Ling 151/Psych 156A: Acquisition of Language II

Lecture 16

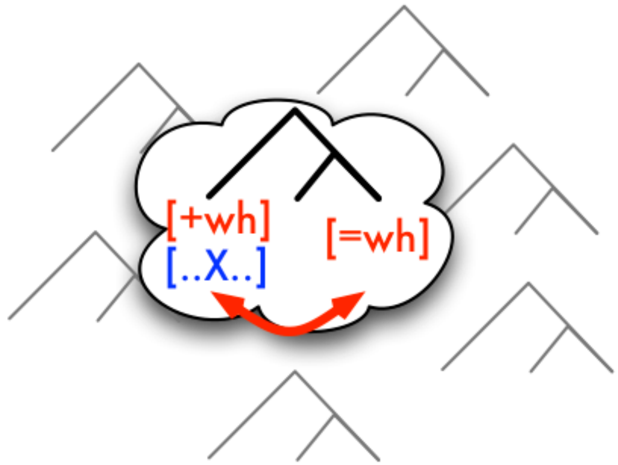
Syntax II

Announcements

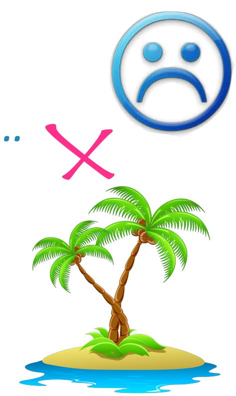
Be working on HW6 (due: 2/26/18)

Be working on review questions for syntax & sentences

Syntax: *wh*-dependencies



Who does…



Syntax: *wh*-dependencies

This kitty was bought as a present for someone.

Lily thinks this kitty is pretty.

What's going on here?

Who does Lily think the kitty for is pretty?

What does Lily think is pretty, and who does she think it's for?

Syntax: *wh*-dependencies

syntax

Who does

Lily think the kitty for is pretty?

What's going on here?

There's a dependency between the *wh*-word *who* and where it's understood (the gap)

Who does Lily think the kitty for ___ is pretty?

This dependency is not allowed in English.

One explanation: The dependency crosses a "syntactic island" (Ross 1967)

Syntax: *wh*-dependencies

syntax

Who does

Lily think the kitty for is pretty?

What's going on here?

syntactic island

Who does Lily think the kitty for ___ is pretty?

http://www.thelingspace.com/episode-66
https://www.youtube.com/watch?v=01uH4XfJx3g

0:39 - 1:34

Syntax: *wh*-dependencies

syntax

Who does *Lily think the kitty for is pretty?*

What's going on here? syntactic island (Ross 1967)

Who does Lily think the kitty for ___ is pretty?

Jack is somewhat tricksy.

He claimed he bought something.

What did Jack make the claim that he bought ___ ?

Syntax: *wh*-dependencies

syntax

Who does ☹
Lily think the kitty for is pretty?

What's going on here? syntactic island (Ross 1967)

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

Jack is somewhat tricksy.

He claimed he bought something.

Elizabeth wondered if he actually did and what it was.

What did Elizabeth wonder whether Jack bought ___ ?

Syntax: *wh*-dependencies

syntax

Who does
Lily think the kitty for is pretty?

What's going on here? syntactic island (Ross 1967)

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

What did Elizabeth wonder whether Jack bought ___ ?

Jack is somewhat tricksy.

He claimed he bought something.

Elizabeth worried it was something dangerous.

What did Elizabeth worry if Jack bought ___ ?

Syntax: *wh*-dependencies

syntax

Who does

Lily think the kitty for is pretty?

What's going on here? syntactic island (Ross 1967)

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

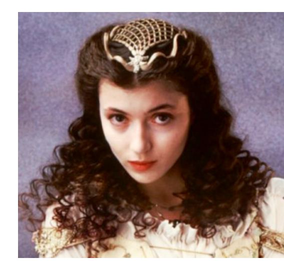
What did Elizabeth wonder whether Jack bought ___ ?

What did Elizabeth worry if Jack bought ___ ?

Jack bought something.

Elizabeth met him afterwards.

Lily asks Elizabeth about it.



What did you meet the pirate who bought ___ ?

Syntax: *wh*-dependencies

syntax

Who does
Lily think the kitty for is pretty?

What's going on here? syntactic island (Ross 1967)

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

What did Elizabeth wonder whether Jack bought ___ ?

What did Elizabeth worry if Jack bought ___ ?

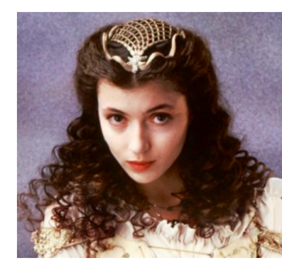
What did you meet the pirate who bought ___ ?

Jack bought something.

Elizabeth was surprised by it.

What did that Jack bought ___ surprise you?

Lily asks Elizabeth about it.



Syntax: *wh*-dependencies

syntax

Who does
Lily think the kitty for is pretty?

What's going on here?  syntactic island (Ross 1967)

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

What did Elizabeth wonder whether Jack bought ___ ?

What did Elizabeth worry if Jack bought ___ ?

What did you meet the pirate who bought ___ ?

What did that Jack bought ___ surprise you ?

Jack bought two things - a kitty and something else.

Elizabeth wants to know about the other thing.

What did you buy a kitty and ___?

Syntax: *wh*-dependencies

syntax

Who does Lily think the kitty for is pretty?

What's going on here? syntactic island (Ross 1967)

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

What did Elizabeth wonder whether Jack bought ___ ?

What did Elizabeth worry if Jack bought ___ ?

What did you meet the pirate who bought ___ ?

What did that Jack bought ___ surprise you ?

What did you buy a kitty and ___ ?

Jack bought a specific kind of kitty.

Elizabeth wants to know about the kind.

Which did you buy ___ kitty ?

Syntax: *wh*-dependencies

syntax

Who does
Lily think the kitty for is pretty?

What's going on here? syntactic island (Ross 1967)

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

What did Elizabeth wonder whether Jack bought ___ ?

What did Elizabeth worry if Jack bought ___ ?

What did you meet the pirate who bought ___ ?

What did that Jack bought ___ surprise you ?

What did you buy a kitty and ___ ?

Which did you buy ___ kitty ?

Important: It's not about the length of the dependency.

(Chomsky 1965, Ross 1967)

Syntax: *wh*-dependencies

syntax

Who does

Lily think the kitty for is pretty?

What's going on here? syntactic island

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

What did Elizabeth wonder whether Jack bought ___ ?

What did Elizabeth worry if Jack bought ___ ?

What did you meet the pirate who bought ___ ?

What did that Jack bought ___ surprise you ?

What did you buy a kitty and ___ ?

Which did you buy ___ kitty ?

Elizabeth

What did Elizabeth think ___ ?

✓

It's not about the length of the dependency.

Syntax: *wh*-dependencies

syntax

Who does
Lily think the kitty for is pretty?

What's going on here? syntactic island

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

What did Elizabeth wonder whether Jack bought ___ ?

What did Elizabeth worry if Jack bought ___ ?

What did you meet the pirate who bought ___ ?

What did that Jack bought ___ surprise you ?

What did you buy a kitty and ___ ?

Which did you buy ___ kitty ?

Jack

Elizabeth

What did Elizabeth think Jack said ___ ?

✓

It's not about the length of the dependency.

Syntax: *wh*-dependencies

syntax

Who does ☹
Lily think the kitty for is pretty?

What's going on here?

syntactic island

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

What did Elizabeth wonder whether Jack bought ___ ?

What did Elizabeth worry if Jack bought ___ ?

What did you meet the pirate who bought ___ ?

What did that Jack bought ___ surprise you ?

What did you buy a kitty and ___ ?

Which did you buy ___ kitty ?

Jack

Elizabeth

Lily

What did Elizabeth think Jack said Lily saw ___ ?

✓

It's not about the length of the dependency.

Syntax: *wh*-dependencies

syntax

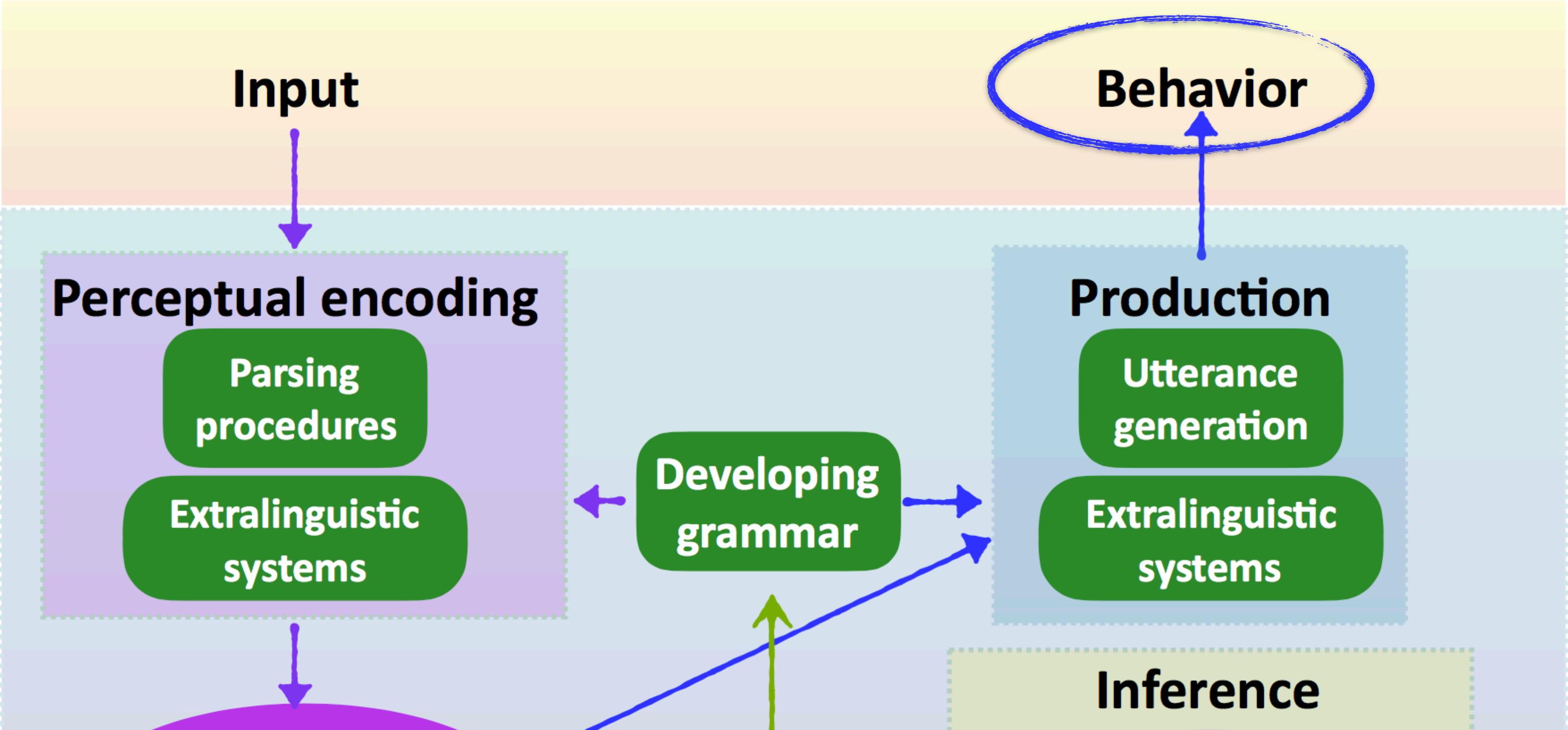
Who does 

Lily think the kitty for is pretty?

syntactic island

Who does Lily think the kitty for ___ is pretty?

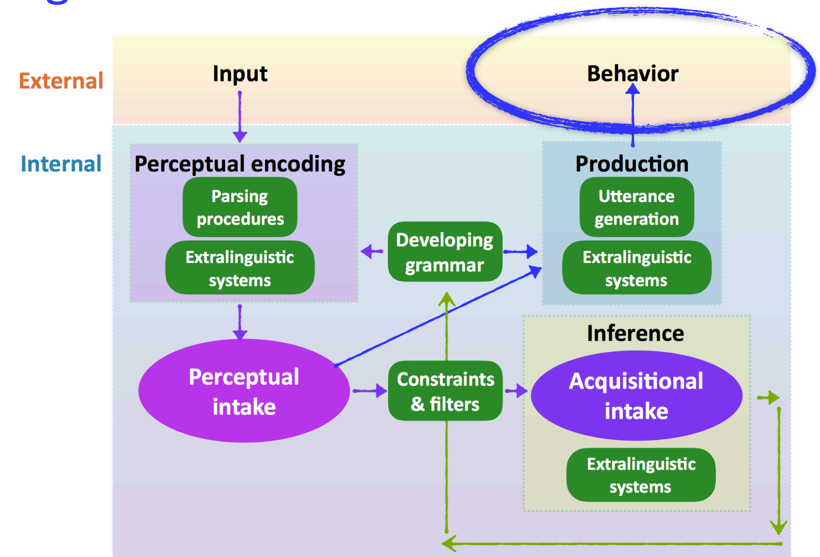
Adults judge these dependencies to be far worse than many others, including others that are very similar except that they don't cross syntactic islands (Sprouse et al. 2012).



