Psych 215L: Language Acquisition

Lecture 17

Complex Systems I: Parameters

Complex Linguistic Systems

What is the generative system that creates the observed (structured) data of language (ex: syntax, metrical phonology)?

Observable data: word order Subject Verb Object

Complex Linguistic Systems

What is the generative system that creates the observed (structured) data of language (ex: syntax, metrical phonology)?

Observable data: word order Subject Verb Object

Kannada

Subject t_{Object} Verb Object

English

Subject Verb Object

German

Subject Verb $t_{Subject}$ Object t_{Verb}

Complex Linguistic Systems

What is the generative system that creates the observed (structured) data of language (ex: syntax, metrical phonology)?

Observable data: stress contour EMphasis

Complex Linguistic Systems

What is the generative system that creates the observed (structured) data of language (ex: syntax, metrical phonology)?

Observable data: stress contour EMphasis

(S S) S
EM pha sis

(H L) H
EM pha sis

(H L L)
EM pha sis

(S S S)
EM pha sis

General Problems with Learning Complex Linguistic Systems

What children encounter: the output of the generative linguistic system

Subject Verb Object EMphasis

General Problems with Learning Complex Linguistic Systems

What children must learn: the components of the system that combine to generate this observable output

Subject Verb Object EMphasis

Does the Verb move? When/where?

Does the Object move? When/where?

Does the Subject move? When/where?

Are syllables differentiated?

Are all syllables included in larger units?

Which syllable(s) of a larger unit is/are stressed?

General Problems with Learning Complex Linguistic Systems

Why this is tricky: There is often a non-transparent relationship between the observable form of the data and the underlying system that produced it. ***Hard to know what parameters of variation to consider.***

Subject Verb Object EMphasis

Does the Verb move? When/where?

Does the Object move? When/where?

Does the Subject move? When/where?

Are syllables differentiated?

Are all syllables included in larger units?

Which syllable(s) of a larger unit is/are stressed?

General Problems with Learning Complex Linguistic Systems

Why this is tricky: There is often a non-transparent relationship between the observable form of the data and the underlying system that produced it. ***Hard to know what parameters of variation to consider.***

Subject Verb Object

Does the Verb move?
When/where?

Does the Object move?
When/where?

Does the Subject move?
When/where?

EMphasis

Are syllables differentiated?

Are all syllables included in larger units?

Which syllable(s) of a larger unit is/are stressed?

Observation: Languages only differ in constrained ways from each other. Not all generalizations are possible.

Idea: Bias on hypothesis space - children's hypotheses are constrained so they only consider generalizations that are possible in the world's languages.

Chomsky (1981), Halle & Vergnaud (1987), Tesar & Smolensky (2000)

Parameters

A parameter is meant to be something that can account for multiple observations in some domain.

Parameter for a statistical model: determines what the model predicts will be observed in the world in a variety of situations

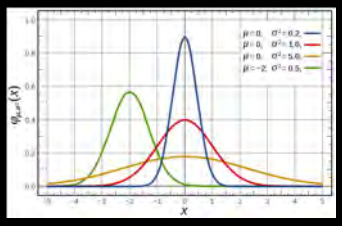
Parameter for our minds (and language): determines what *we* predict will be observed in the world in a variety of situations

Statistical Parameters

The normal distribution is a statistical model that uses two parameters:

- μ for the mean
- σ for the standard deviation

$$\varphi_{\mu,\sigma^2}(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$$

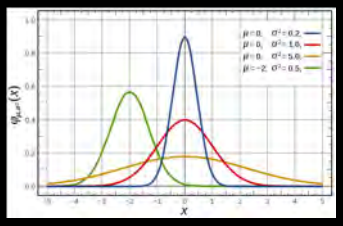


If we know the values of these parameters, we can make predictions about the likelihood of data we rarely or never see.

Statistical Parameters

Suppose this is a model of how many minutes late you'll be to class.

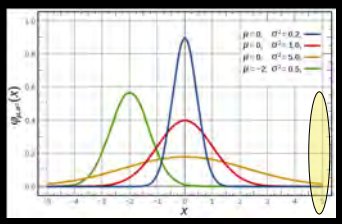
Let's use the model with $\mu = 0$, and $\sigma^2 = 0.2$. (blue line)



Statistical Parameters

Suppose this is a model of how many minutes late you'll be to class.

Let's use the model with $\mu = 0$, and $\sigma^2 = 0.2$. (blue line)



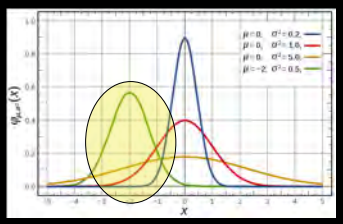
How likely are you to be 5 minutes late, given these parameters?

