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: stress contour  Emphasis
**Complex Linguistic Systems**

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

<table>
<thead>
<tr>
<th>Observable data: stress contour</th>
<th>EMphasis</th>
</tr>
</thead>
<tbody>
<tr>
<td>(H L) H EM pha sis</td>
<td></td>
</tr>
<tr>
<td>(S S) S EM pha sis</td>
<td></td>
</tr>
<tr>
<td>(S S S) EM pha sis</td>
<td></td>
</tr>
</tbody>
</table>

**General Problems with Learning Complex Linguistic Systems**

What children encounter: the output of the generative linguistic system

<table>
<thead>
<tr>
<th>Subject</th>
<th>Verb</th>
<th>Object</th>
<th>EMphasis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**General Problems with Learning Complex Linguistic Systems**

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

<table>
<thead>
<tr>
<th>Subject</th>
<th>Verb</th>
<th>Object</th>
<th>EMphasis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**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.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Verb</th>
<th>Object</th>
<th>EMphasis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Verb</th>
<th>Object</th>
<th>EMphasis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Subject</th>
<th>Verb</th>
<th>Object</th>
<th>EMphasis</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
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.

<table>
<thead>
<tr>
<th>Subject</th>
<th>Verb</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does the Verb move?</td>
<td>When/where?</td>
<td></td>
</tr>
<tr>
<td>Does the Object move?</td>
<td>When/where?</td>
<td></td>
</tr>
<tr>
<td>Does the Subject move?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>When/where?</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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.


Parameter

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:
- \( \mu \) for the mean
- \( \sigma \) for the standard deviation

\[ f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \]

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 \( \sigma^2 = 2 \). (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 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 - \( q(X) \) - as well as the associated parameters - \( \mu \) and \( \sigma^2 \) - 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.

Linguistic parameters correspond to the properties that vary across human languages. Comparison: \( \mu \) and \( \sigma^2 \) 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.
A note on identifying universal linguistic principles
Nevins 2010

“...the study of impossible languages and their acquisition... By creating artificial and controlled examples of these unattested patterns we can observe whether they are unattested because of pure historico-geographic accident or due to more principled reasons, such as Universal Grammar – a set of analytic biases that prefer certain language types over others... it only takes a few skeptics to say that we simply haven’t found enough languages to know whether this is a true generalization or not, and that perhaps waiting for us in the Amazon is a language that violates exactly the universal we take to be central to human language structure... It is my contention that one of the most effective ways of examining whether there is a true analytic and cognitive bias for one type of linguistic structure over another is in teaching it to experimental participants who have neither in their native language, and seeing whether they learn or prefer one to the other.” - Nevins (2010)

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.

<table>
<thead>
<tr>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>a</td>
<td>b</td>
<td>a</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>P1</th>
<th>P2</th>
<th>P3</th>
<th>P4</th>
<th>P5</th>
</tr>
</thead>
<tbody>
<tr>
<td>a a a a a a a a...</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Why Hard-To-Learn Structures Are Easier

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

\[ P \] a or b?

\[ \text{P1, P2, P3, P4, P5} \]

---

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.

\[ P \] a or b?

---

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.

\[ P \] a or b?

\[ \text{P1, P2, P3, P4, P5} \]

---

Why Hard-To-Learn Structures Are Easier

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

\[ P \] a

\[ \text{P1, P2, P3, P4, P5} \]
Why Hard-To-Learn Structures Are Easier

Step 3: Since parameter \( P \) is connected to \( P_4 \), we can predict that \( P_4 \) will also show behavior \( a \) - even though we’ve never seen any examples of it! (We can also infer \( P_3 \) and \( P_5 \) 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: \( P_1 \) and \( P_2 \)) and infer the right behavior for the remaining structures (\( P_3, P_4, \) and \( P_5 \)). That is, instead of having to make 5 decisions (one for \( P_1, P_2, P_3, P_4, \) and \( P_5 \)), 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

<table>
<thead>
<tr>
<th>English</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject Verb</td>
<td>Subject Verb</td>
</tr>
<tr>
<td>Jareth will come.</td>
<td>Jareth verrá</td>
</tr>
<tr>
<td>Jareth will come.</td>
<td>Jareth will-come</td>
</tr>
</tbody>
</table>

grammatical
One potential parameter

<table>
<thead>
<tr>
<th>English</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Verb Subject</td>
<td>Verb Subject</td>
</tr>
<tr>
<td>*Will arrive Jareth</td>
<td>Verrà Jareth</td>
</tr>
<tr>
<td>ungrammatical</td>
<td>grammatical</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>English</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>*Verb</td>
<td>Verb</td>
</tr>
<tr>
<td>Will come</td>
<td>Verrà He-will-come</td>
</tr>
<tr>
<td>ungrammatical</td>
<td>grammatical</td>
</tr>
</tbody>
</table>

One potential parameter

<table>
<thead>
<tr>
<th>English</th>
<th>Italian</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expletive subjects: words without content (may be more difficult to notice)</td>
<td>Piove. It-rains. “It’s raining.”</td>
</tr>
<tr>
<td>English</td>
<td>Italian</td>
</tr>
<tr>
<td>Raining.</td>
<td>Piove.</td>
</tr>
<tr>
<td>“It’s raining.”</td>
<td>“It’s raining.”</td>
</tr>
<tr>
<td>Not okay to leave out expletive subject “it”.</td>
<td>Okay to leave out expletive subject “it”.</td>
</tr>
</tbody>
</table>
One potential parameter
That-trace effect for subject questions

English
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?

