

# GRS LX 865

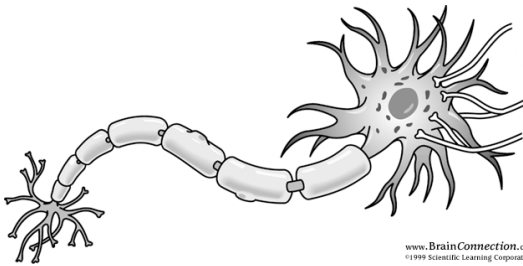
## Topics in Linguistics

Week 8. Neurons and impoverished stimuli

### Rules and brains

- (Generative) linguistics has traditionally been done in term of symbolic rules.
- $S \rightarrow NP VP$
- $V_{[past]} \rightarrow V + -ed / -d / -t$
- But people studying neurophysiology complain that there's no obvious way to "write" that in neurons.

### A neuron

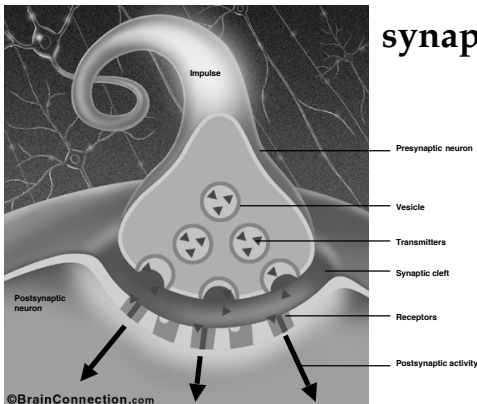


www.BrainConnection.com  
©1999 Scientific Learning Corporation

### Neural connections

- Individual neurons are connected to one another via excitatory and inhibitory connections, and has a certain level of activation. When a neuron's level of activation reaches a critical threshold, the neuron fires, spreading positive activation to other neurons that it is excitatorily connected to and negative activation to neurons that it is inhibitorily connected to.
- "Neurons that fire together wire together". Connections are developed or strengthened between neurons whose firings temporally coincide. Function has changed. Memory. It becomes likely now that if one fires the other will too. Long-term memory?

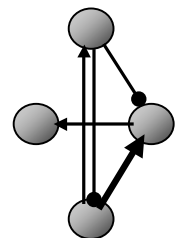
### synapse



©BrainConnection.com

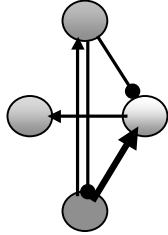
### Connectionism

- A connectionist system consists of a set of interconnected nodes ("neurons").
- Each connection has a certain strength and polarity.



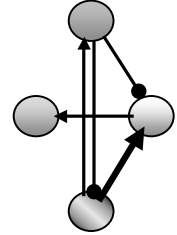
## Connectionism

- Each node has an *activation level* and a *threshold value*.



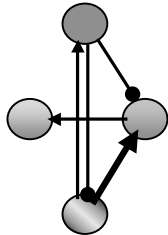
## Connectionism

- Each node has an *activation level* and a *threshold value*.
- When a node reaches the threshold level, it *fires*—and transfers its activation (additively or subtractively) along the connections.



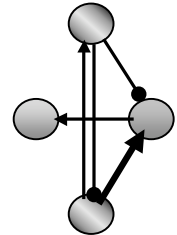
## Connectionism

- If this pushes a connected node over its threshold, *it* fires.



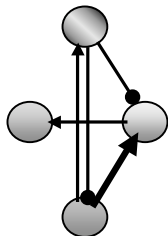
## Connectionism

- If this pushes a connected node over its threshold, *it* fires.



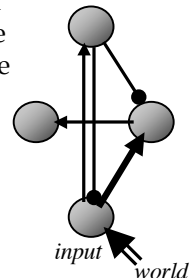
## Connectionism

- If this pushes a connected node over its threshold, *it* fires.
- And so forth...
- Clearly, the interactions can quickly become mind-bogglingly complex.



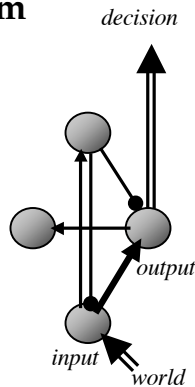
## Connectionism

- Certain nodes are designated as *input nodes*. These have an activation level driven by the perceptual system. So, maybe the node will be active if the currently perceive word starts with *t*...



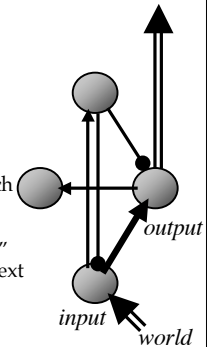
## Connectionism

- Other nodes are designated as *output nodes*. The status of these nodes determine the system's reaction to the input.
- So, it's a complex way to compute a *function* from input (patterns) to output (patterns).