Not very likely! We can tell this just by knowing the values of the two statistical parameters. These parameter values allow us to infer the likelihood of some observed behavior.

Statistical Parameters

Observing different quantities of data with particular values can tell us which values of μ and σ^2 are most likely, if we know we are looking to determine the values of μ and σ^2 in function $\varphi(X)$

$$\varphi_{\mu,\sigma^2}(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(X-\mu)^2}{2\sigma^2}}$$

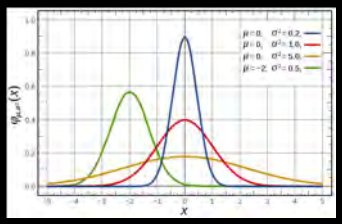


Observing data points distributed like the green curve tells us that μ is likely to be around 2, for example.

Statistical vs. Linguistic Parameters

Important similarity:

We do not see the process that generates the data, but only the data themselves. This means that in order to form our expectations about X, we are, in effect, reverse engineering the observable data.



Our knowledge of the underlying function/principle that generates these data - $\varphi(X)$ - as well as the associated parameters - μ, and σ^2 - allows us to represent an infinite number of expectations about the behavior of variable X.

Linguistic principles vs. linguistic parameters

Both principles and parameters are often thought of as innate domain-specific abstractions that connect to many structural properties about language.

Linguistic principles correspond to the properties that are invariant across all human languages. Comparison: the equation's form– it is the statistical "principle" that explains the observed data.

$$\varphi_{\mu,\sigma^2}(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(X-\mu)^2}{2\sigma^2}}$$

Linguistic parameters correspond to the properties that vary across human languages. Comparison: μ and σ^2 determine the exact form of the curve that represents the likelihood of observing certain data. While different values for these parameters can produce many different curves, these curves share their underlying form due to the common invariant function.

The utility of connecting to multiple properties

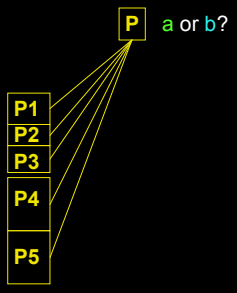
The fact that parameters connect to multiple structural properties then becomes a very good thing from the perspective of someone trying to acquire language. This is because a child can learn about that parameter's value by observing many different kinds of examples in the language.

"The richer the deductive structure associated with a particular parameter, the greater the range of potential 'triggering' data which will be available to the child for the 'fixing' of the particular parameter" – Hyams (1987)

Parameters can be especially useful when a child is trying to learn the things about language structure that are otherwise hard to learn, perhaps because they are very complex properties themselves or because they appear very infrequently in the available data

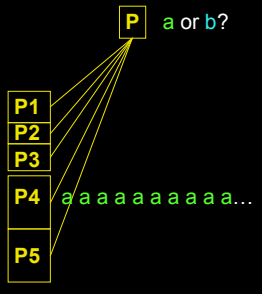
Why Hard-To-Learn Structures Are Easier

Let's assume a number of properties are all connected to parameter P, which can take one of two values: a or b.



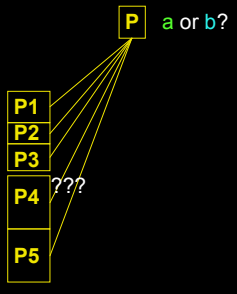
Why Hard-To-Learn Structures Are Easier

How do we learn whether P4 shows behavior a or b?
One way is to observe many instances of P4.



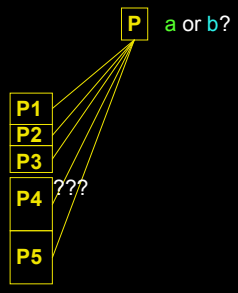
Why Hard-To-Learn Structures Are Easier

But what if P4 occurs very rarely? We might never see any examples of P4.



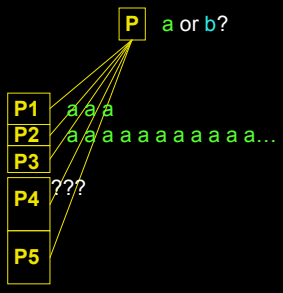
Why Hard-To-Learn Structures Are Easier

Fortunately, if P4 is connected to P, we can learn the value for P4 by learning the value of P. Also fortunately, P is connected to P1, P2, P3, and P5.