Adult judgments: Target behavior

syntactic island

Adult knowledge as measured by acceptability judgment behavior



Sprouse et al. (2012) collected magnitude estimation judgments for four different islands, using a factorial definition that controlled for two salient properties of island-crossing dependencies:

- length of dependency (**matrix** vs. **embedded**)
- presence of an island structure (**non-island** vs. **island**)

Note: matrix is another word for "main" when talking about clause structure

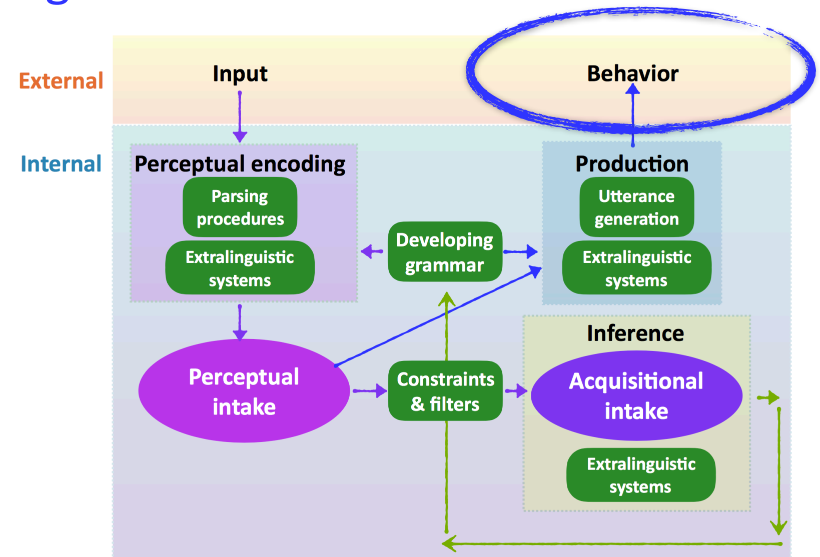
Pearl & Sprouse 2013a, 2013b, 2015

Adult judgments: Target behavior

syntactic island

Adult knowledge as measured by acceptability judgment behavior

Sprouse et al. (2012)
length of dependency
(**matrix** vs. **embedded**)
presence of an **island** structure
(**non-island** vs. **island**)



Complex NP island stimuli

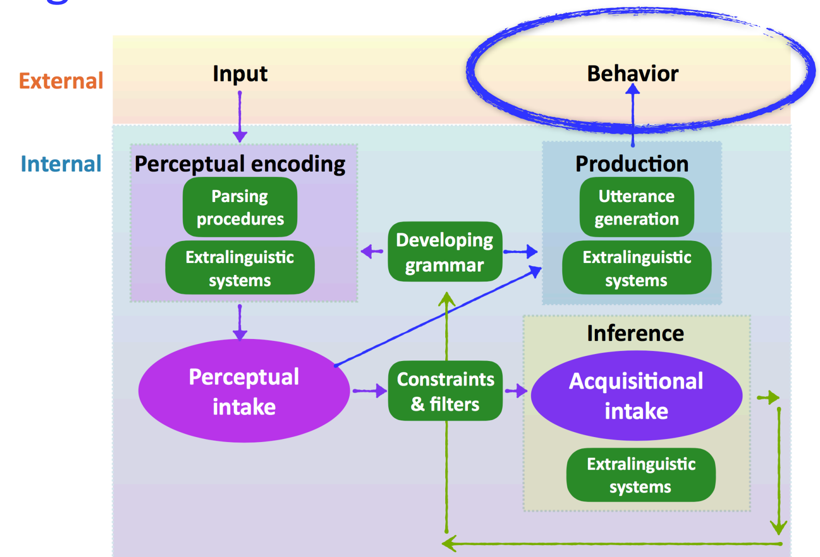
Who __ claimed that Lily forgot the necklace?	matrix | non-island
What did the teacher claim that Lily forgot __?	embedded | non-island
Who __ made the claim that Lily forgot the necklace?	matrix | island
*What did the teacher make the claim that Lily forgot __?	embedded | island

Pearl & Sprouse 2013a, 2013b, 2015

Adult judgments: Target behavior

Adult knowledge as measured by acceptability judgment behavior

Sprouse et al. (2012)
length of dependency
(**matrix** vs. **embedded**)
presence of an **island** structure
(**non-island** vs. **island**)



Subject island stimuli

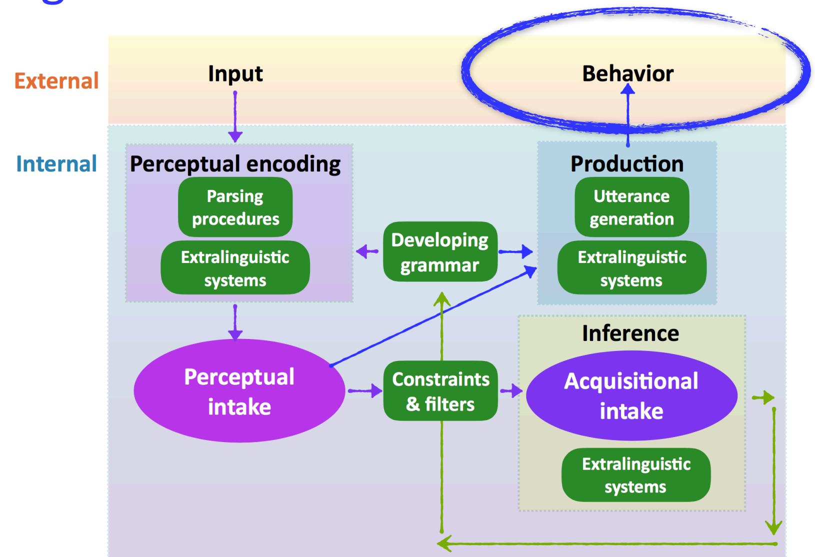
Who __ thinks the necklace is expensive?	matrix | non-island
What does Jack think __ is expensive?	embedded | non-island
Who __ thinks <u>the necklace for Lily</u> is expensive?	matrix | island
*Who does Jack think <u>the necklace for __</u> is expensive?	embedded | island

Pearl & Sprouse 2013a, 2013b, 2015

Adult judgments: Target behavior

Adult knowledge as measured by acceptability judgment behavior

Sprouse et al. (2012)
length of dependency
(**matrix** vs. **embedded**)
presence of an **island** structure
(**non-island** vs. **island**)



Whether island stimuli

Who __ thinks that Jack stole the necklace?	matrix | non-island
What does the teacher think that Jack stole __ ?	embedded | non-island
Who __ wonders whether Jack stole the necklace?	matrix | island
*What does the teacher wonder whether Jack stole __ ?	embedded | island

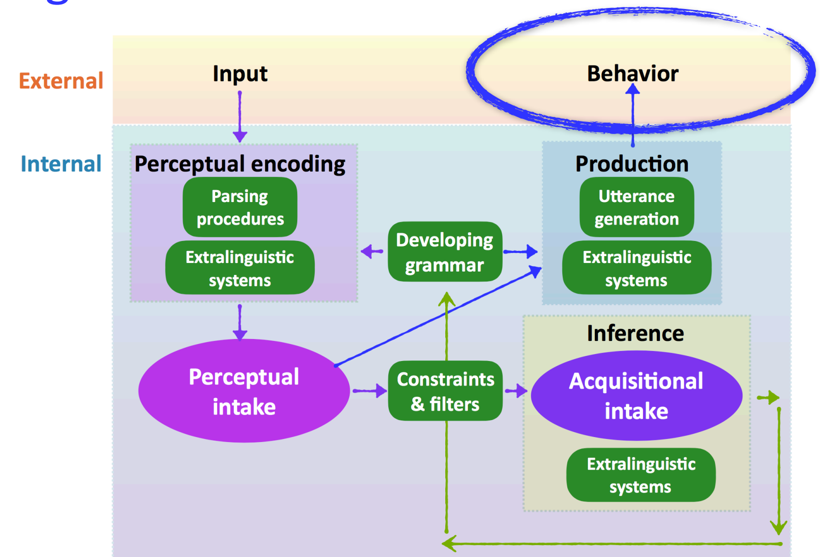
Pearl & Sprouse 2013a, 2013b, 2015

Adult judgments: Target behavior

syntactic island

Adult knowledge as measured by acceptability judgment behavior

Sprouse et al. (2012)
length of dependency
(**matrix** vs. **embedded**)
presence of an island structure
(**non-island** vs. **island**)



Adjunct island stimuli

Who __ thinks that Lily forgot the necklace?	matrix | non-island
What does the teacher think that Lily forgot __ ?	embedded | non-island
Who __ worries <u>if Lily forgot the necklace</u>?	matrix | island
*What does the teacher worry <u>if Lily forgot __</u> ?	embedded | island

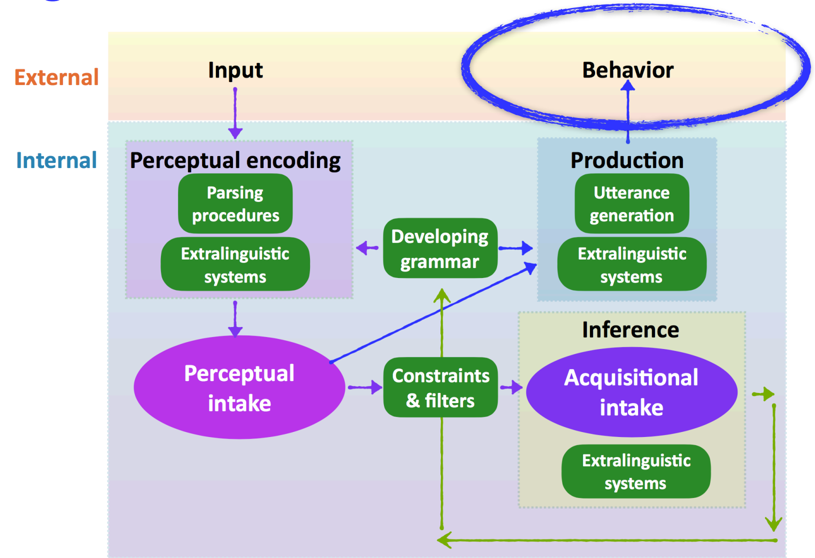
Pearl & Sprouse 2013a, 2013b, 2015

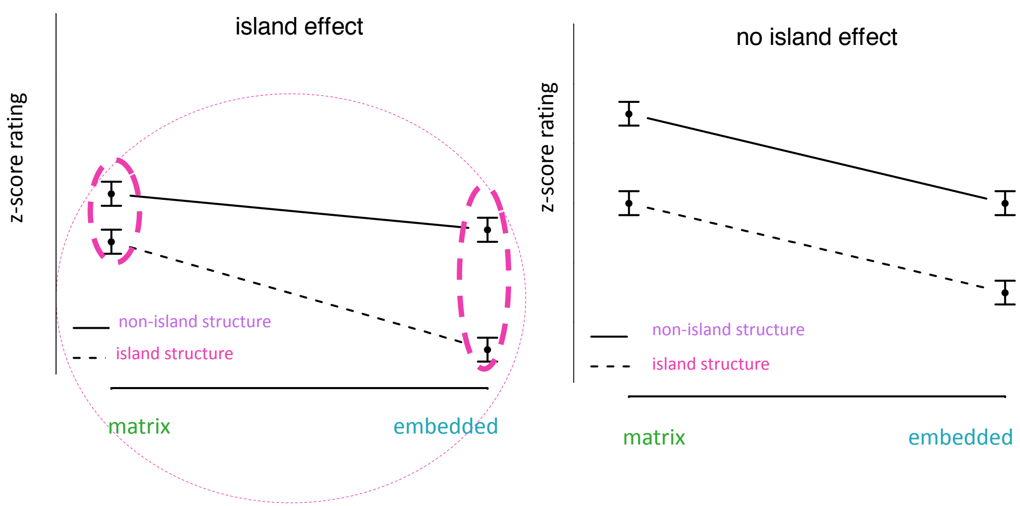
Adult judgments: Target behavior

syntactic island

Adult knowledge as measured by acceptability judgment behavior

Syntactic island = **superadditive** interaction of the two factors (additional unacceptability that arises when the two factors — **length** & presence of an **island** structure — are combined, above and beyond the independent contribution of each factor).





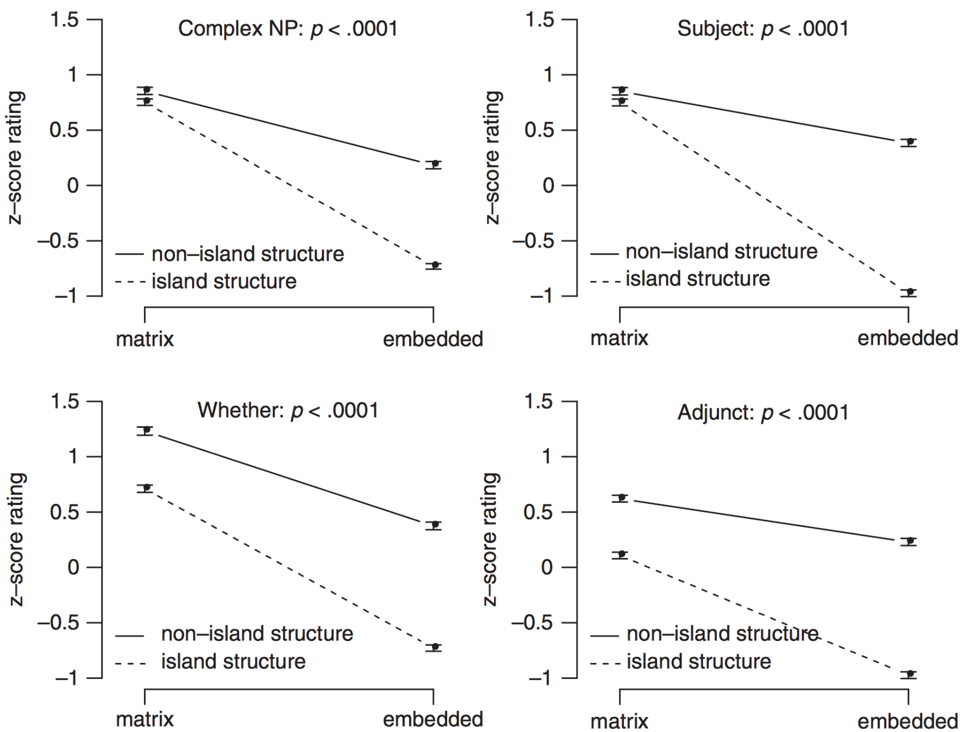
Pearl & Sprouse 2013a, 2013b, 2015

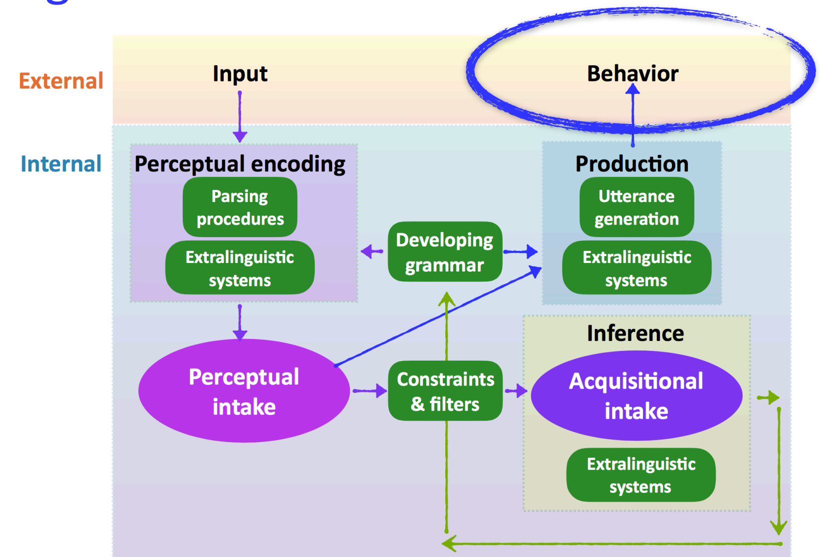
Adult judgments: Target behavior

syntactic island

Adult knowledge as measured by acceptability judgment behavior

Sprouse et al. (2012): acceptability judgments from 173 adult subjects





Superadditivity present for all islands tested = Knowledge that dependencies can't cross these island structures is part of adult knowledge about syntactic islands.