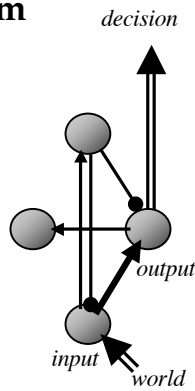
## Connectionism

- Finally, the *learning* aspect. The way neural nets are trained is generally:
  - Provide an input with a known "correct" output.
  - Check the output the system provides.
  - If the system's output doesn't match the correct output, adjust the connection weights in the network using a general "back-propagation" algorithm to make it come closer next time.



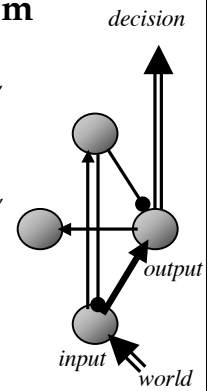
## Connectionism

- After a *lot* of training, the neural net can produce the appropriate outputs from the given inputs.
- The neural net has *abstracted out the systematicity* in the input data—but in ways that are almost always far too mathematically complicated to fathom.



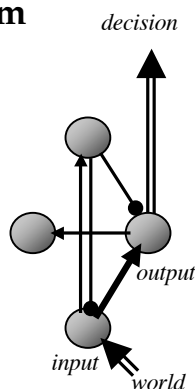
## Connectionism

- Then, when presented with novel inputs, the neural net will *generalize* its training to make decisions. This was previously considered to be a sure signal of following a rule.
- When trained on rule-governed material, it tends to "follow the rule" even with novel forms. Neural nets are also great *pattern recognizers*— they latch onto any kind of statistical regularity. Medical diagnosis, image reconstruction, ...



## Connectionism

- What's the point?
- The point is that neural nets can "learn" rule-like behavior from statistical regularity without being taught the rule (and in fact without there even *being* a rule). There are just neurons and connections (vaguely like the human brain).
- So maybe those rules were just *approximations*.



## Connectionism is too hard

- Connectionist research often has this property, looking at a very small problem on the boundary between grammatical knowledge and lexical memorization, where it is not at all clear that we could generalize the results to language as a whole grammatical system (or even come close to understanding what the network is even *doing*).
- At this point, connectionism is too hard—for a network large enough to do anything interesting and predictive, the generalizations it reaches will be completely inaccessible to us analyzing it from the outside.

## Rumelhart & McClelland

- R&M 1986 created a connectionist network to learn English past tense.
- English past tense forms come in *regular* and *irregular* kinds.
  - Walk/walked, kick/kicked
  - Tow/towed, rub/rubbed
  - Melt/melted, right/righted
- Break/broke, sing/sang, light/lit, grow/grew

## Regularities

- The *regular* kind are the ones that are easily described in terms of rules.
- $V_{[past]} \rightarrow V + -ed / -d / -t$ 
  - (depends on voicing and place of the last consonant of the stem).
- The irregular kind you have to memorize. The regular kind you can build on the fly.
- Claim of the connectionists is that they're all really the same kind of thing.

## Subregularities

- The thing is, the “irregulars” have some similarities as well, or at least can be grouped. (*sing/sang, ring/rang, sit/sat, ...*)
- This has the feel of the kind of thing an associationist network would be good at—seeing similarities, regularities and subregularities.

## He goed.

- One of the arguments in favor of rules in morphology is the “overregularization” that kids are seen to do.
  - For a while they seem get things right, including irregulars (“he went”)
  - But then they start saying things like “he goed” (overregularizing).
  - Eventually, they get it right again.
- Idea: In that middle step, they learned the rule. In the first step, they'd just memorized everything.

## RM

- The connectionist model (taking advantage of the fact that regulars are also overwhelmingly common) will also tend to over-regularize for a while. Like kids?
- There were various problems in the RM model, amply documented by Pinker & Prince (1988). Among them, the apparent overregularization RM saw seemed to be more a function of the *input* changing.

## Generalization

- Eventually RM's model could come up with the right answers for everything it was trained on. Then, the test is: What will it do with novel words? Will it do what people do?
- 72 new regular verbs.

## Generalization in RM

- 6: refused to answer
  - jump, pump, soak, warm, trail, glare
- 4: grossly bizarre (not a human-type mistake)
  - Squat/squakt, mail/membled, tour/toueder, mate/maded
- 7: double-marked
  - Type/typeded, snap/snappeded, smoke/smokeded
- ...This is a pretty poor simulation of adult knowledge.

## Dual mechanisms

- Pinker and various others have championed a version of morphology in which there are *both* rules and an associationist network.
- Rules are for regulars
- Network (memory) is for irregulars.

## Logical problem of language acquisition

- The grammar that people end up with is very complicated, and underdetermined by the data.
- The main argument for this (“poverty of the stimulus”) is that there are many generalizations that a kid *could* make on the basis of the input data that would be wrong, that would not result in a language that conforms to the principles that we’ve discovered seem to hold true of all adult languages.

## Language is really complicated

- 1) Frasier threw out Martin’s chair.
- 2) Frasier threw Martin’s chair out.
- 3) Daphne walked out the door.
- 4) \*Daphne walked the door out.
- 5) What did Roz say Niles bought?
- 6) What did Roz say that Niles bought?
- 7) Who did Roz say bought an espresso doppio?
- 8) \*Who did Roz say that bought an espresso doppio?