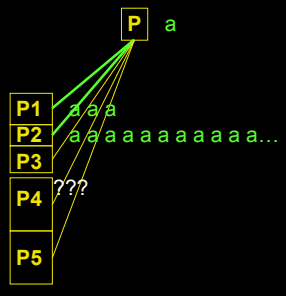
Why Hard-To-Learn Structures Are Easier

Step 1: Observe P1, P2, P3, *or* P5. In this case, all the observed examples of these structures are behavior a.



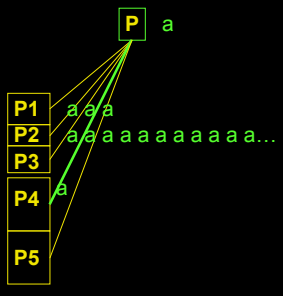
Why Hard-To-Learn Structures Are Easier

Step 2: Use this knowledge to set the value of parameter P to a.

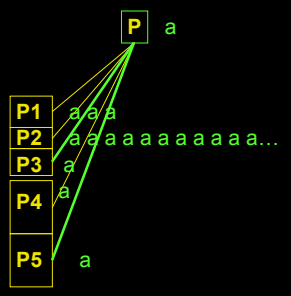


Why Hard-To-Learn Structures Are Easier

Step 3: Since parameter P is connected to P4, we can predict that P4 will also show behavior a - even though we've never seen any examples of it! (We can also infer P3 and P5 the same way.)



Why Acquisition Is Easier



This highlights another benefit of parameters - we don't have to learn the behavior of each structure individually. Instead, we can observe some structures (ex: P1 and P2) and infer the right behavior for the remaining structures (P3, P4, and P5).

That is, instead of having to make 5 decisions (one for P1, P2, P3, P4, and P5), we actually only need to make one decision - is P a or b?

What linguistic parameters are supposed to be

Parameter property 1:

Governs many different observable linguistic structures

Parameter property 2:

Varies in a constrained way from language to language

Used both in the theory of language acquisition and the theory of grammar typology to condense the representation of the language, thereby structuring the learning task for the child in such a way as to reduce the range of observations required to construct a grammar. In theory, this works by connecting together observations that might otherwise need to be accounted for independently from each other.

One potential parameter

English	Italian
Subject Verb	Subject Verb Jareth verrá *Jareth will-come*
"Jareth will come."	"Jareth will come."
grammatical	grammatical

One potential parameter

English	Italian
*Verb Subject	Verb Subject Verrá Jareth *Will-arrive Jareth*
*Will arrive Jareth	"Jareth will arrive"
ungrammatical	grammatical

One potential parameter

English	Italian
*Verb	Verb
Will come	Verrá *He-will-come*
	"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 verrá.
You think that Jareth will-come.
"You think that Jareth will come."

Che credi che __ verrá?
Who think-you that will-come?
"Who do you think will come?"

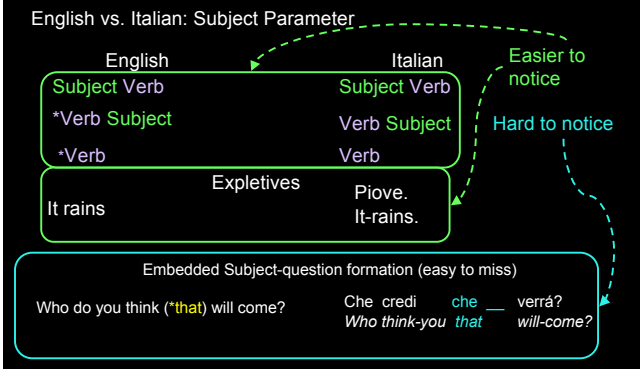
Allows "that" in the embedded clause of a subject question (and declarative clauses).

One potential parameter

English	Italian
Subject Verb	Subject Verb
*Verb Subject	Verb Subject
*Verb	Verb
Not okay to leave out expletive subject "it".	Okay to leave out expletive subject "it".
Requires special action for embedded subject questions.	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

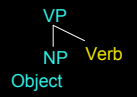


Another possible parameter

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



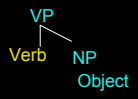
Postpositions:
Noun Phrase Postposition

Another possible parameter

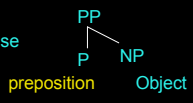
Syntax: the Head Directionality parameter (Baker 2001, Cook & Newson 1996): heads of phrases (ex: Nouns of Noun Phrases, Verbs of Verb Phrases, Prepositions of Preposition Phrases) are consistently in either the leftmost or rightmost position

