar has a compatible contour

OCtopus
Three knowledge representations

Constraint-ranking systems


9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

Nonfinality, Parse-σ
Foot binarity
Trochaic
Weight-to-Stress
Align left, Align right
*Sonorant nucleus

Best candidate for the correct grammar has a compatible contour

OCtopus

Only one candidate left, and it has a compatible contour.

Sample candidates

A sample grammar that is a version of the English “grammar”:

(OC to) (PUS)

(OC to) pus

(oc TO) (PUS)

oc (TO pus)
Three knowledge representations

**Constraint-ranking systems**


9 violable constraints

Hypothesis space: 9! rankings = 362,880 grammars

Principle (Rooting): All words must have stress

---

**A sample grammar that is a version of the English “grammar”:**

---

**Best candidate for the correct grammar has a compatible contour**

---

**English “grammar”**
Knowledge representation comparison

**HV**: 5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

**Hayes**: 8 parameters
Hypothesis space: 768 grammars

**OT**: 9 violable constraints
Hypothesis space: 362,880 grammars
**English instantiations**

**HV**: 5 parameters & 4 sub-parameters  
Hypothesis space: 156 grammars

**Hayes**: 8 parameters  
Hypothesis space: 768 grammars

**OT**: 9 violable constraints  
Hypothesis space: 362,880 grammars  
(English = 26 grammars)
Knowledge representation comparison

Each representation assumes certain syllabic distinctions.

**HV**: 5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars

**Hayes**: 8 parameters
Hypothesis space: 768 grammars

**OT**: 9 violable constraints
Hypothesis space: 362,880 grammars
Knowledge representation comparison

**HV**: 5 parameters & 3 sub-parameters
Hypothesis space: 156 grammars
Syllabic distinctions: 3
(short, closed, long)

**Hayes**: 8 parameters
Hypothesis space: 768 grammars
Syllabic distinctions: 4
(short, potentially short, closed, long)

**OT**: 9 violable constraints
Hypothesis space: 362,880 grammars
Syllabic distinctions: 8
(short, sonorant, 4 closed variants, long, super-long)