

Psych 156A/ Ling 150:
Acquisition of Language II

Lecture 15
Introduction to Language Structure

Announcements

HW3 is due by the end of class today

Review questions are available for structure

Online course evaluations are available for this class - please fill them out! :)

Computational Problem:
Figure out the order of words (syntax)



Jareth juggles crystals
Subject Verb Object
Noun Verb Noun
NP NP

Depends on grammatical categories like Nouns and Verbs (and their associated phrases (NP)), but also on more precise distinctions like Subjects and Objects.

Some Noun Phrase distinctions:

Subject = usually the agent/actor of the action, "doer": Jareth
Object = usually the recipient of the action, "done to": crystals

Computational Problem:
Figure out the order of words (syntax)



Jareth juggles crystals
Subject Verb Object

Important idea: The observable word order speakers produce (like Subject Object Verb) is the result of a system of word order rules that speakers unconsciously use when they speak. This system of word order rules is called **syntax**.

Computational Problem: Figure out the order of words (syntax)



Jareth juggles crystals
Subject Verb Object

A third way to generate Subject Verb Object order:
The linguistic system specifies Subject Object Verb as the general pattern, but the Object moves after the Verb in certain contexts (the Object is unexpected information). Kannada is a language like this.

Kannada Subject Object Verb

Computational Problem: Figure out the order of words (syntax)



Jareth juggles crystals
Subject Verb Object

A third way to generate Subject Verb Object order:
The linguistic system specifies Subject Object Verb as the general pattern, but the Object moves after the Verb in certain contexts (the Object is unexpected information). Kannada is a language like this.

Kannada Subject ^{movement rule} Object Verb Object

Computational Problem: Figure out the order of words (syntax)



Jareth juggles crystals
Subject Verb Object

English
Subject Verb Object

German
Subject Verb Subject Object Verb

Kannada
Subject Object Verb Object

The learning problem: How do children know which system their language uses?

Computational Problem: Figure out the order of words (syntax)



Jareth juggles crystals
Subject Verb Object

English
Subject Verb Object

German
Subject Verb Subject Object Verb

Kannada
Subject Object Verb Object

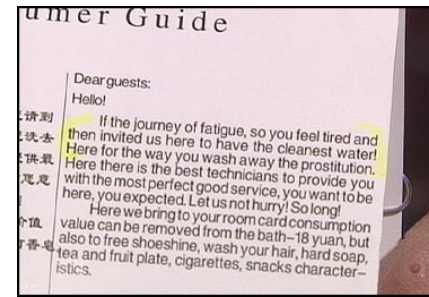
This is a hard question!

Children only see the output of the system (the observable word order of Subject Verb Object).



Syntax: One reason why translation is so hard

Translation is not so easy:
more than just word-by-word



http://www.nbc.com/nbc/The_Tonight_Show_with_Jay_Leno/headlines/

Translation is not so easy:
more than just word-by-word

translate.google.com

Through dangers untold and hardships unnumbered, I have fought my way here to the castle beyond the goblin city to take back the child that you have stolen.

Hebrew

דרך סכנות עצומות וקשיים לא ממוספרים, יש לי נלחם בדרך שלי כאן לטירה מעבר לעיר גובלין לקחת בחזרה את הילד שיש לך נגנב.

Literally:

Through dangers immense and difficulties not numbered, there-is to-me fighting through my here castle transition city goblin take back you child there-was-to-you stolen.

Translation is not so easy:
more than just word-by-word

translate.google.com

Through dangers untold and hardships unnumbered, I have fought my way here to the castle beyond the goblin city to take back the child that you have stolen.

Haitian Creole

Atravè danje inonbrabl ak difikilte inonbrabl, mwen te goumen jan m 'isit la yo chato la pi lwen pase lavil la Goblin yo pran tounen timoun nan ke ou te vòlè li.

Literally:

Through danger countless and difficulties countless, I was fight how me here they mansion the more far than cities the Goblin they take back children of that you was thief it.

Translation is not so easy: more than just word-by-word

translate.google.com

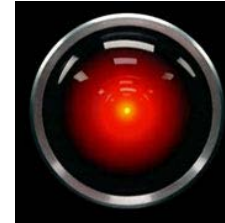
Through dangers untold and hardships unnumbered, I have fought my way here to the castle beyond the goblin city to take back the child that you have stolen.