Italian
Credi che Jareth verrà.
"You think that Jareth will come."
Che credi che ... verrà?
"Who do you think will come?"

English vs. Italian: Subject Parameter

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

Easier to notice

Hard to notice

Embedded Subject-question formation (easy to miss)
Who do you think (*that*) will come?
Che credi che ... verrà?
Who think you ... will-come?
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: VP Object Verb

Postpositions: PP Noun Phrase Postposition

Edo/English: Head-First

Verb Phrase: VP NP Verb

Prepositions: PP Preposition NP

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[riniples]&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.

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)?
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.

Raising vs. Control Verbs
(Mitchener & Becker 2011)
Verbs that have specific syntactic behavior with specific semantic connotations.

The girl seems to be running.

Who is doing the running?
Verbs that have specific syntactic behavior with specific semantic connotations.

The gir__l__ seems __ to be running __.

Who is doing the running?
The girl. (The girl is the AGENT of the verb RUN.)

Who is doing the seeming?
Is it the girl?

"It seems that the girl is running." (expletive if)
SEEM is called a raising verb, since the subject (the girl) can "raise" to the main clause without changing the meaning of the sentence.
Verbs that have specific syntactic behavior with specific semantic connotations.

The girl is trying to run.

Who is doing the running?
The girl. The girl is the AGENT of RUN.

Who is doing the trying?
The girl? Probably, since we can’t use expletive it:
“IT tries that the girl is running.” The girl is the AGENT of TRY. TRY is called a control verb, since the subject of the control verb (try) seems to also control the subject of the embedded verb (run).

Some verbs are ambiguous between raising and control.

It began to rain.

BEGIN seems to be acting like a raising verb, since expletive it is the subject.
Raising vs. Control Verbs  
(Mitchener & Becker 2011)

Some verbs are ambiguous between raising and control.

He began to talk.

BEGIN seems to be acting like a control verb since he is the SUBJECT of BEGIN (“It begins him to talk”) and also the SUBJECT of TALK.

So how does a child learn how to use a novel verb, and what its semantics are?

He plorged to talk.

He is the SUBJECT of TALK…. …but is he also the SUBJECT of PLORG (control)? …or did he raise to that position (raising)?

See Mitchener & Becker (2011) for other semantic cues (animacy, eventivity) children may key into in order to determine what a verb’s syntax and semantics are. In addition, they assess whether biologically plausible learning algorithms could use this information to classify verbs.

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 thought 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.

Imitation task results with 2½ 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 2½ 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, it
- himself, he, him
- themselves, we, us
- same forms for “plain” pronouns

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.

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.

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.

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.

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.

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.
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.

**Some**one teases **every**one. (Don’t let it get you down!)

**Every**one teases **some**one. (Don’t let it get you down!)

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.

**Some**one teases **every**one. (Don’t let it get you down!)

**Some**one teases **every**one. (Don’t let it get you down!)

Compatible with this situation:

- **Jareth** teases **Hoggle**,
- **Sarah** teases **Sir Didymus**,
- **Hoggle** teases **Jareth**,
- **Ludo** teases **Jareth**.

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.

**Some**one teases **every**one. (Don’t let it get you down!)

**Some**one teases **every**one. (Don’t let it get you down!)

Compatible with this situation:

- **Jareth** teases **Hoggle**,
- **Sarah** teases **Sir Didymus**,
- **Hoggle** teases **Jareth**,
- **Ludo** teases **Jareth**.

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.

**Some**one teases **every**one. (Don’t let it get you down!)

**Some**one teases **every**one. (Don’t let it get you down!)

Compatible with this situation:

- **Jareth** teases **Hoggle**,
- **Sarah** teases **Sir Didymus**,
- **Hoggle** teases **Jareth**,
- **Ludo** teases **Jareth**.

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.

**Some**one teases **every**one. (Don’t let it get you down!)

**Some**one teases **every**one. (Don’t let it get you down!)

Compatible with this situation:

- **Jareth** teases **Hoggle**,
- **Sarah** teases **Sir Didymus**,
- **Hoggle** teases **Jareth**,
- **Ludo** teases **Jareth**.