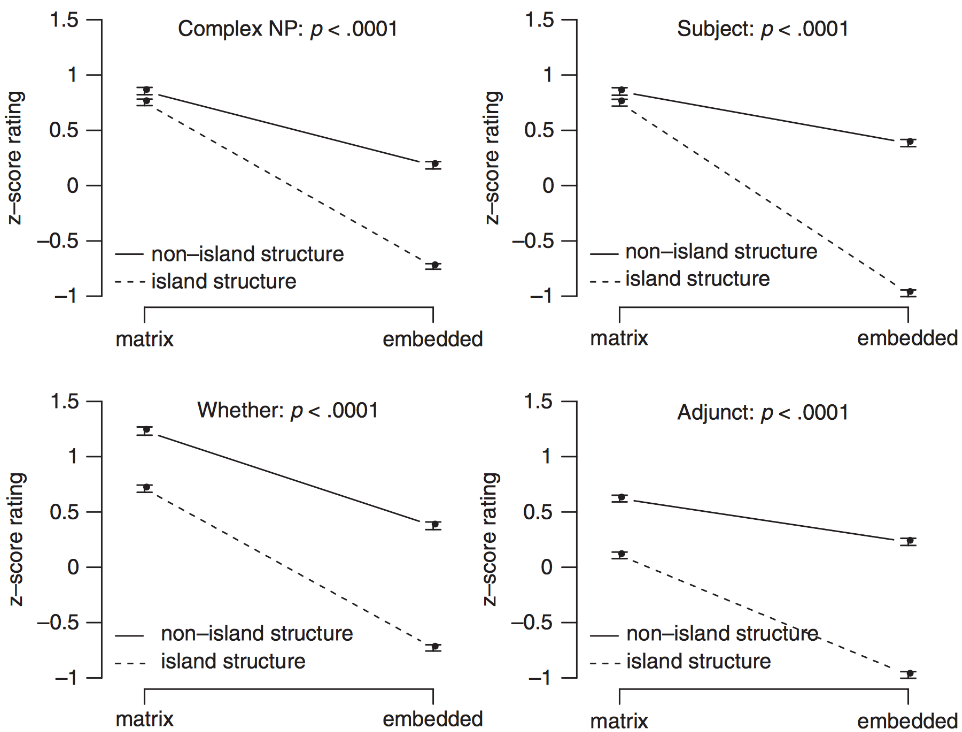
Pearl & Sprouse 2013a, 2013b, 2015

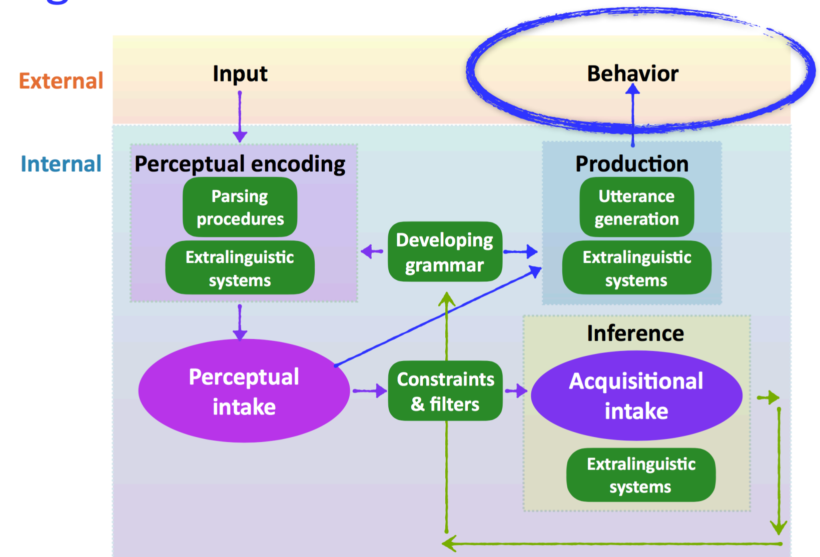
Adult judgments: Target behavior

syntactic island

Adult knowledge as measured by acceptability judgment behavior

Sprouse et al. (2012): acceptability judgments from 173 adult subjects





Importance for acquisition: This is one kind of target behavior that we'd like a modeled child to produce.

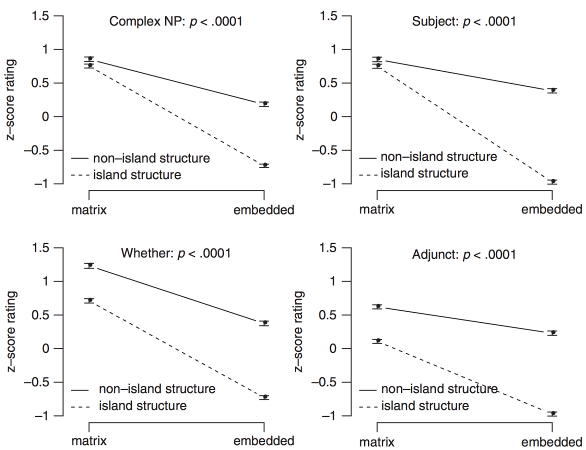
Pearl & Sprouse 2013a, 2013b, 2015

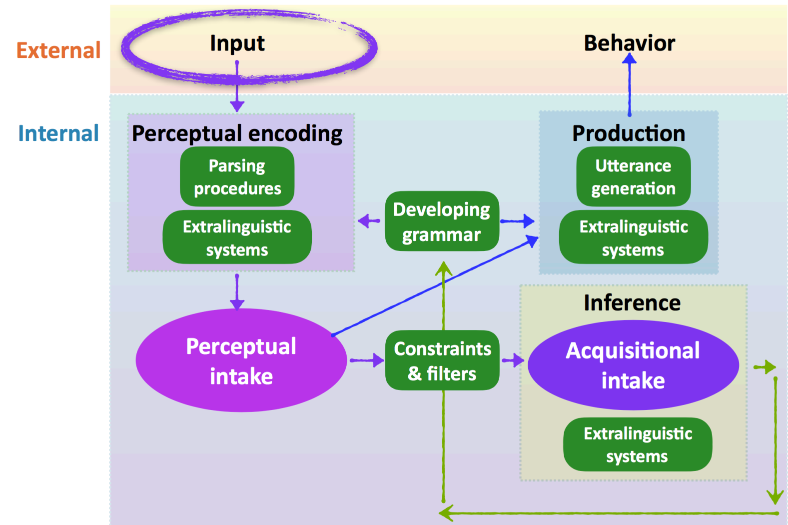
Adult judgments: Target behavior

syntactic island

Adult knowledge as measured by acceptability judgment behavior

Sprouse et al. (2012): acceptability judgments from 173 adult subjects





So if we're focusing on these *wh*-dependencies and that specific target state, what does children's input look like?

Pearl & Sprouse 2013a, 2013b, 2015

Children's input

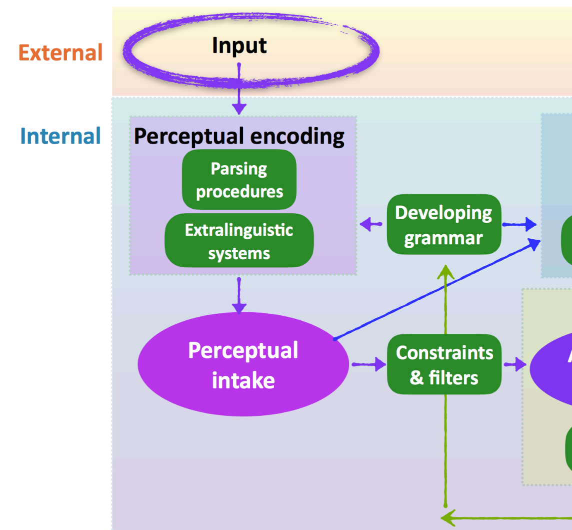
syntactic island

Children's input really doesn't look so helpful

Data from five corpora of child-directed speech (Brown-Adam, Brown-Eve, Brown-Sarah, Suppes, Valian) from CHILDES (MacWhinney 2000): speech to 25 children between the ages of one and five years old.

= 813,036 words

= 31,247 utterances containing a *wh*-dependency



Pearl & Sprouse 2013a, 2013b, 2015

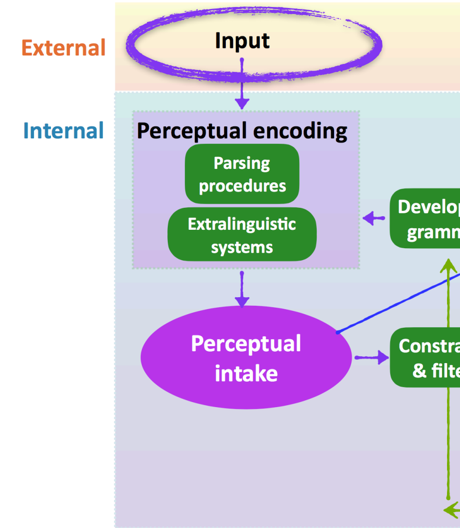
Children's input

syntactic island

Children's input really doesn't look so helpful

Data from five corpora of child-directed speech =
31,247 utterances containing a *wh*-dependency

		grammatical stimuli		*syntactic island*
	MATRIX + NON-ISLAND	EMBEDDED + NON-ISLAND	MATRIX + ISLAND	*EMBEDDED + ISLAND*
Complex NP	7	295	0	0
Subject	7	29	0	0
Whether	7	295	0	0
Adjunct	7	295	15	0



These kinds of utterances are fairly rare in general - the most frequent appears about 0.9% of the time (295 of 31,247.)

Pearl & Sprouse 2013a, 2013b, 2015

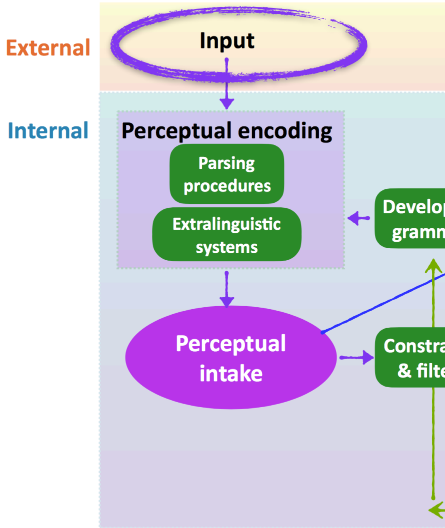
Children's input

syntax

syntactic island

Children's input really doesn't look so helpful

Data from five corpora of child-directed speech =
31,247 utterances containing a *wh*-dependency



grammatical stimuli *syntactic island*

	MATRIX + NON-ISLAND	EMBEDDED + NON-ISLAND	MATRIX + ISLAND	*EMBEDDED + ISLAND*
Complex NP	7	295	0	0
Subject	7	29	0	0
Whether	7	295	0	0
Adjunct	7	295	15	0

Being grammatical doesn't necessarily mean an utterance will appear in the input at all.

Pearl & Sprouse 2013a, 2013b, 2015

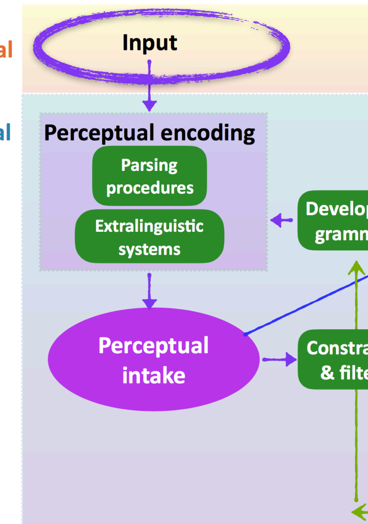
Children's input

syntax

syntactic island

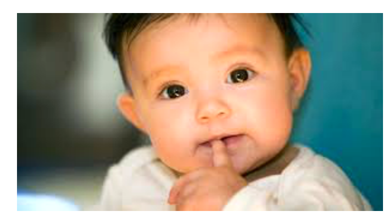
Children's input really doesn't look so helpful

Data from five corpora of child-directed speech =
31,247 utterances containing a *wh*-dependency



grammatical stimuli | *syntactic island*

	MATRIX + NON-ISLAND	EMBEDDED + NON-ISLAND	MATRIX + ISLAND	*EMBEDDED + ISLAND*
Complex NP	7	295	0	0
Subject	7	29	0	0
Whether	7	295	0	0
Adjunct	7	295	15	0



Unless the child is sensitive to very small frequencies, it's difficult to tell the difference between grammatical and ungrammatical dependencies sometimes…

Pearl & Sprouse 2013a, 2013b, 2015

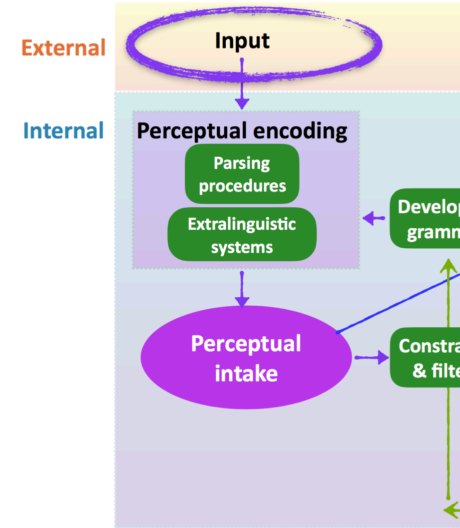
Children's input

syntax

syntactic island

Children's input really doesn't look so helpful

Data from five corpora of child-directed speech =
31,247 utterances containing a *wh*-dependency



grammatical stimuli *syntactic island*

	MATRIX + NON-ISLAND	EMBEDDED + NON-ISLAND	MATRIX + ISLAND	*EMBEDDED + ISLAND*
Complex NP	7	295	0	0
Subject	7	29	0	0
Whether	7	295	0	0
Adjunct	7	295	15	0

…and impossible to tell no matter what the rest of the time.
This looks like an **induction problem** for the language learner if we're looking for direct evidence in the input.