Hindi

अनकहा और बेशुमार कठिनाइयों खतरों के माध्यम से, मैं तुम्हें चुराया है कि बच्चे को वापस लेने के लिए भूत शहर पर महल को यहाँ अपने तरीके से लड़ाई लड़ी है.

Literally:

Untold and uncountable difficulties threats medium through, I you stole is that children back take the ghost city beyond palace the here your methods from fight fought.



Syntax: One reason why natural language comprehension is so hard for computers

Solving the Language Problem (Artificial Intelligence)

HAL 9000 from 2001: A Space Odyssey (1968)

Perfect production and comprehension of English.



1960s: Language not considered one of the “hard” problems of artificial intelligence.

2010: Getting better but still not perfect.

<http://www.research.att.com/~ttsweb/tts/demo.php>

Solving the Language Problem (Artificial Intelligence)

2012: Apple’s Siri is getting closer, though still has problems ...



http://bits.blogs.nytimes.com/2012/07/15/with-apple's-siri-a-romance-gone-sour/?_php=true&_type=blogs&_r=0

Late last summer, I was introduced to a new special someone. I wasn’t looking to meet this new muse; it all just kind of happened.

We met at an Apple product announcement in Cupertino, Calif. She was helpful, smart and even funny, cracking sarcastic jokes and making me laugh. What more could a guy ask for?

Since then, we have had some major communication issues. She frequently misunderstands what I’m saying. Sometimes she is just unavailable. Often, she responds with the same, repetitive statement.

Her name is Siri.



Solving the Language Problem (Artificial Intelligence)

Contrast: Chess-playing.

In 1997, a program named Deep Blue beat the reigning world champion in chess. It did this by having enough computational resources to investigate every move option before it actually made the chess move. This shows that computers' poor performance on language is not about insufficient computational power, since there is enough computational power to solve the chess-playing problem (which some people might consider a very difficult problem).



Solving the Language Problem (Artificial Intelligence)

Update for 2011 on a machine's abilities to do what humans do:

Man vs. Machine (Watson) in Jeopardy
& how hard a problem language comprehension and production is

<http://www.youtube.com/watch?v=dr7lxQeXr7g>
(approximately 9 min video)

Watson vs. all humanity
https://www.youtube.com/watch?v=WFR3IOm_xhE
(approximately 4 min video)

Solving the Language Problem (Artificial Intelligence)

2013: True on-the-fly language comprehension is still pretty hard, as well as determining the answer to "commonsense" questions that are phrased naturally.

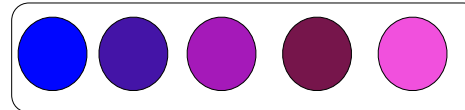
<http://www.sciencedaily.com/releases/2013/07/130715151059.htm>

"One of the hardest problems in building an artificial intelligence, Sloan said, is devising a computer program that can make sound and prudent judgment based on a simple perception of the situation or facts-the dictionary definition of commonsense.

Commonsense has eluded AI engineers because it requires both a very large collection of facts and what Sloan calls implicit facts — things so obvious that we don't know we know them. A computer may know the temperature at which water freezes, but we know that ice is cold." - Jeanne Galatzer-Levy

"We're still very far from programs with commonsense-AI that can answer comprehension questions with the skill of a child of 8," said Sloan. He and his colleagues hope the study will help to focus attention on the "hard spots" in AI research.

About human knowledge:
Language & variation



Navajo Code Talkers



Crucial cryptographic method used in World War II

http://en.wikipedia.org/wiki/Code_talker#Use_of_Navajo

“...Johnston saw Navajo as answering the military requirement for an undecipherable code. Navajo was spoken only on the Navajo lands of the American Southwest, and its syntax and tonal qualities, not to mention dialects, made it unintelligible to anyone without extensive exposure and training. One estimate indicates that at the outbreak of World War II fewer than 30 non-Navajos could understand the language...”

<https://www.youtube.com/watch?v=5rSvm3m8ZUA>
(~3 min video)

Navajo Code Talker Paradox (Baker 2001)



English must be very different from Navajo
Japanese could decode English, but couldn't decode Navajo when they didn't know it was Navajo.