Edo/English: Head-First

Verb Phrase:
Verb Object



Prepositions:
Preposition Noun Phrase



Some other current thoughts on parameters

Lasnik & Lohndal 2009: An important distinction

"…a sharp break from earlier approaches, under which universal grammar specified an infinite array of possible grammars, and …required an unfeasible search procedure to find the highest-valued one, given primary linguistic data….[now] There is no enumeration of the array of possible grammars. There are only finitely many targets for acquisition, and no search procedure apart from valuing parameters"

Lightfoot 2010: Cue-based learning with parameters

"Children are insensitive to the set of sentences generated by any grammar and the approach makes strong predictions about the 'learning path,' the sequence of structures in the growing internal language…one can view historical change as taking place when external language comes to express [parametric] cues differently…cues are abstract pieces of structure in the child's I-language and they are expressed by sentences"

Some other current thoughts on parameters

Neske 2010: Parameters & Universal Grammar

"UG is a theory of the initial state of the language faculty, which, in the P[rinciples]&P[arameters] model, undergoes a setting of parameters driven by external linguistic data. This is a selectionist account of learning…"

"What exactly does it mean for external linguistic data to 'set' a 'parameter'? The answer might involve a probabilistic component. In one learning scheme, UG represents the hypothesis space of grammars and parameter setting would involve the discarding of hypotheses that are inconsistent with external linguistic data….the child has a representation of many possible grammars, not just one (Crain & Pietroski 2001)….have a probability that is either increased or decreased depending on consistency or inconsistency with the linguistic input (Yang 2004)."

Some tricky language phenomena that children have to learn that are (likely) part of larger systems of knowledge [most from O'Grady 2005]

Complicated silent things

Sentences that have both an implied subject and implied object.

The girl is *eager* to see .

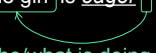
Who/what is doing the seeing (subject of see)?

Who/what is being seen (object of see)?

Complicated silent things

Sentences that have both an implied subject and implied object.

The girl is *eager* [ ] to see .



Who/what is doing the seeing (subject of see)?
The girl.

Who/what is being seen (object of see)?

Complicated silent things

Sentences that have both an implied subject and implied object.

The girl is *eager* [ ] to see [ ] .

Who/what is doing the seeing (subject of see)?
The girl.

Who/what is being seen (object of see)?
Something unspecified.
This sentence means approximately something like
“The girl is eager to see (something).”

Complicated silent things

Sentences that have both an implied subject and implied object.

The girl is *easy* to see .

Who/what is doing the seeing (subject of see)?

Who/what is being seen (object of see)?

Complicated silent things

Sentences that have both an implied subject and implied object.

The girl is *easy* to see [ ] .

Who/what is doing the seeing (subject of see)?

Who/what is being seen (object of see)?
The girl.

Complicated silent things

Sentences that have both an implied subject and implied object.

The girl is *easy* [ ] to see [ ] .

Who/what is doing the seeing (subject of see)?
Someone not mentioned.
This sentence means the same thing as
"It is easy (for someone) to see the girl."
Who/what is being seen (object of see)?
The girl.

Pronouns

Pronouns are energy-saving devices that allow us to refer to someone or something (whose identity we know) without using a name (like "Sarah" or "Jareth") or other noun phrase (like "the girl" or "a very impressive goblin king").

Sarah though that *she* could save her brother.

Jareth was surprised the girl summoned him, and resolved to show her he was a very impressive goblin king.

Pronouns

Young children seem to know how to use pronouns – they like to use them if a preceding noun has already established what they refer to.

Imitation task results with 2 ½ and 3-year-old children (Lust 1981):

Experimenter says a sentence with two names:
"Because Sam was thirsty, Sam drank some soda."

Child replaces second name with a pronoun:
"Because Sam was thirsty, he drank some soda."

Pronouns

Young children seem to know how to use pronouns – they like to use them if a preceding noun has already established what they refer to.

Imitation task results with 2 ½ and 3-year-old children (Lust 1981):

Experimenter says a sentence with a pronoun before a name:
"Because he was thirsty, Sam drank some soda."

Child replaces name and pronoun so name comes first:
"Because Sam was thirsty, he drank some soda."

Trickier Pronouns

Reflexive pronouns have different forms than "plain" pronouns

myself	me, I	herself	she, her
yourself	you	itself	it
himself	he, him	ourselves	we, us
themselves	they, them		

Trickier Pronouns

Reflexive pronouns behave differently than "plain" pronouns: they are interpreted differently

Jareth thought that Hoggle tricked himself.

Jareth thought that Hoggle tricked him.