Pearl & Sprouse 2013a, 2013b, 2015

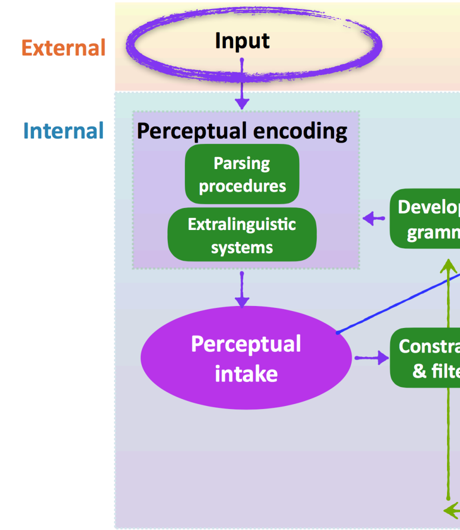
Children's input

syntax

syntactic island

Children's input really doesn't look so helpful

Data from five corpora of child-directed speech =
31,247 utterances containing a *wh*-dependency



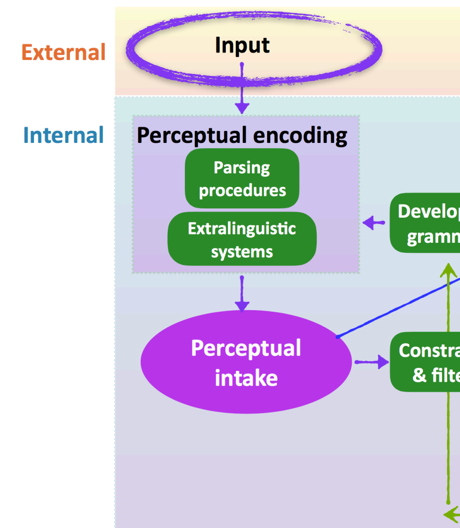
Important: Some grammatical utterances never appeared at all. This means that **only a subset of grammatical utterances appeared**, and the child has to **generalize appropriately from this subset**.

Pearl & Sprouse 2013a, 2013b, 2015

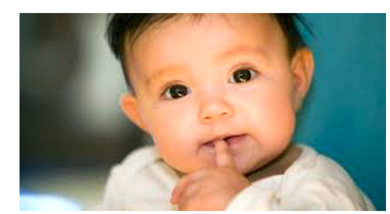
Children's input

syntactic island

Data from five corpora of child-directed speech =
31,247 utterances containing a *wh*-dependency



So what kinds of dependencies *are* in the input?



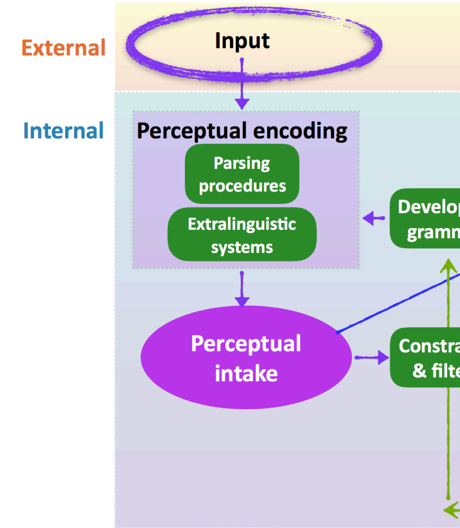
Pearl & Sprouse 2013a, 2013b, 2015

Children's input

syntactic island

So what kinds of dependencies *are* in the input?

Data from five corpora of child-directed speech = **31,247** utterances containing a *wh*-dependency



A lot of simpler ones!

76.7%	*What did you see __?*
12.8%	*What __ happened?*
5.6%	*What did she want to do __?*
2.5%	*What did she read from __?*
1.1%	*What did she think he said __?*

…

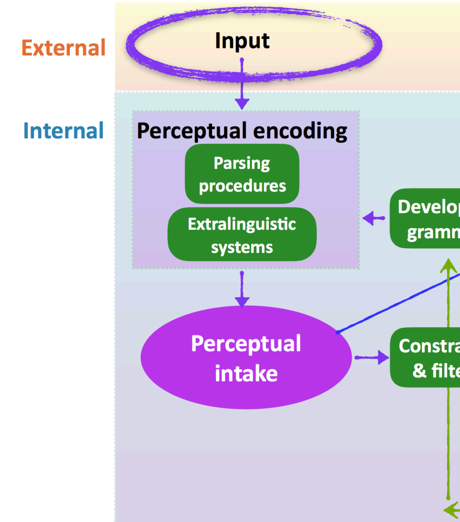
Pearl & Sprouse 2013a, 2013b, 2015

Children's input

syntax

syntactic island

The induction problem



Items Encountered

wh-questions in input (usually fairly simple)

What did you see __?

What __ happened?

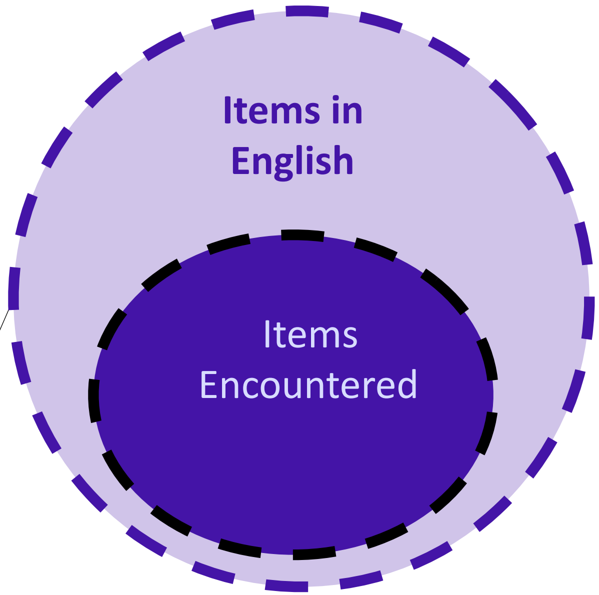
…

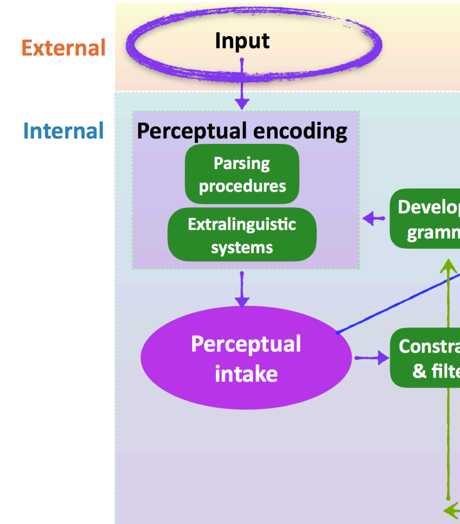
Pearl & Sprouse 2013a, 2013b, 2015

Children's input

syntactic island

The induction problem





Grammatical *wh*-questions

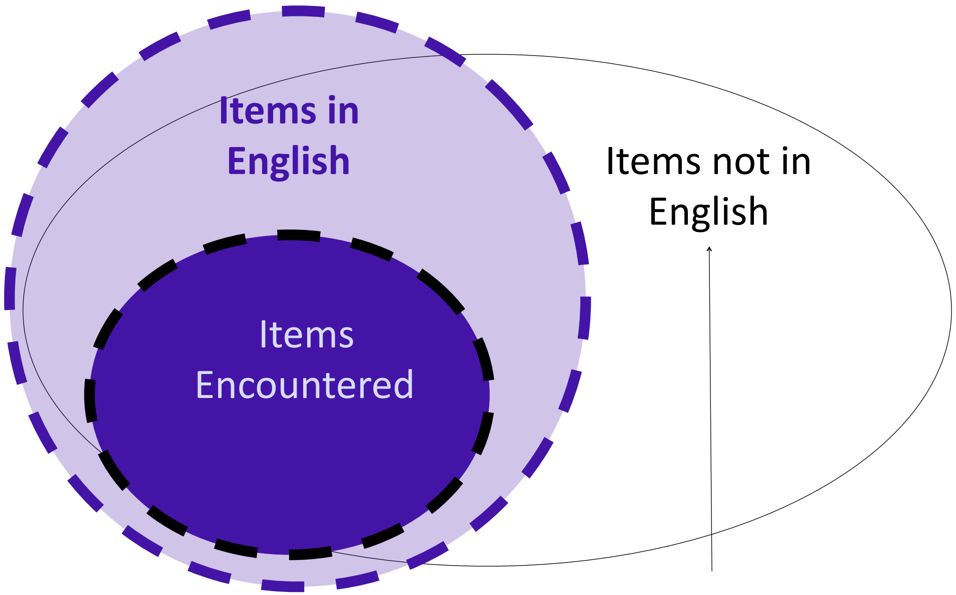
What did you see __?
What __ happened?
Who did Jack think that Lily saw __?
What did Jack think __ happened?

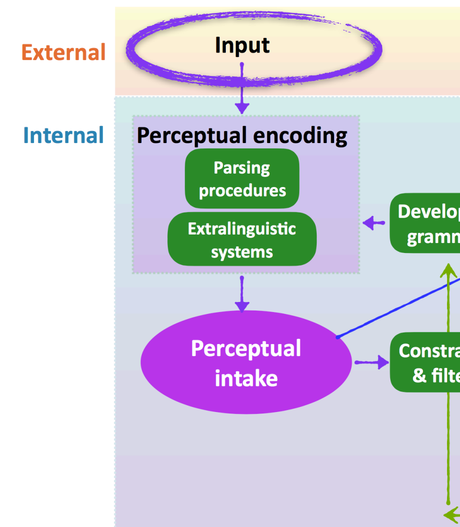
Pearl & Sprouse 2013a, 2013b, 2015

Children's input

syntactic island

The induction problem





Ungrammatical *wh*-questions: Syntactic islands

Who does Lily think the kitty for ___ is pretty?

What did Jack make the claim that he bought ___ ?

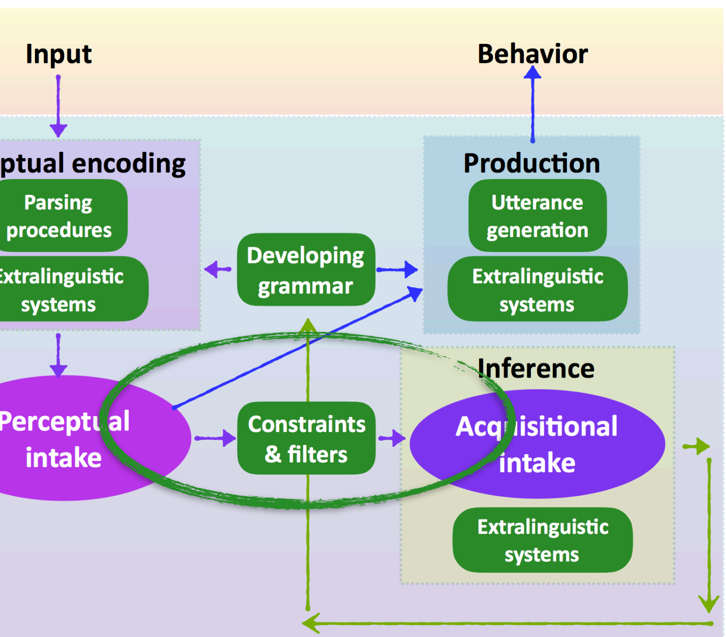
What did Elizabeth wonder whether Jack bought ___ ?

What did Elizabeth worry if Jack bought ___ ?

Pearl & Sprouse 2013a, 2013b, 2015

Learning strategies

Previous learning theories suggested children need syntactic-island-specific innate knowledge.



Learning strategies

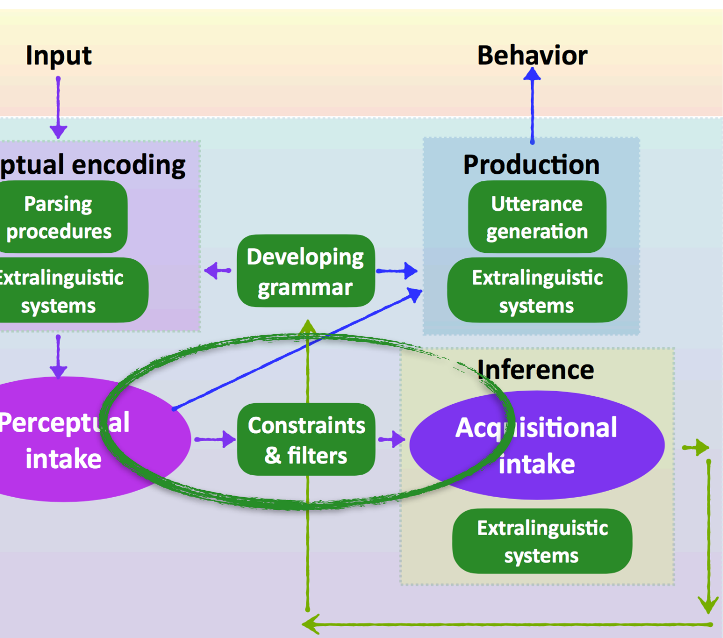
syntax

syntactic island

Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

A dependency cannot cross two or more bounding nodes.

Wh … $[_{BN1}$ … $[_{BN2}$ … ___]]



Learning strategies

syntax

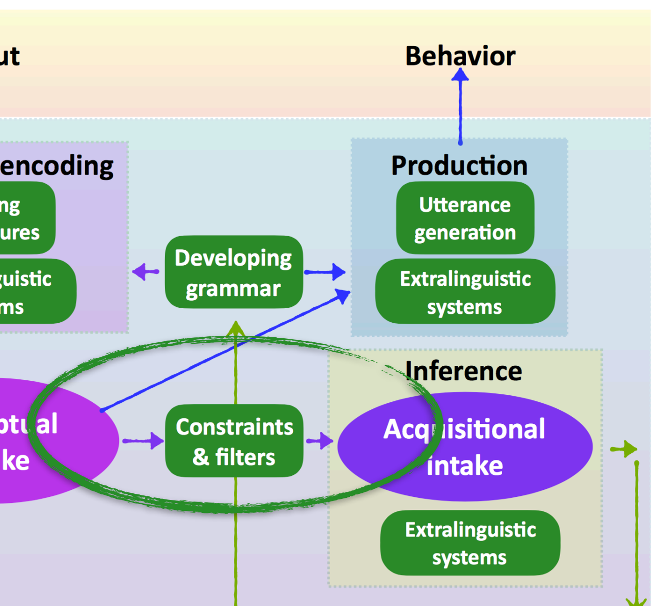
syntactic island

Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

Bounding nodes come from a fixed set of phrase structure nodes (CP, IP, and/or NP). The ones that act as a bounding nodes for a given language must be learned.

Wh … $[_{BN1}$ … $[_{BN2}$ … ___]]

{CP, IP, NP}?