English must be similar to Navajo

English can be translated into Navajo and back with no loss of meaning. (Languages are not just a product of the culture - pastoral Arizona lifestyle couldn't have prepared the code talkers for Pacific Island high tech warfare. Yet, translation was still possible.)

Types of variation

Vocabulary

English “think” verbs: think, know, wonder, suppose, assume, ...

Multiple types of the action verb “think”. Each has certain uses that are appropriate.

“I wonder whether the girl saved her little brother from the goblins.” [grammatical]

* “I suppose whether the girl saved her little brother from the goblins.” [ungrammatical]

Types of variation

Vocabulary

English “think” verbs: think, know, wonder, suppose, assume, ...

Navajo “carry” verbs: depends on object being carried
aaah (carry a solid round-ish object)



kaah (carry an open container with contents)



lé (carry a flexible object)



Types of variation

Sounds: Each language uses a particular subset of the sounds in the International Phonetic Alphabet, which represents all the sounds used in all human languages. There's often overlap (ex: "m", "p" are used in many languages), but languages also may make use of the less common sounds.

less common English sounds: "th" [θ], "th" [ð]

less common Navajo sounds: "whispered l", "nasalized a", ...

	Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Retroflex	Palatal	Velar	Uvular	Pharyngeal	Glottal
Plosive	p b			t d		ʈ ɖ	c ɟ	k ɡ	q ɢ		ʔ
Nasal	m	ɱ		n		ɳ	ɲ	ŋ	ɴ		
Trill				r					ʀ		
Tap or Flap				ɾ		ɽ					
Fricative	ɸ β	f v	θ ð	s z	ʃ ʒ	ʂ ʐ	ç ʝ	x ɣ	χ ʁ	ħ ʕ	h ɦ
Lateral fricative				ɬ ɮ							
Approximant		ʋ		ɹ		ɻ	j	ɰ			
Lateral approximant				l		ɭ	ʎ	ʟ			

Types of variation

Morphology (word forms)

English: invariant word forms

"the girl is crying", "I am crying"

Navajo: no invariant forms (there may be 100-200 prefixes for verb stems)

At'ééd **yicha**. "Girl crying"

Yishcha. "I am crying"

(yi + sh + cha)

Ninááhwiishdlaad. "I am again plowing"

(ni + náá + ho + hi + sh + l + dlaad)

Types of variation

Word order (syntax)

English: **Subject Verb Object** (invariant word order)

"The boy saw the girl"

Navajo: **Subject Object Verb, Object Subject Verb**

(varying word orders, meaning depends only on verb's form)

Ashkii at'ééd **yiyiiltsá**

boy girl saw

"The boy saw the girl"

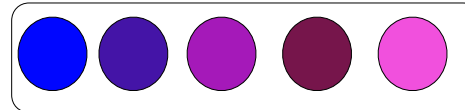


Ashkii at'ééd **biilstá**

boy girl saw

"The girl saw the boy"

Thinking about syntactic variation

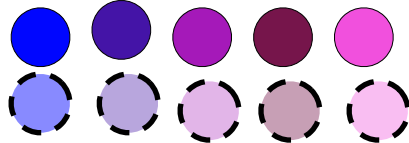


Similarities & differences: Parameters

Chomsky: Different combinations of different basic elements (**parameters**) would yield the observable languages (similar to the way different combinations of different basic elements in chemistry yield many different-seeming substances).



Big Idea: A relatively small number of syntax parameters yields a large number of different languages' syntactic systems.

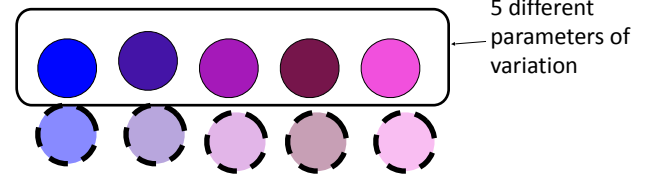


Similarities & differences: Parameters

Chomsky: Different combinations of different basic elements (**parameters**) would yield the observable languages (similar to the way different combinations of different basic elements in chemistry yield many different-seeming substances).



