Psych 215L: Language Acquisition

Lecture 18 Complex Systems

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 $\frac{\text{English}}{\text{Subject}}$ Verb Object $\frac{\text{German}}{\text{Subject}}$ Verb Object $\frac{\text{German}}{\text{Subject}}$ Object $\frac{t_{\text{Verb}}}{t_{\text{Subject}}}$ Object $\frac{t_{\text{Verb}}}{t_{\text{Subject}}}$

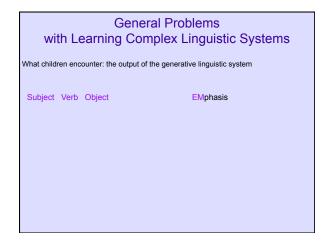
Complex Linguistic Systems

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

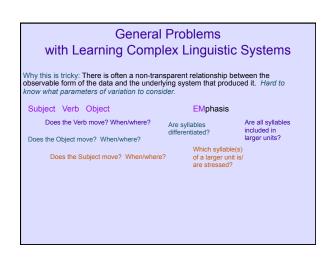
Observable data: stress contour

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 (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 must learn: the components of the system that combine to generate this observable output Subject Verb Object EMphasis Does the Verb move? When/where? Are syllables differentiated? Included in larger units? Does the Object move? When/where? Which syllables(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 Subject move? When/where?

Are syllables differentiated?

Does the Object move? When/where?

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

Are all syllables

included in larger units?

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

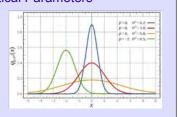
Statistical Parameters

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

- $\boldsymbol{\mu}$ for the mean

- σ for the standard deviation

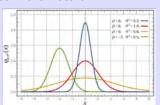




If we know the values of these parameters, we can make predictions about the likelihood of data we rarely or never see.

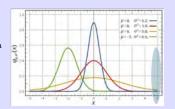
Statistical Parameters

Suppose this is a model of how many minutes late you'll be to class. Let's use the model with μ = 0, and σ^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 μ = 0, and σ^2 = 0.2. (blue

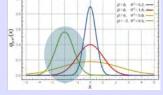


How likely are you to be 5 minutes late, given these parameters?

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 $\phi(X)$



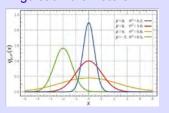
 $\varphi_{\mu,\sigma^2}(X) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(X-\mu)^2}{2\sigma^2}}$

Observing data points distributed like the green curve tells us that $\boldsymbol{\mu}$ 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 – $\phi(X)$ – as well as the associated parameters – $\mu_{\rm l}$ 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 domainspecific 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.

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 historicogeographic 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

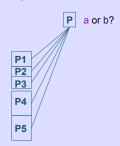
The fact that parameters connect to multiple structural properties then becomes a very good thing from the perspective of someone trying to acquire language. This is because a child can learn about that parameter's value by observing many different kinds of examples in the language.

"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

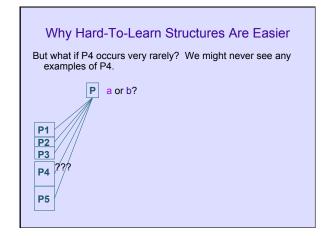
Let's assume a number of properties are all connected to parameter P, which can take one of two values: a or b.

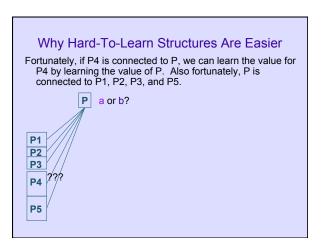


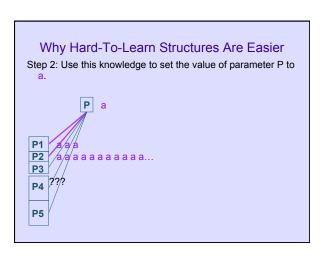
Why Hard-To-Learn Structures Are Easier

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

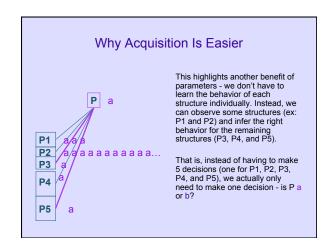








Why Hard-To-Learn Structures Are Easier Step 3: Since parameter P is connected to P4, we can predict that P4 will also show behavior a - even though we've never seen any examples of it! (We can also infer P3 and P5 the same way.) P a P1 P2 a a a a a a a a a a a a a ... P3 P4 P5



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

Subject Verb

Subject Verb

Jareth verrá Jareth will-come

Italian

"Jareth will come."

"Jareth will come."

grammatical

grammatical

One potential parameter

English Italian

*Verb Subject Verb Subject

Verb Subject

Verrá Jareth

*Will arrive Jareth

Will-arrive Jareth

"Jareth will arrive"

ungrammatical grammatical

One potential parameter

English Italian

*Verb Verb Verrá
Will come He-will-come

"He will come"

ungrammatical grammatical

One potential parameter

English Italian
Subject Verb

*Verb Subject

*Verb Subject

*Verb

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 Subject Verb

*Verb Subject

*Verb

Not okay to leave out expletive subject "it".

Requires special action for embedded subject questions.