Learning strategies

syntax

syntactic island

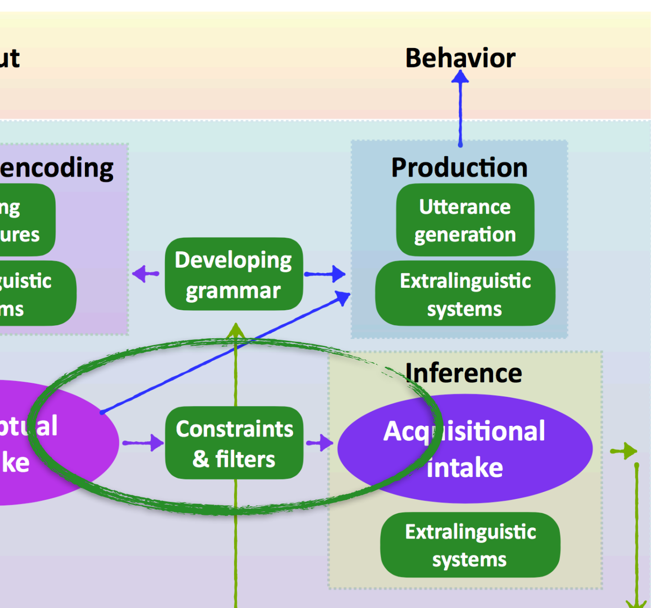
Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

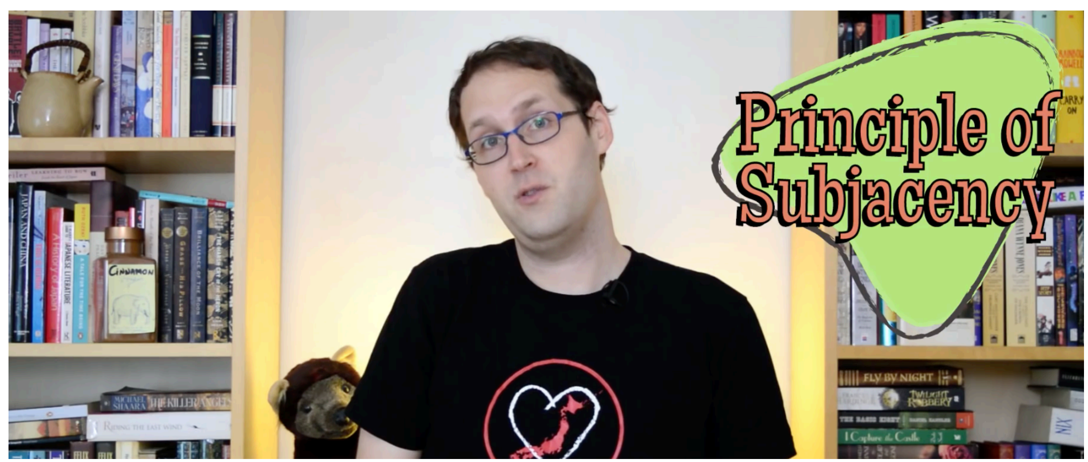
can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

Wh … $[_{BN1}$ … $[_{BN2}$ … __]]

http://www.thelingspace.com/episode-66
https://www.youtube.com/watch?v=01uH4XfJx3g

1:34 - 4:20





Learning strategies

syntax

syntactic island

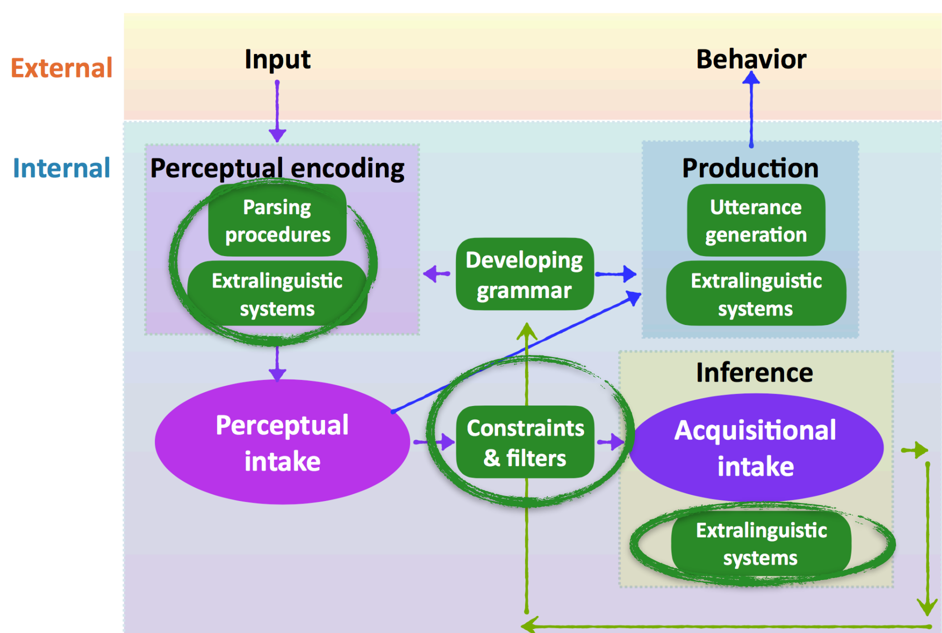
Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

$\mathit{Wh} \ \ldots \ [_{BN1} \ \ldots \ [_{BN2} \ \ldots \ \_\_]]$

An alternative learning strategy proposes children need less-specific linguistic prior knowledge along with probabilistic learning.

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)



Learning strategies

syntax

syntactic island

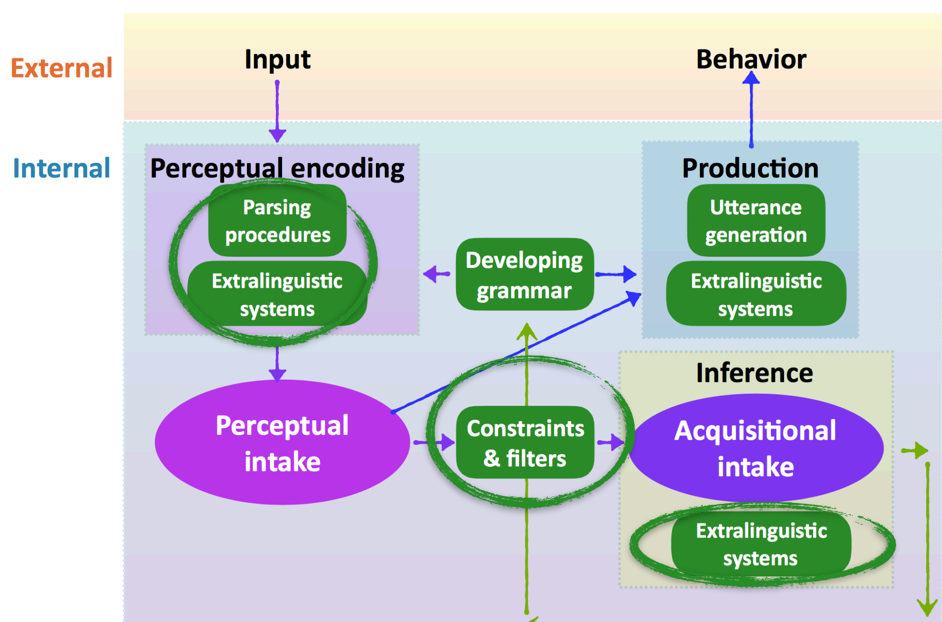
Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

$\textit{Wh} \ \ldots \ [_{BN1} \ \ldots \ [_{BN2} \ \ldots \ \_\_]]$

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure



Learning strategies


syntax

syntactic island

Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

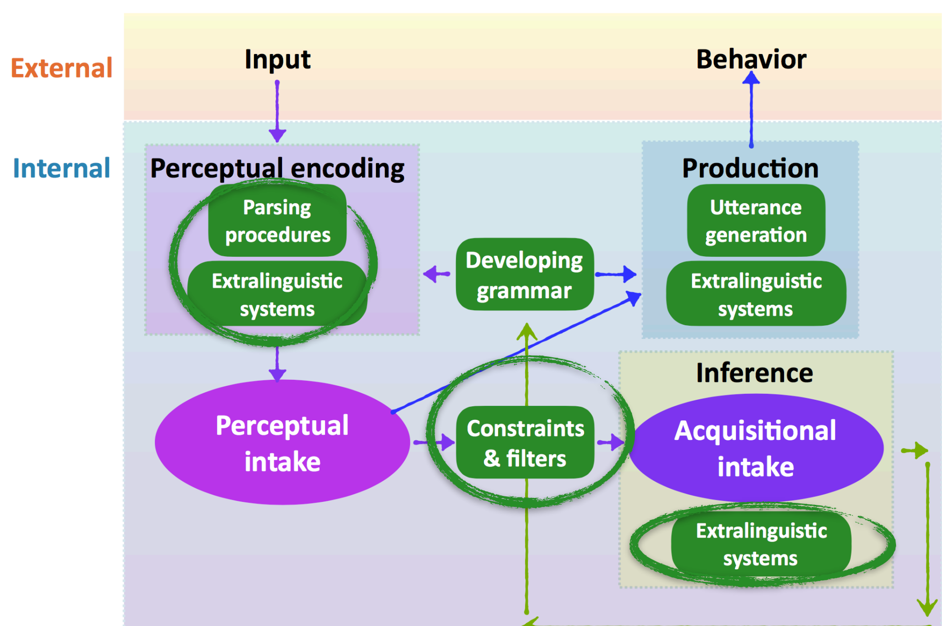
can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

Wh … $[_{BN1}$ … $[_{BN2}$ … ___]]

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

Dependencies represented as a sequence of **container nodes**



Container nodes

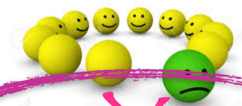
syntax

syntactic island

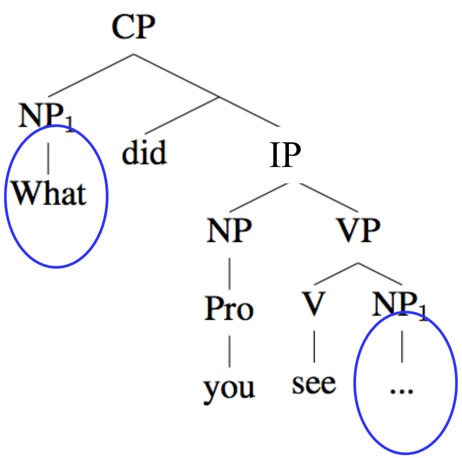
Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

Dependencies represented as a sequence of **container nodes**



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … ___]]



How to describe this dependency:
What phrases is the gap inside but the *wh*-word isn't inside?

Container nodes


syntax

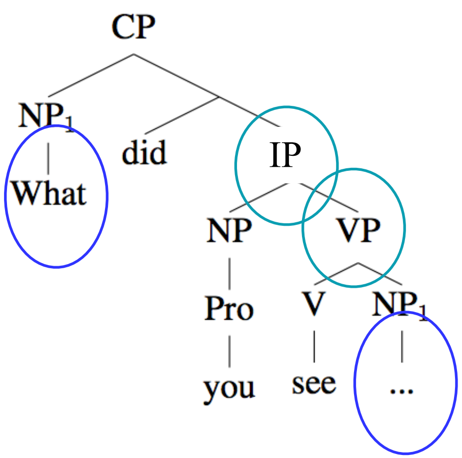
syntactic island

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

Dependencies represented as a sequence of **container nodes**

Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … __]]



How to describe this dependency:
What phrases is the gap inside but the *wh*-word isn't inside?

What did you see __?
= What did [$_{IP}$ you [$_{VP}$ see __]]?
= IP-VP

Container nodes

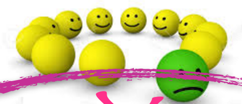

syntax

syntactic island

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

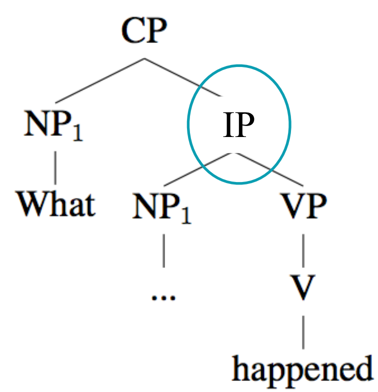
Dependencies represented as a sequence of **container nodes**



Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … __]]

What did you see __?
= What did [$_{IP}$ you [$_{VP}$ see __]]?
= IP-VP

What __ happened?
= What [$_{IP}$ __ happened]?
= IP



Container nodes

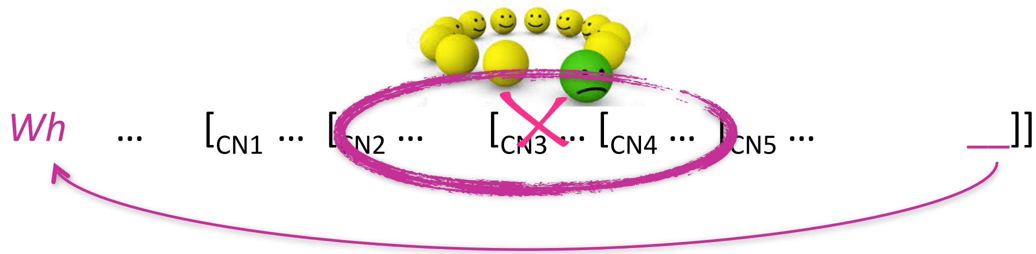

syntax

syntactic island

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

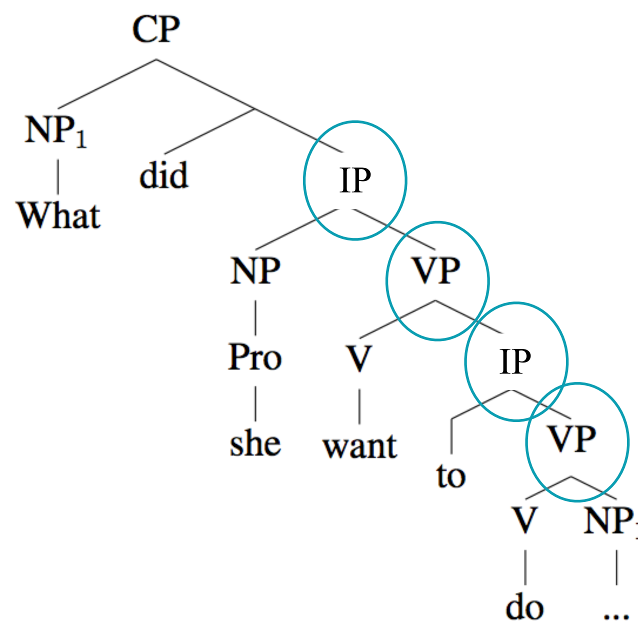
Dependencies represented as a sequence of **container nodes**