Big Idea: A relatively small number of syntax parameters yields a large number of different languages' syntactic systems.

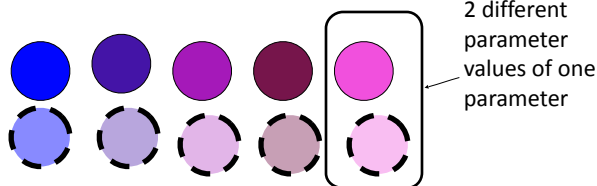


Similarities & differences: Parameters

Chomsky: Different combinations of different basic elements (**parameters**) would yield the observable languages (similar to the way different combinations of different basic elements in chemistry yield many different-seeming substances).



Big Idea: A relatively small number of syntax parameters yields a large number of different languages' syntactic systems.

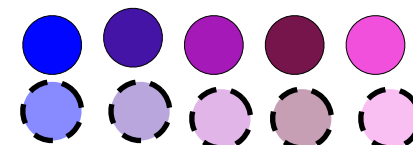


Similarities & differences: Parameters

Chomsky: Different combinations of different basic elements (**parameters**) would yield the observable languages (similar to the way different combinations of different basic elements in chemistry yield many different-seeming substances).



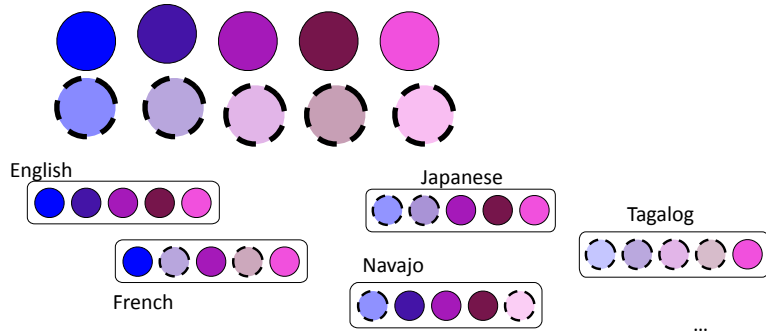
Big Idea: A relatively small number of syntax parameters yields a large number of different languages' syntactic systems.



Total languages that can be represented
 $= 2 * 2 * 2 * 2 * 2$
 $= 2^5$
 $= 32$

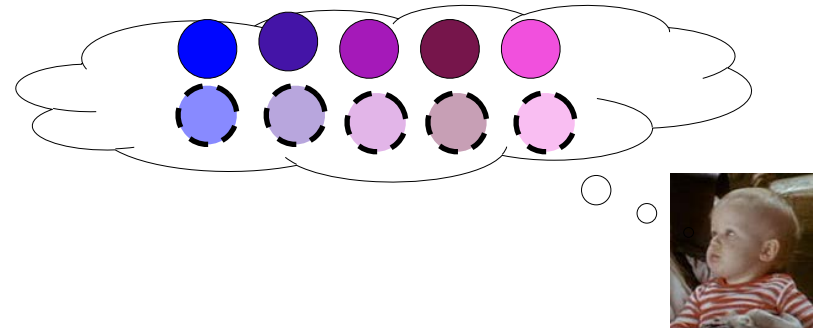
Similarities & differences: Parameters

Big Idea: A relatively small number of syntax parameters yields a large number of different languages' syntactic systems.



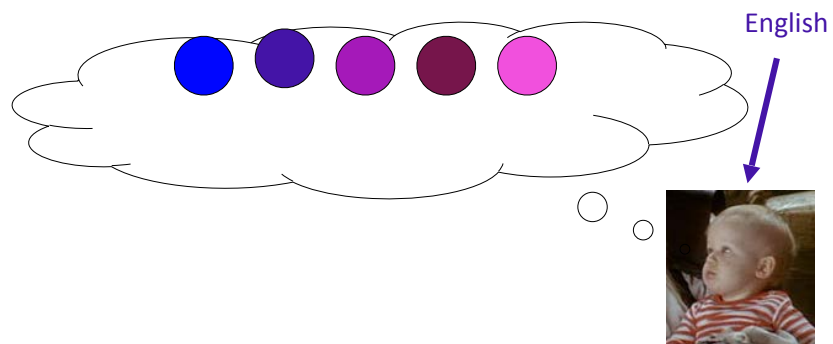
Learning language structure

Chomsky: Children are born knowing the parameters of variation. This is part of **Universal Grammar**. Input from the native linguistic environment determines what values these parameters should have.