Trickier Pronouns

Reflexive pronouns behave differently than "plain" pronouns: they are interpreted differently

Jareth thought that Hoggle tricked himself.
= Jareth thought that Hoggle tricked Hoggle.

Jareth thought that Hoggle tricked him.
= Jareth thought that Hoggle tricked Jareth.

Trickier Pronouns

Reflexive pronouns behave differently than "plain" pronouns: they are interpreted differently

Jareth thought that Hoggle tricked himself.

must refer to NP in same clause

Jareth thought that Hoggle tricked him.

must not refer to NP in same clause, but can refer to NP in different clause

Rule: Reflexive pronouns must refer to a noun phrase inside the same clause while regular pronouns must not.

Quantifiers

Quantifiers are words that express quantities, like *a, some, every, none*, and *most*.

When two (or more) quantifiers are in a sentence, they interact semantically to determine the sentence's meaning.

Everyone saw a movie last night.

Quantifiers

Quantifiers are words that express quantities, like *a, some, every, none*, and *most*.

When two (or more) quantifiers are in a sentence, they interact semantically to determine the sentence's meaning.

Everyone saw a movie last night.

every >> a:
For each person p, that person saw a movie m.

Compatible with this situation:
Lisa watched *Labyrinth*, Joseph watched *Troy*, and Benjamin watched *Serenity*.

Quantifiers

Quantifiers are words that express quantities, like *a, some, every, none*, and *most*.

When two (or more) quantifiers are in a sentence, they interact semantically to determine the sentence's meaning.

Everyone saw a movie last night.

a >> every:
For a movie m, every person saw m.

Compatible with this situation:
Lisa, Joseph, and Benjamin watched *Labyrinth*.

Quantifiers

Quantifiers are words that express quantities, like *a, some, every, none,* and *most*.

When two (or more) quantifiers are in a sentence, they interact semantically to determine the sentence's meaning.

Someone teases everyone. (Don't let it get you down!)

Quantifiers

Quantifiers are words that express quantities, like *a, some, every, none,* and *most*.

When two (or more) quantifiers are in a sentence, they interact semantically to determine the sentence's meaning.

Someone teases everyone. (Don't let it get you down!)

some >> every:
For some person p, p teases every (other) person.

Compatible with this situation:
Jareth teases Sarah, Jareth teases Hoggle, and Jareth teases Ludo.

Quantifiers

Quantifiers are words that express quantities, like *a, some, every, none,* and *most*.

When two (or more) quantifiers are in a sentence, they interact semantically to determine the sentence's meaning.

Someone teases everyone. (Don't let it get you down!)

every >> some:
For every person p, some (other) person p' teases them.

Compatible with this situation:
Jareth teases Hoggle, Sarah teases Sir Didymus, and Hoggle teases Sarah.

Syntactic Islands

Lily thought the letter (from the soldier) inspired the students.

Syntactic Islands

Lily thought the letter (from the soldier) inspired the students.

Questions formed from this statement:

Human ratings

Ask about Lily (simple subject extraction: main clause)
Who __ thought [the letter inspired the students]?

Who __ thought [the letter from the soldier inspired the students]?

Syntactic Islands

Lily thought the letter (from the soldier) inspired the students.

Questions formed from this statement:

Human ratings

Ask about Lily (simple subject extraction: main clause)
Who __ thought [the letter inspired the students]?

Who __ thought [the letter from the soldier inspired the students]?

Ask about the letter (simple subject extraction: embedded clause)
What did Lily think [__ inspired the students]?

Syntactic Islands

Lily thought the letter (from the soldier) inspired the students.

Questions formed from this statement:

Human ratings

Ask about Lily (simple subject extraction: main clause)
Who __ thought [the letter inspired the students]?

Who __ thought [the letter from the soldier inspired the students]?

Ask about the letter (simple subject extraction: embedded clause)
What did Lily think [__ inspired the students]?

Ask about the soldier (complex subject extraction: embedded clause)
* Who did Lily think [the letter from __] inspired the students?

Syntactic Islands

Lily thought the letter (from the soldier) inspired the students.

Questions formed from this statement:

Human ratings

Ask about Lily (simple subject extraction: main clause)
Who __ thought [the letter inspired the students]?

Who __ thought [the letter from the soldier inspired the students]?

Ask about the letter (simple subject extraction: embedded clause)
What did Lily think [__ inspired the students]?

Ask about the soldier (complex subject extraction: embedded clause)
* Who did Lily think [the letter from __] inspired the students?

Why can't we ask the last question? Conventional answer: You're trying to extract the question word from a "subject island". This is bad.