What did you see __?
= What did [$_{IP}$ you [$_{VP}$ see __]]?
= IP-VP

What __ happened?
= What [$_{IP}$ __ happened]?
= IP

What did she want to do __ ?
= What did [$_{IP}$ she [$_{VP}$ want [$_{IP}$ to [$_{VP}$ do __]]]]?
= IP-VP-IP-VP



Container nodes

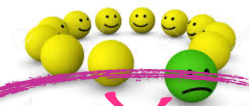

syntax

syntactic island

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

Dependencies represented as a sequence of **container nodes**



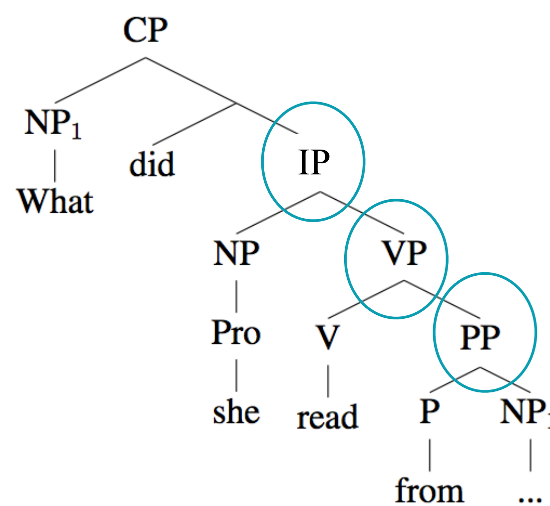
Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … __]]

What did you see __?
= What did [$_{IP}$ you [$_{VP}$ see __]]?
= IP-VP

What __ happened?
= What [$_{IP}$ __ happened]?
= IP

What did she want to do __ ?
= What did [$_{IP}$ she [$_{VP}$ want [$_{IP}$ to [$_{VP}$ do __]]]]?
= IP-VP-IP-VP

What did she read from __ ?
= What did [$_{IP}$ she [$_{VP}$ read [$_{PP}$ from __]]]]?
= IP-VP-PP



Learning strategies

syntax

syntactic island

Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

Wh … $[_{BN1}$ … $[_{BN2}$ … ___]]

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

Dependencies represented as a sequence of **container nodes**

Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … ___]]

Container node: phrase structure node that contains dependency

$[_{CP}$ *What* *do* $[_{IP}$ *you* $[_{VP}$ *like* ___ $[_{PP}$ *in this picture?*]]]]

Learning strategies

syntax

syntactic island

Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

Wh … [$_{BN1}$ … [$_{BN2}$ … __]]

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

Dependencies represented as a sequence of **container nodes**

Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … __]]

Sequence of container nodes characterizes dependencies

[$_{CP}$ *What* *do* [$_{IP}$ *you* [$_{VP}$ *like* __ [$_{PP}$ *in this picture?*]]]]

start-IP-VP-*end*

Learning strategies

syntax

syntactic island

Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

Wh … $[_{BN1}$ … $[_{BN2}$ … __]]

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

Dependencies represented as a sequence of **container nodes**

Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

Ungrammatical dependencies have low probability segments

$[_{CP}$ *Who* *did* $[_{IP}$ *Lily* $[_{VP}$ *think* $[_{CP}$ $[_{IP}$ $[_{NP}$ *the kitty* $[_{PP}$ *for* __]] *was pretty ?*]]]]

start-IP-VP-CP-IP-NP-PP-*end*

Learning strategies


syntax

syntactic island

Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

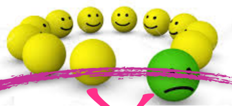
can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

Wh … $[_{BN1}$ … $[_{BN2}$ … ___]]

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability region of structure

Dependencies represented as a sequence of **container nodes**



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … ___]]

Low probability container node sequences have to be learned for the language

Learning strategies

syntax

syntactic island

Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

Wh … $[_{\text{BN1}}$ … $[_{\text{BN2}}$ … ___]]

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very
low probability sequence of
container nodes

Wh … $[_{\text{CN1}}$ … $[_{\text{CN2}}$ … $[_{\text{CN3}}$ … $[_{\text{CN4}}$ … $[_{\text{CN5}}$ … ___]]

In common: Local structural anomaly is the problem

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

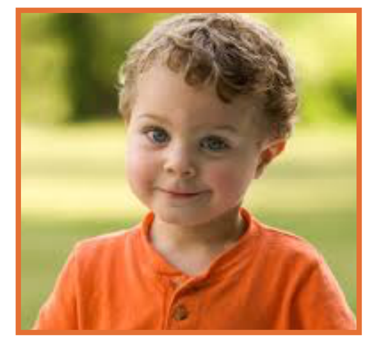
syntax

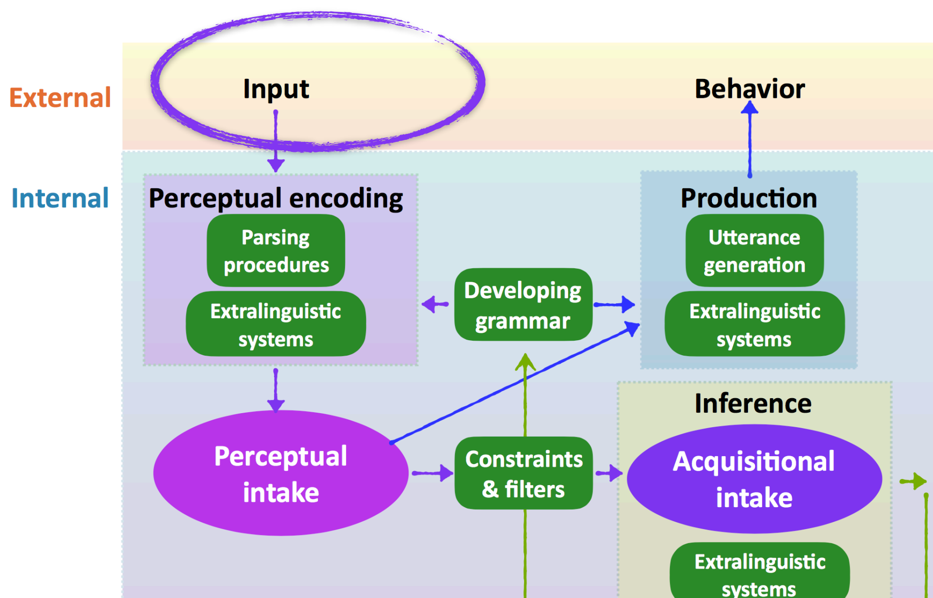
syntactic island

A dependency can't cross a very low probability sequence of container nodes

Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

Implemented in an algorithmic-level learning model that learned from realistic samples of child-directed speech.



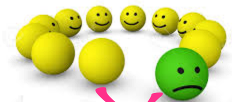


Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)


syntax

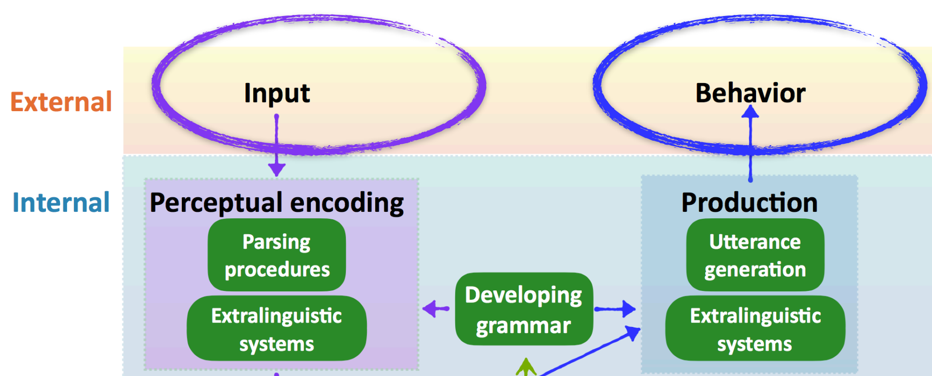
syntactic island

A dependency can't cross a very low probability sequence of container nodes



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

Intuition: Learn what you can from the dependencies you do actually observe in the data and apply it to make a judgment about the dependencies you haven't seen before, like these syntactic islands.

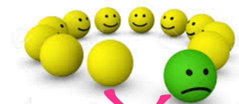


Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island

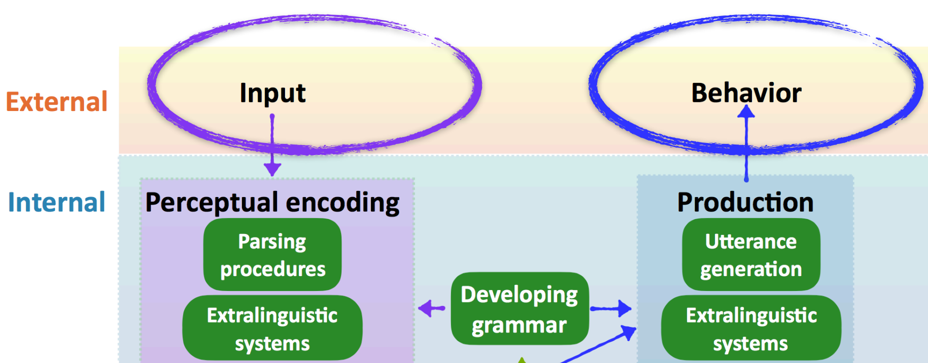
A dependency can't cross a very low probability sequence of container nodes



Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … ___]]

Intuition: Learn what you can from the dependencies you do actually observe in the data and apply it to make a judgment about the dependencies you haven't seen before, like these syntactic islands.

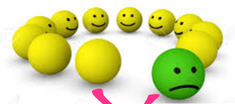
That is, leverage a broader set of data to make syntactic generalizations.



Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

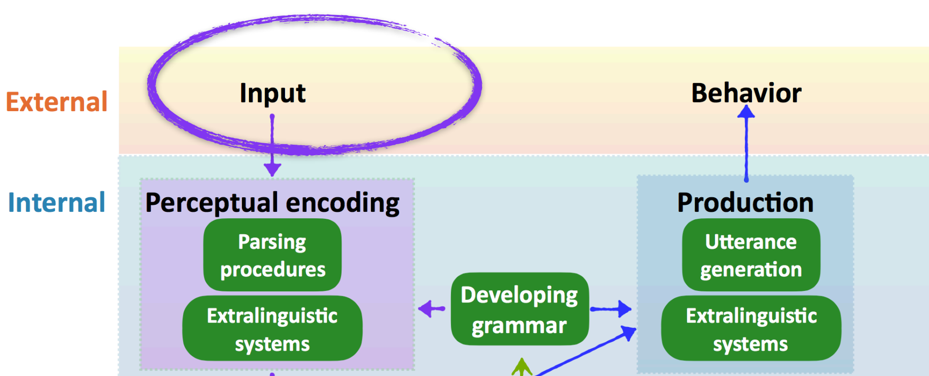

syntax

syntactic island



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … ___]]

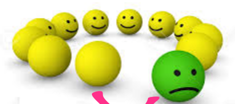
What information is there to leverage exactly?



Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)


syntax

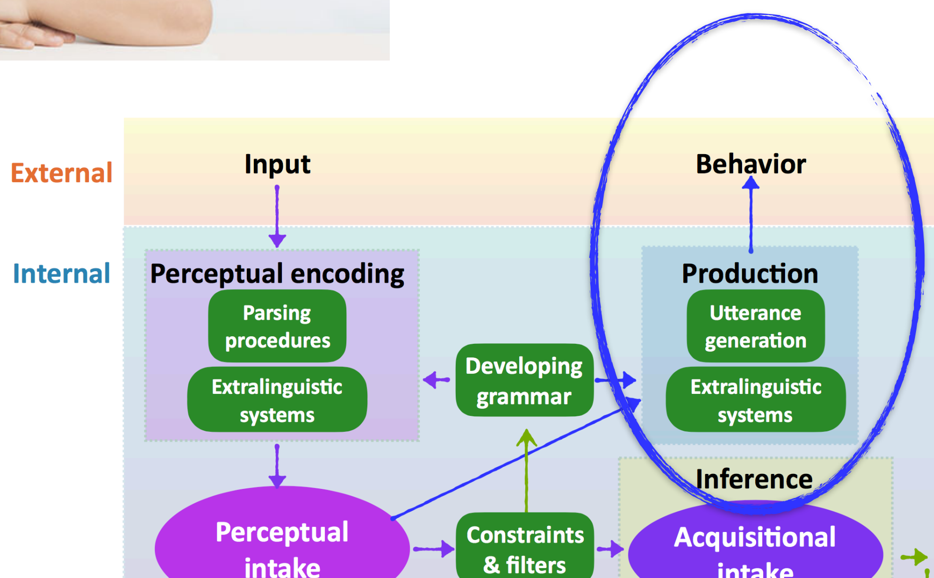
syntactic island



Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … __]]

What information is there to leverage exactly?

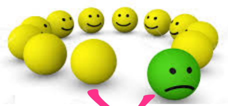
This relates to the strategy children use for learning and then generating predictions about the grammaticality of dependencies.



Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



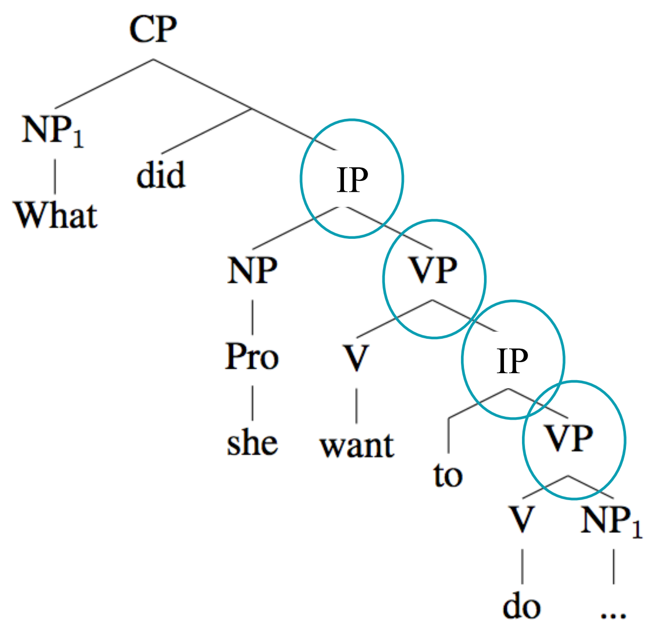
Wh … [$_{CN1}$ … [$_{CN2}$ … ~~[$_{CN3}$~~ … [$_{CN4}$ … [$_{CN5}$ … __]]

What information is there to leverage exactly?

Strategy

(1) Pay attention to the structure of dependencies.