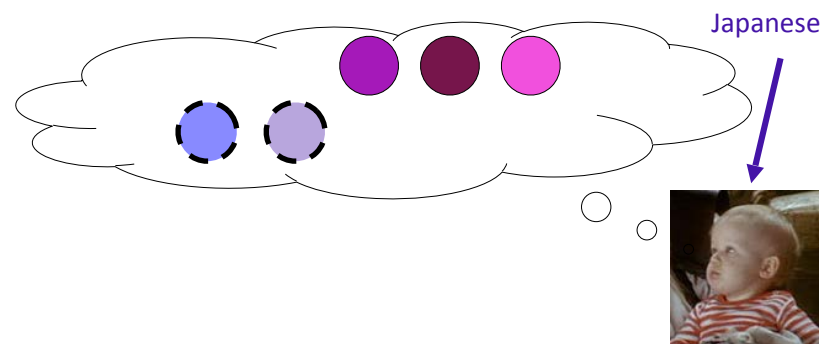
Learning language structure

Chomsky: Children are born knowing the parameters of variation. This is part of **Universal Grammar**. Input from the native linguistic environment determines what values these parameters should have.



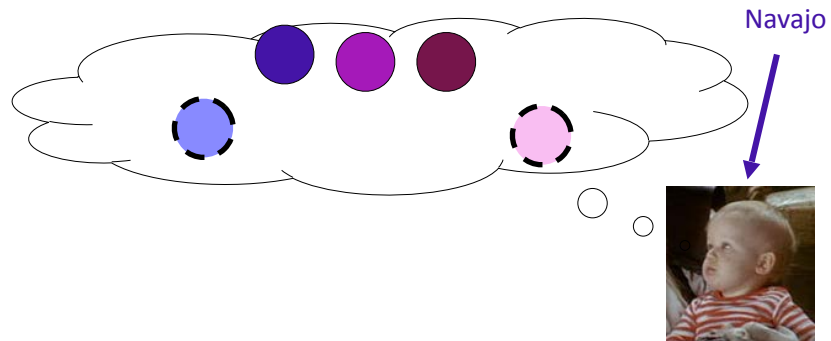
Learning language structure

Chomsky: Children are born knowing the parameters of variation. This is part of **Universal Grammar**. Input from the native linguistic environment determines what values these parameters should have.



Learning language structure

Chomsky: Children are born knowing the parameters of variation. This is part of **Universal Grammar**. Input from the native linguistic environment determines what values these parameters should have.



Generalizations about language structure

Greenberg's word order generalizations

Navajo

Japanese

Greenberg's word order generalizations

Navajo

Japanese

Basic word order:

Subject Object Verb

Basic word order:

Subject Object Verb

Ashkii at'ééd yiyiiltsá
boy girl saw

Jareth-ga Hoggle-o butta
Jareth Hoggle hit

"The boy saw the girl"

"Jareth hit Hoggle"

Greenberg's word order generalizations

Navajo

Postpositions:

Noun Phrase Postposition

'ée' biih náásdzá
clothing into I-got-back
"I got back into (my) clothes."

Japanese

Postpositions:

Noun Phrase Postposition

Jareth-ga Sarah to kuruma da
Jareth Sarah with car by

London ni itta
London to went

"Jareth went to London with Sarah by car."

Greenberg's word order generalizations

Navajo

Possessor before Possessed

Possessor Possession

Chidí bi-jáád
Car its-leg

"the car's wheel"

Japanese

Possessor before Possessed

Possessor Possession

Toby-no imooto-ga
Toby's sister

"Toby's sister"

Greenberg's word order generalizations

Navajo

Basic word order:

Subject Object Verb

Postpositions:

Noun Phrase Postposition

Possessor before Possessed

Possessor Possession

Japanese

Basic word order:

Subject Object Verb

Postpositions:

Noun Phrase Postposition

Possessor before Possessed

Possessor Possession

Despite the differences in the languages (and their cultural histories), both Japanese and Navajo are very similar when viewed through these three structural descriptions.