Italian
Subject Verb

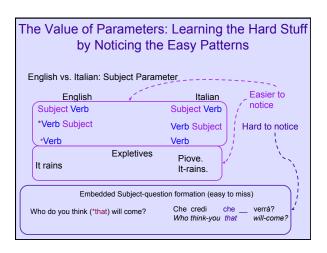
Verb Subject

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.



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

NP Verb

Postpositions: Noun Phrase Postposition

NP P Object 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 VP Verb NP Object

Prepositions: Preposition Noun Phrase

P NP
Preposition Object

Some other current thoughts on parameters

Baker 2008

"History has not been kind to the [Subject] Parameter as originally stated. It is now well-known that the cluster of properties that this parameter claimed to be related fragments in various ways when one looks at a wider range of languages, either dialects or languages from other families (Gilligan, 1987). Moreover, few new proposals for parameters with the same kind of scope as the [Subject] Parameter have gained currency since. As a result, thought about parameters has gone in two different directions."

"A large portion of the field has moved away from looking for classical parameters to looking for microparameters...tends to look for very localized differences with small but nontrivial effects in the grammars of very closely related languages (alias dialects)...has proven to be very fertile, creating a steady stream of results that shows no signs of slowing down."

Some other current thoughts on parameters

Baker 2008

Macro-parametric approach

"...perhaps the most important parameters are best found by a larger scale comparison, comparing languages from different language families. Whereas classical parametric theory compared French, Spanish, and Italian, and microparametric theory compares Trentino, Fiorentino, and Piedmontese, macroparametric theory might compare Mohawk and Mayali with Spanish and Swahili."

"I then present briefly two new parameters of a more or less macroparametric sort, discovered by comparing Bantu languages with Indo-European languages...My moral: that there is some reason to think that there are macroparameters out there hiding in a forest of microparameters, and the generative linguistic community should be trying harder to seek them out."

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

Complicated silent things

Sentences that have both an implied subject and implied object.

The girl is <u>eager</u> to see .

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

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

		يزا م		- ~	Ail.			h	_	~	
٠.	om	ЮII	orali	ea	SII	en	П	m	ш	01	į

Sentences that have both an implied subject and implied object.

The girl is \underline{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 <u>eager</u> 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 <u>easy</u> 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 <u>easy</u> 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 \underline{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?

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?

The girl. (*The girl* is the AGENT of the verb RUN.)

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?

The girl. (*The girl* is the AGENT of the verb RUN.)

Who is doing the seeming? Is it 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?

The girl. (*The girl* is the AGENT of the verb RUN.)

Who is doing the seeming?

Is it the girl? Maybe not, since we can say

"It seems that the girl is running." (expletive it)

SEEM is called a raising verb, since the subject (the girl) can "raise" to the main clause without changing the meaning of the sentence.

Raising vs. Control Verbs (Mitchener & Becker 2011)

Verbs that have specific syntactic behavior with specific semantic connotations.

The girl is trying to run .

Who is doing the running?

Raising vs. Control Verbs (Mitchener & Becker 2011)

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.

Raising vs. Control Verbs (Mitchener & Becker 2011)

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?

Raising vs. Control Verbs (Mitchener & Becker 2011)

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*).

Raising vs. Control Verbs (Mitchener & Becker 2011)

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.

Raising vs. Control Verbs (Mitchener & Becker 2011)

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 psychologically plausible learning algorithms could use this information to classify verbs.

Other difficult syntactic & semantic phenomena

Lidz 2010

Learning constituent structure

ancient history teacher

[[ancient history] teacher] vs. [ancient [history teacher]]

"I teach ancient history" "I am very old and I teach history"

Learning the interpretation of bare plurals

Crocodiles live in the swamp.

Generic: "It's generally true that crocodiles live in swamps."

Existential: "There are some crocodiles that live in that swamp over there."

Other difficult syntactic & semantic phenomena

Lidz 2010 + Viau & Lidz 2011

Ditransitive verbs

English John sent the book to Mary.

John sent Mary the book.

- nala
 Hari rashmi-go pustaka-vannu kalis-id-a
 Hari rashmi-DAT nois-ACC send-PST-ISM
 Hari rashmi-DAT nois-ACC send-PST-ISM
 Hari rashmi-go pustaka-vannu kalis-id-AKOTT-a
 Hari rashmi-go pustaka-vannu sahmi-ge kalis-id-a
 Hari pustaka-vannu rashmi-ge kalis-id-a
 Hari pustaka-vannu rashmi-ge kalis-id-ACC rashmi-DAT arad-PST-ISM
 Hari pustaka-vannu rashmi-ge kalis-id-ACC rashmi-DAT arad-PST-ISM
 Hari pustaka-vannu rashmi-ge kalis-id-ACC rashmi-DAT arad-PST-ISM
 Hari sent a book to Rashmi-

- Carmen envidel libro a su profesora.

 Carmen sent the book to her profesor

 Carmen le envidel libro a su profesora

 Carmen CL sent the book to her profesora
- Carmen envió a su profesora el libro.
 Carmen sent to her profesor the book
 Carmen be envió a su profesor el libro.
 Carmen CL sent to her profesor the book

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.

Pronouns

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

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,*

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,*

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

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.