What did she want to do __ ?
= What did [$_{IP}$ she [$_{VP}$ want [$_{IP}$ to [$_{VP}$ do __]]]]?
= IP-VP-IP-VP

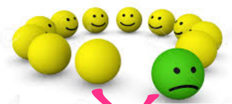


Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

What information is there to leverage exactly?

Strategy

(1) Pay attention to dependency structure.

(2) Break these dependency structures into smaller pieces made up of three units (trigrams) that you can track the frequency of in the input you encounter.

IP-VP =
begin-IP-VP
IP-VP-*end*

IP =
begin-IP-*end*

IP-VP-IP-VP
= *begin*-IP-VP
IP-VP-IP
VP-IP-VP
IP-VP-*end*

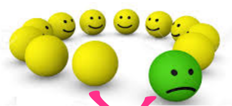
IP-VP-PP
= *begin*-IP-VP
IP-VP-PP
VP-PP-*end*

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … __]]

What information is there to leverage exactly?

Strategy

(1) Pay attention to dependency structure.

(2) Break these dependency structures into smaller pieces made up of three units (trigrams) that you can track the frequency of in the input you encounter.

IP-VP =
begin-IP-VP
IP-VP-*end*

IP =
begin-IP-*end*

IP-VP-IP-VP
= *begin*-IP-VP
IP-VP-IP
VP-IP-VP
IP-VP-*end*

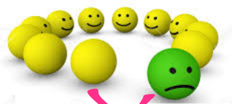
IP-VP-PP
= *begin*-IP-VP
IP-VP-PP
VP-PP-*end*

begin-IP-VP = 86/225
IP-VP-*end* = 83/225
begin-IP-*end* = 13/225
IP-VP-IP = 6/225
VP-IP-VP = 6/225
IP-VP-PP = 3/225
VP-PP-*end* = 3/225
…

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … ___]]

What information is there to leverage exactly?

Strategy

(1) Pay attention to dependency structure.

(2) Break these dependency structures into smaller pieces made up of three units (trigrams) that you can track the frequency of in the input you encounter.

IP-VP =
begin-IP-VP
IP-VP-*end*

IP =
begin-IP-*end*

IP-VP-IP-VP
= *begin*-IP-VP
IP-VP-IP
VP-IP-VP
IP-VP-*end*

IP-VP-PP
= *begin*-IP-VP
IP-VP-PP
VP-PP-*end*

begin-IP-VP = 86/225
IP-VP-*end* = 83/225
begin-IP-*end* = 13/225
IP-VP-IP = 6/225
VP-IP-VP = 6/225
IP-VP-PP = 3/225
VP-PP-*end* = 3/225
…

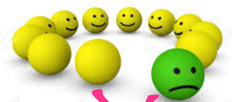
Note that some of these trigrams appear in multiple dependencies that commonly occur in children's input. This will be helpful!

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … ___]]

What information is there to leverage exactly?

Strategy

(1) Pay attention to dependency structure.

(2) Break dependency structures into trigrams that you can track the frequency of.

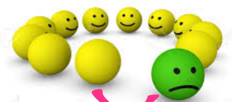
(3) Use trigram frequency to calculate the probability of that trigram occurring in a dependency.

begin-IP-VP = 86/225	p(*begin*-IP-VP) = 0.38
IP-VP-*end* = 83/225	p(IP-VP-*end*) = 0.37
begin-IP-*end* = 13/225	p(*begin*-IP-*end*) = 0.06
IP-VP-IP = 6/225	p(IP-VP-IP) = 0.03
VP-IP-VP = 6/225	p(VP-IP-VP) = 0.03
IP-VP-PP = 3/225	p(IP-VP-PP) = 0.01
VP-PP-*end* = 3/225	p(VP-PP-*end*) = 0.01
…	

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)


syntax

syntactic island



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

What information is there to leverage exactly?

Strategy

(1) Pay attention to dependency structure.

(2) Break dependency structures into trigrams that you can track the frequency of.

(3) Calculate the trigram probability in a dependency.

(4) When you see a new dependency, break it down into its trigrams and then calculate its probability, based on the trigram probabilities.

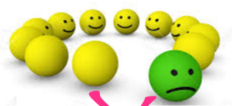
What does Jack want __?
= What does $[_{IP}$ Jack $[_{VP}$ want __]]?
= IP-VP
= *begin*-IP-VP
IP-VP-*end*

p(IP-VP) = p(*begin*-IP-VP)*p(IP-VP-*end*)
= 0.38 * 0.37 = 0.14

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)


syntax

syntactic island



Wh … [$_{CN1}$ … [$_{CN2}$ … ~~[$_{CN3}$~~ … [$_{CN4}$ … [$_{CN5}$ … __]]

What information is there to leverage exactly?

Strategy

(1) Pay attention to dependency structure.

(2) Break dependency structures into trigrams that you can track the frequency of.

(3) Calculate the trigram probability in a dependency.

(4) When you see a new dependency, break it down into its trigrams and then calculate its probability, based on the trigram probabilities.

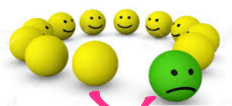
What does Jack want to do that for __?
= What does [$_{IP}$ Jack [$_{VP}$ want [$_{IP}$ to [$_{VP}$ do that [$_{PP}$ for __]]?
= IP-VP-IP-VP-PP
= *begin*-IP-VP
IP-VP-IP
VP-IP-VP
IP-VP-PP
VP-PP-*end*

p(IP-VP-IP-VP-PP) = p(*begin*-IP-VP)*p(IP-VP-IP)*p(VP-IP-VP)*p(IP-VP-PP)*p(VP-PP-*end*)
= 0.38*0.03*0.03*0.01*0.01 = 0.000000034

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … __]]

What information is there to leverage exactly?

Strategy

(1) Pay attention to dependency structure.

(2) Break dependency structures into trigrams that you can track the frequency of.

(3) Calculate the trigram probability in a dependency.

(4) When you see a new dependency, break it down into its trigrams and then calculate its probability, based on the trigram probabilities.

Subject island dependency

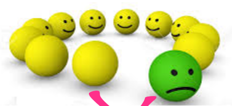
What do you think that the joke about __ offended Jack?
= What do [$_{IP}$ you [$_{VP}$ think [$_{CP}$ that [$_{IP}$ [$_{NP}$ the joke [$_{PP}$ about __]]]]]] offended Jack?
= IP-VP-CP-NP-PP
= *begin*-IP-VP
IP-VP-CP
VP-CP-IP
CP-IP-NP
IP-NP-PP
NP-PP-*end*

p(IP-VP-CP-IP-NP-PP) = p(*begin*-IP-VP)*p(IP-VP-CP)*p(VP-CP-S)*p(CP-IP-NP)*p(IP-NP-PP)*p(NP-PP-*end*)
= 0.86*0.01*0.001*0.00*0.00*0.02 = 0.00

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

What information is there to leverage exactly?

Strategy

(1) Pay attention to dependency structure.

(2) Break dependency structures into trigrams that you can track the frequency of.

(3) Calculate the trigram probability in a dependency.

(4) Break a new dependency into its trigrams and calculate its probability.

(5) Use calculated dependency probabilities as the basis for grammaticality judgments. Lower probability dependencies are dispreferred, compared to higher probability dependencies.

p(IP-VP) = 0.14

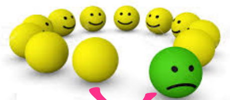
p(IP-VP-IP-VP-PP) = 0.000000034

p(IP-VP-CP-IP-NP-PP) = 0.00

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)


syntax

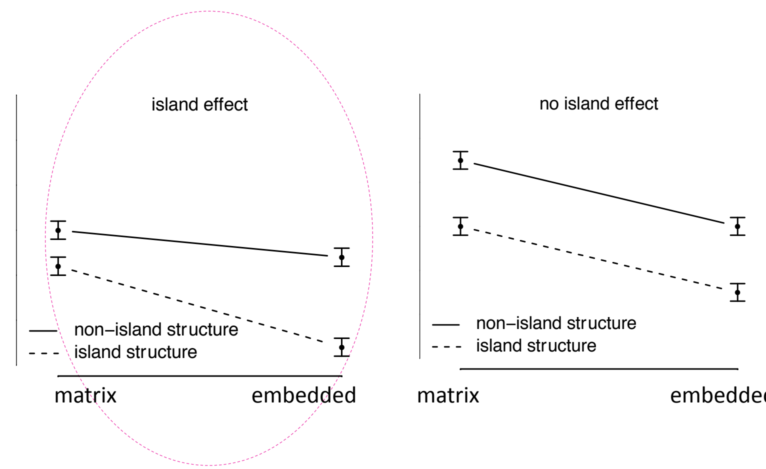
syntactic island



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

Use calculated dependency probabilities as the basis for grammaticality judgments. Lower probability dependencies are dispreferred, compared to higher probability dependencies.

For each set of island stimuli from Sprouse et al. (2012), we generate grammaticality preferences for the modeled learner based on the dependency's perceived probability and use this as a stand-in for acceptability.

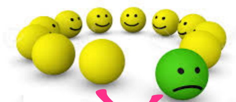


Looking for superadditivity as a sign of syntactic island knowledge

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

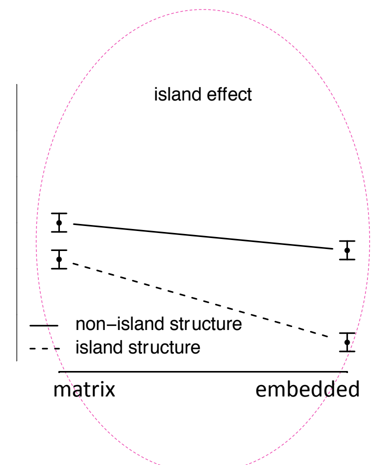
syntactic island

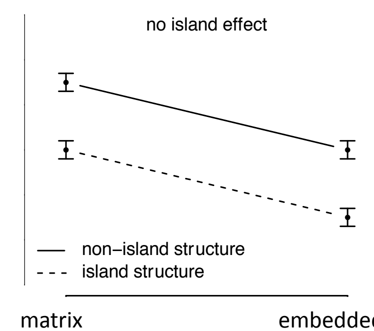


Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

Use calculated dependency probabilities as the basis for grammaticality judgments. Lower probability dependencies are dispreferred, compared to higher probability dependencies.

	matrix	embedded
non-island	*Who __ claimed that Lily forgot the necklace?*	*What did the teacher claim that Lily forgot __?*
island	*Who __ made the claim that Lily forgot the necklace?*	**What did the teacher make the claim that Lily forgot __?*



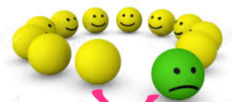


Looking for superadditivity as a sign of syntactic island knowledge

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

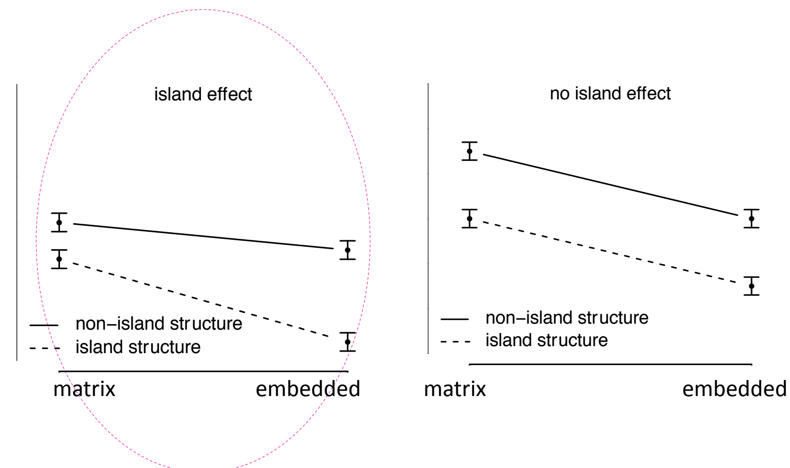
syntactic island



Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … __]]

Use calculated dependency probabilities as the basis for grammaticality judgments. Lower probability dependencies are dispreferred, compared to higher probability dependencies.

	matrix	embedded
non-island	*IP*	*IP-VP-CP$_{that}$-IP-VP*
island	*IP*	**IP-VP-NP-CP$_{that}$-IP-VP*

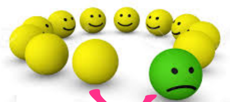


Each dependency is characterized by a container node sequence, whose probability can be calculated and then plotted.

✓ Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)


syntax

syntactic island

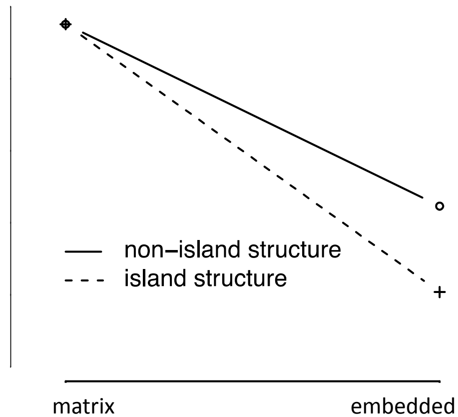


Wh … [$_{CN1}$ … [$_{CN2}$ … [$_{CN3}$ … [$_{CN4}$ … [$_{CN5}$ … __]]

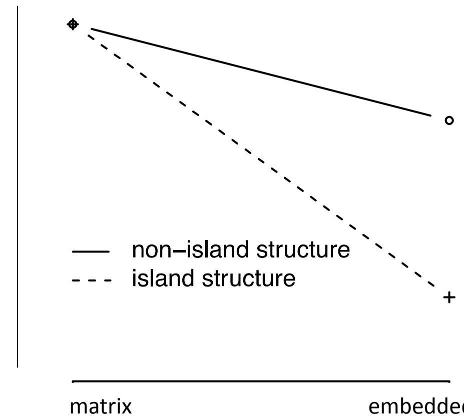
Superadditivity observed for all four islands — the qualitative behavior suggests that this learner has knowledge of these syntactic islands.

The Subjacency-ish representation that relies on container node trigram probabilities can solve this learning problem using this learning strategy.