Greenberg's word order generalizations

English

Edo (Nigeria)

Greenberg's word order generalizations

English

Edo (Nigeria)

Basic word order:
Subject Verb Object

Basic word order:
Subject Verb Object

Sarah found Toby

Òzó mién Adésuwá
Ozo found Adesuwa

Greenberg's word order generalizations

English

Edo (Nigeria)

Prepositions:
Preposition Noun Phrase

Prepositions:
Preposition Noun Phrase

Jareth gave the crystal to Sarah

Òzó rhié néné ebé né Adésuwá
Ozo gave the book to Adesuwa

Greenberg's word order generalizations

English

Edo (Nigeria)

Possessed before Possessor

Possessed before Possessor

Possession Possessor

Possession Possessor

quest of Sarah

Omo Ozó
child Ozo

(alternative: Sarah's quest)

"child of Ozo"

Greenberg's word order generalizations

English

Edo (Nigeria)

Basic word order:
Subject Verb Object

Basic word order:
Subject Verb Object

Prepositions:
Preposition Noun Phrase

Prepositions:
Preposition Noun Phrase

Possessed before Possessor
Possession Possessor

Possessed before Possessor
Possession Possessor

Again, despite the differences in the languages (and their cultural histories), both English and Edo are very similar when viewed through these three structural descriptions.

Greenberg's word order generalizations

Greenberg found forty-five “universals” of languages - patterns overwhelmingly followed by languages with unshared history (Navajo & Japanese, English & Edo)

Not all combinations are possible - some patterns rarely appear
Ex: Subject Verb Object language (English/Edo-like) + postpositions (Navajo/Japanese-like)

Moral: Languages may be more similar than they first appear “on the surface”, especially if we consider their structural properties.

One potential parameter

English	Italian
Subject Verb	Subject Verb
	Jareth verrá
	<i>Jareth will-come</i>
“Jareth will come.”	“Jareth will come.”
grammatical	grammatical

One potential parameter

English	Italian
*Verb Subject	Verb Subject
	Verrá Jareth
<i>*Will arrive Jareth</i>	<i>Will-arrive Jareth</i>
	“Jareth will arrive”
ungrammatical	grammatical

One potential parameter

English	Italian
*Verb	Verb
	Verrá
<i>Will come</i>	<i>He-will-come</i>
	“He will come”
ungrammatical	grammatical

One potential parameter

English

Italian

Subject Verb

Subject Verb

*Verb Subject

Verb Subject

*Verb

Verb

These word order patterns might be fairly easy to notice. They involve the combinations of Subject and Verb that are grammatical in the language. A child might be able to notice the prevalence of some patterns and the absence of others.

One potential parameter

Expletive subjects: words without content
(may be more difficult to notice)

English

Italian

Raining.

Piove.

It-rains.

“It’s raining.”

“It’s raining.”

Not okay to leave out
expletive subject “it”.

Okay to leave out
expletive subject “it”.

One potential parameter

That-trace effect for subject questions

English

Italian

Who do you think (*that) will come?

Requires no “that” in embedded clause, despite allowing “that” in declaratives and object questions

I think (that) Hoggle will save Sarah.

Who did you think (that) Hoggle would save?

One potential parameter

That-trace effect for subject questions

English

Italian

Credi *che* Jareth verrà.

You think that Jareth will-come.

“You think that Jareth will come.”

Che credi *che* __ verrà?

Who think-you that will-come?

“Who do you think will come?”

Allows “that” in the embedded clause of a subject question (and declarative clauses).

One potential parameter

English

Subject Verb

*Verb Subject

*Verb

Not okay to leave out
expletive **subject** "it".

Requires special action for
embedded **subject** questions.