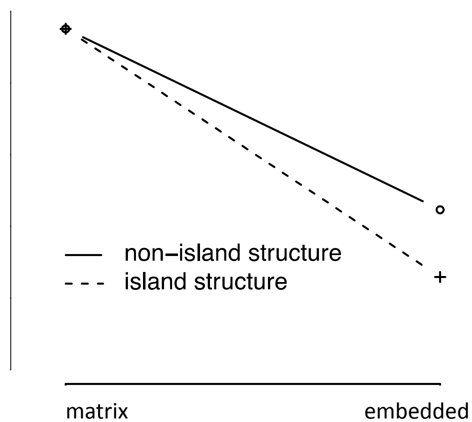
Complex NP



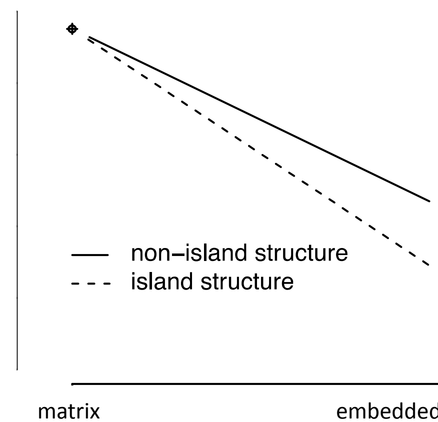
Subject



Whether



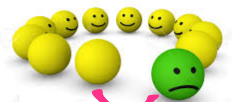
Adjunct



✓ Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

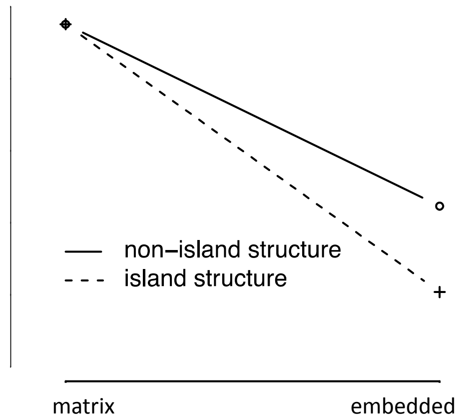
syntactic island



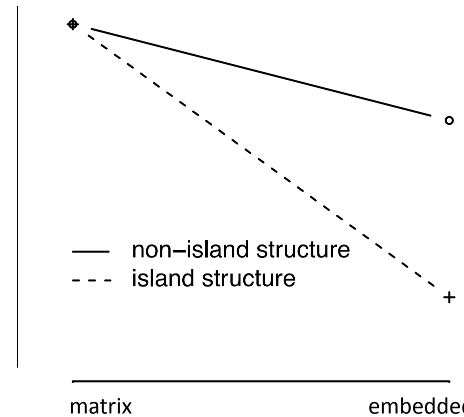
Wh … [$_{CN1}$ … [$_{CN2}$ … ~~[$_{CN3}$~~ … [$_{CN4}$ … [$_{CN5}$ … __]]

Note: We're careful to say "qualitative" behavior fit because there are lots of other factors that impact acceptability judgment behavior, and we've only modeled one (presumably) large part of them, which is the grammaticality of the dependency.

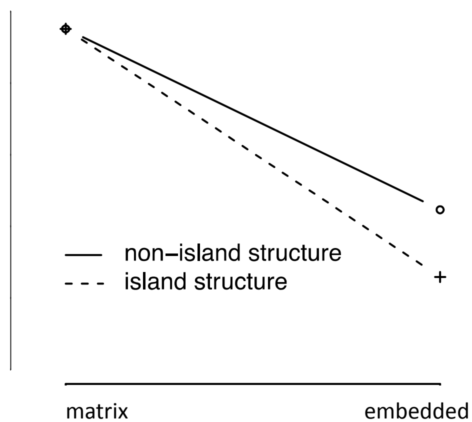
Complex NP



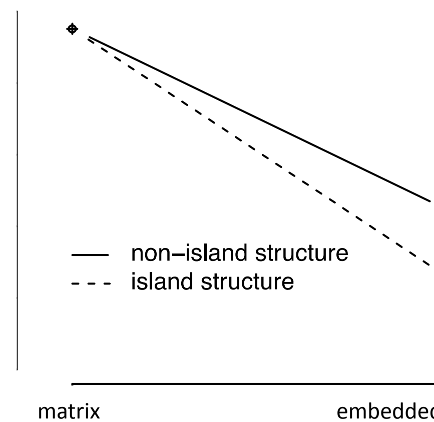
Subject



Whether



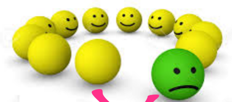
Adjunct



✓ Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)


syntax

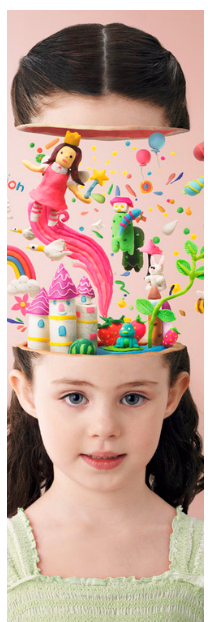
syntactic island



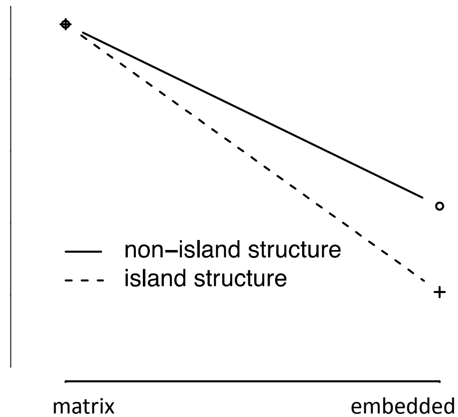
Wh … [$_{CN1}$ … [$_{CN2}$ … ~~[$_{CN3}$~~ … [$_{CN4}$ … [$_{CN5}$ … ___]]

But is this all we can say?

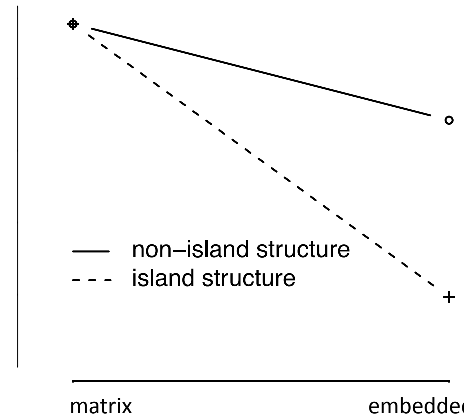
No! One useful aspect of models is that we can look inside the modeled child to see *why* it's behaving the way that it is. (This is something that's harder to do with real children — that is, opening up their minds and seeing how they work.)



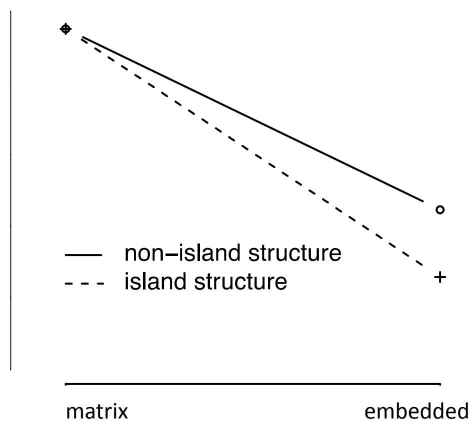
Complex NP



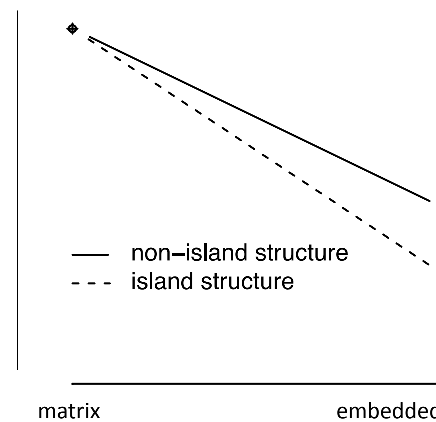
Subject



Whether



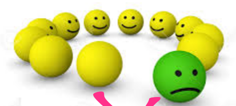
Adjunct



✓ Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



Wh … $[_{CN1}$ … $[_{CN2}$ … $\cancel{[_{CN3}}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

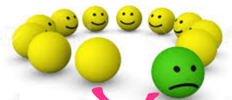
What's going on?

Why are the island-spanning dependencies so much worse than the grammatical ones?

syntax

syntactic island

✓ Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{\cancel{CN3}}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

What's going on?

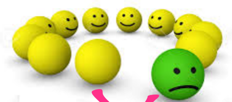
Why are the island-spanning dependencies so much worse than the grammatical ones?

Let's look inside them and see!

✓ Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{\cancel{CN3}}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

Let's look inside them and see!

It turns out that each island-spanning dependency contains at least one very low probability container node trigram. So these are the relevant "island" representations.

a. Complex NP

(i) * What did $[_{IP}$ the teacher $[_{VP}$ make $[_{NP}$ the claim $_{CP_{that}}$ that $[_{IP}$ Lily $_{VP}$ forgot __]]]]]?

(ii) *start*-IP-VP-NP-CP_{that}-IP-VP-*end*

(iii) Low probability:

VP-NP-CP_{that}

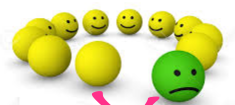
NP-CP_{that}-IP

syntax

syntactic island

✓ Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)



Wh … $[_{CN1}$ … $[_{CN2}$ … $\cancel{[_{CN3}}$ … $[_{CN4}$ … $[_{CN5}$ … __]]

Let's look inside them and see!

It turns out that each island-spanning dependency contains at least one very low probability container node trigram. So these are the relevant "island" representations.

b. Subject

(i) * Who does $[_{IP}$ Jack $[_{VP}$ think $[_{CP_{null}}$ $[_{IP}$ $[_{NP}$ the necklace $[_{PP}$ for __]] is expensive]]]]?

(ii) *start*-IP-VP-CP_{null}-IP-NP-PP-*end*

(iii) Low probability:

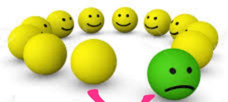
CP_{null}-IP-NP

syntax

syntactic island

✓ Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … ___]]

Let's look inside them and see!

It turns out that each island-spanning dependency contains at least one very low probability container node trigram. So these are the relevant "island" representations.

c. Whether

(i) * What does $[_{IP}$ the teacher $[_{VP}$ wonder $[_{CP_{whether}}$ whether $[_{IP}$ Jack $[_{VP}$ stole __]]]]]?

(ii) *start*-IP-VP-$CP_{whether}$-IP-VP-*end*

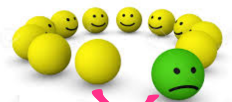
(iii) Low probability:

IP-VP-$CP_{whether}$
VP-$CP_{whether}$-IP
$CP_{whether}$-IP-VP

✓ Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

syntax

syntactic island



Wh ... $[_{CN1}$ ... $[_{CN2}$ ... $[_{CN3}$ ... $[_{CN4}$ ... $[_{CN5}$ ... __]]

Let's look inside them and see!

It turns out that each island-spanning dependency contains at least one very low probability container node trigram. So these are the relevant "island" representations.

d. Adjunct

(i) * What does $[_{IP}$ the teacher $[_{VP}$ worry $[_{CP_{if}}$ if $[_{IP}$ Lily $[_{VP}$ forgot __]]]]]?

(ii) *start*-IP-VP-CP_{if}-IP-VP-*end*

(iii) Low probability:

IP-VP-CP_{if}

VP-CP_{if}-IP

CP_{if}-IP-VP

Learning strategies


syntax

syntactic island

Subjacency (Chomsky 1973, Huang 1982, Lasnik & Saito 1984)

can't cross 2+ bounding nodes
from a fixed set (CP, IP, and/or NP)

Wh … $[_{BN1}$ … $[_{BN2}$ … ___]]

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)

A dependency can't cross a very low probability sequence of container nodes

Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … ___]]

In common: Local structural anomaly is the problem

The way Subjacency-ish implements this local structural anomaly can allow the development of syntactic island knowledge without relying on prior knowledge about bounding nodes and how many a dependency is limited to crossing.

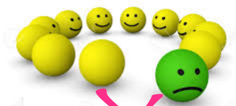
Less reliance on island-specific prior knowledge

Learning strategies

syntax

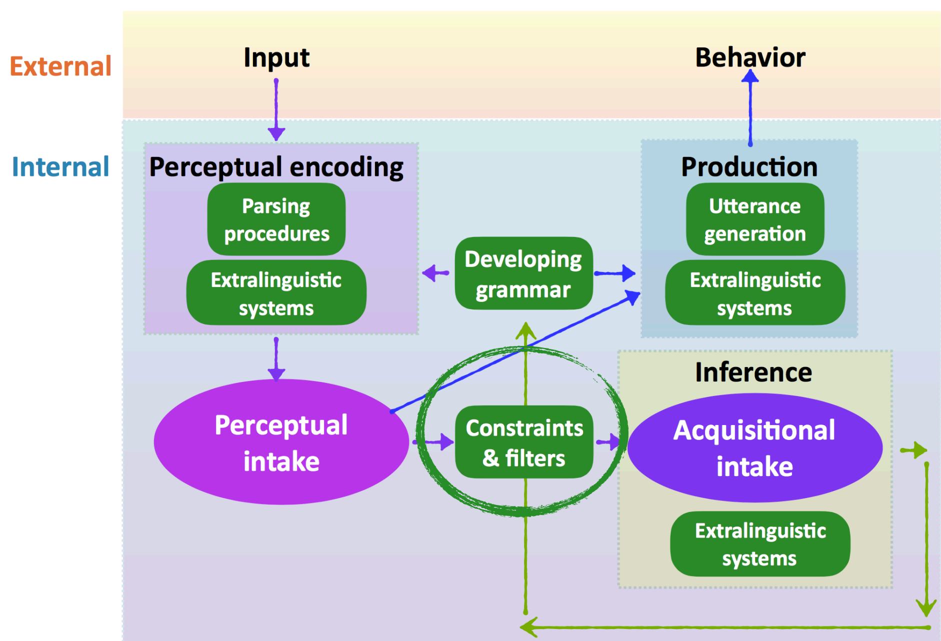
syntactic island

Subjacency-ish (Pearl & Sprouse 2013a, 2013b, 2015)



Wh … $[_{CN1}$ … $[_{CN2}$ … $[_{CN3}$ … $[_{CN4}$ … $[_{CN5}$ … ___]]

Less reliance on island-specific prior knowledge



Recap

- Syntactic islands are pieces of structure that don't allow *wh*-dependencies to cross them, and children have to learn what the syntactic islands are for their language

- It isn't obvious from children's input how they could learn about these syntactic islands — they need to generalize from their experience with only a few types of dependencies.

- One way to overcome this problem is to rely on island-specific innate knowledge in the form of Subjacency.

- Another way is to combine probabilistic learning with knowledge of phrase structure nodes that's not just specific to learning about islands. This strategy encodes islands as pieces of structure that a *wh*-dependency has a very low probability of crossing, based on the child's language experience.

Questions?

You should be able to do up through 2 on HW6
and up through 13 on the review questions for syntax & sentences.