Italian

Subject Verb

Verb Subject

Verb

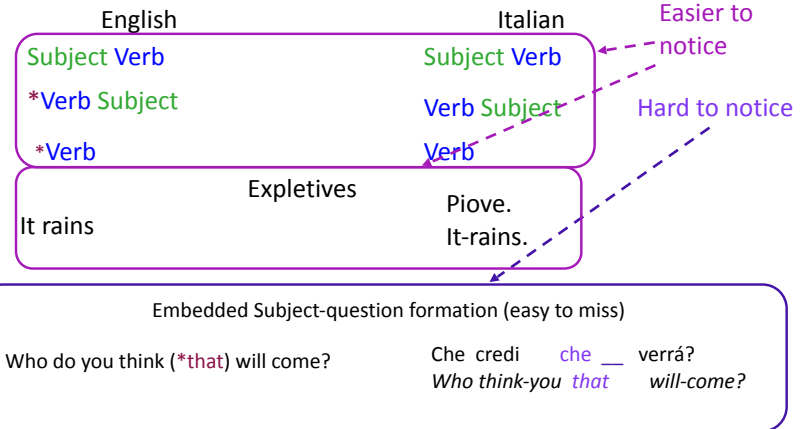
Okay to leave out
expletive **subject** "it".

Does not require special action
for embedded **subject**
questions.

All these involve the subject in some way - coincidence?
Idea: No! There's a language parameter involving the subject.

The Value of Parameters: Learning the Hard Stuff by Noticing the Easy Patterns

English vs. Italian: Subject Parameter



The Value of Parameters: Learning the Hard Stuff by Noticing the Easy Patterns

English vs. Italian: Subject Parameter

Big idea: If all these structural patterns are generated from the same linguistic parameter (e.g. a "subject" parameter), then children can learn the hard-to-notice patterns (like the patterns of embedded subject questions) by being exposed to the easy-to-notice patterns (like the optional use of subjects with verbs). The hard-to-notice patterns are generated by one setting of the parameter, which children can learn from the easy-to-notice patterns.

Children's knowledge of language structure variation is believed by linguistic nativists to be part of **Universal Grammar**, which children are born with.

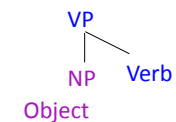
Another possible parameter

Syntax: the **Head Directionality** parameter (Baker 2001, Cook & Newson 1996): heads of phrases (ex: Nouns of Noun Phrases, Verbs of Verb Phrases, Prepositions of Preposition Phrases) are consistently in either the leftmost or rightmost position

Japanese/Navajo: Head-Last

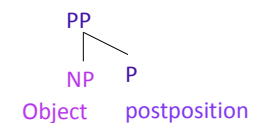
Verb Phrase:

Object Verb



Postpositions:

Noun Phrase Postposition

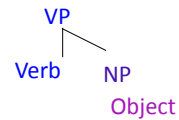


Another possible parameter

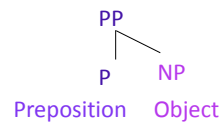
Syntax: the **Head Directionality** parameter (Baker 2001, Cook & Newson 1996): heads of phrases (ex: Nouns of Noun Phrases, Verbs of Verb Phrases, Prepositions of Preposition Phrases) are consistently in either the leftmost or rightmost position

Edo/English: Head-First

Verb Phrase:
Verb Object



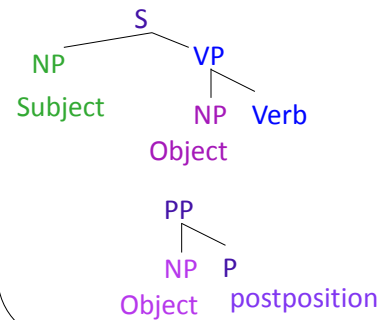
Prepositions:
Preposition Noun Phrase



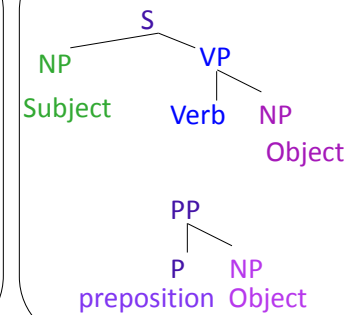
Universal Grammar: Parameters

At this level of structural analysis (parameters), languages differ vary minimally from each other. This makes language structure much easier for children to learn. All they need to do is set the right parameter values for their language, based on the data that are easy to observe.

Japanese/Navajo



Edo/English



Questions?



You should be able to do up through question 9 on the structure review questions