## Language is really complicated

- 9) His mother thinks Bill is a genius.
- 10) He thinks Bill is a genius, too.
- 11) Mary saw her.
- 12) Mary saw her duck.
- 13) I asked Mary to buy rum.
- 14) What did you ask Mary to buy ?
- 15) I saw the book about snakes on the table.
- 16) \*What did you see the book about on the table?

## Language is really complicated

- John ate.
- John ate a fish.
- John is too clever to catch.
- John is too clever to catch a fish.
- Who does Arnold wanna make breakfast for?
- \*Who does Arnold wanna make breakfast?
- Do you know what that’s doing up there?
- \*Do you know what that’s up there?

## Yet people know this stuff...

- Adult native speakers uniformly and overwhelmingly agree.
- To know English is to have knowledge of (how to determine) which sentences are possible and which are impossible in English.
- How one comes to have this knowledge is going to be our primary focus.

## Grammar

- People eventually end up with a *system* with which they can produce (and rate) sentences: a grammar.
- Even if a native speaker of English has never heard either of these sentences before, s/he knows which one is possible in English and which one isn't:
  - 15) Eight very adept sea lions played trombones.
  - 16) Eight sea lions very adept trombones played.

## How do people know this?

- Every native speaker of English knows these things.
- Nobody who speaks English as a first language was explicitly taught (growing up) "You can't question a subject in a complement embedded with *that*" or "You can't use a proper name if it's c-commanded by something coindexed with it."
- Trying to use any simple kind of general learning principle based on (analogy to) the sentences you get seems almost sure to lead you astray.

## That's the setup

- Language involves a complex grammar.
- Adults end up with knowledge of this grammar, quite uniformly.
- Children seem to go through advancing stages of language sophistication; they are learning, the end result being the adult language system.
- Next question: What is the nature of the children's learning?

## Linguists

- As linguists trying to figure out the grammatical system of a language, we...
  - Look at which sentences are grammatical
  - Look at which sentences are ungrammatical
  - Compare them to describe generalizations about what the crucial factors are differentiating the grammatical from the ungrammatical.
  - Check the predictions of the hypothesized generalization by looking at more complex sentences.
- Are kids just little linguists?

## Kids are not just little linguists.

- \*What did you see the book about on the table?
- \*Who did Mary say that bought coffee?
- Eight very adept sea lions played trombones.
- Linguists' theories: built by considering both grammatical *and ungrammatical* sentences (often of a fairly complex type).
- Kids: Don't hear ungrammatical sentences, nor even all of the grammatical sentences (often of a simpler type).

## So how do they do it?

- One hypothesis holds suggests that parents actually help kids along (though not consciously).
- It's well known that people seem to instinctively talk to little kids in kind of a weird way; exaggerated intonation, simpler words, more repetition. "Baby talk" or as it is sometimes known, "Motherese".
- Many have entertained the idea that this simpler, more carefully articulated, speech might guide kids along the path of language acquisition.

## Some properties of "Motherese"

- Slower speech, longer pauses
- Higher pitch, greater pitch range
- Exaggerated intonation and stress
- More varied loudness
- Fewer disfluencies
- More restricted vocabulary
- More rephrasings
- More repetitions
- Shorter, less complex utterances
- More imperatives and questions
- Fewer complex (multiclausal) sentences

## Does "Motherese" drive acquisition?

- Initially tempting, perhaps, but no.
- If "Motherese" were crucial for acquisition, it must be available to all language acquirers, universally.
- Several documented cultures don't even speak to the kids until they reach linguistic sophistication. (Of course, they're exposed to language in the environment, but not directed at them in "Motherese")

## Does "Motherese" drive acquisition?

- If you give a 4-month old the choice of whether to listen to "Motherese" or to normal adult-directed speech, the kid will choose to listen to "Motherese"...
- ...so it is quite likely that "Motherese" forms a significant part of the PLD for the kid, but it can't be *necessary* for successful language acquisition.

## Simpler isn't really better

- Linguists look to complex sentences to differentiate between predictions of different hypotheses about how the grammar works.
- Generally, prior to considering complex sentences, the data underdetermines the grammar; there are (at least) two systems compatible with the data observed so far.
- If linguists need to look to complex sentences to figure out the intricacies of the rules (which all adult native speakers seem to end up with), kids should need this information too.

## Positive and negative evidence

- Kids need to know the grammatical system by the time they are adults.
- Kids hear grammatical sentences (*positive evidence*)
- Kids are not told which sentences are ungrammatical (*no negative evidence*)
- Let's consider *no negative evidence* further...

## Negative evidence

- Negative evidence (information that a given sentence is ungrammatical) could come in various conceivable forms.
  - “The sentence *Bill a cookie ate* is not a sentence in English, Timmy. No sentence with SOV word order is.”
  - Upon hearing *Bill a cookie ate*, an adult might
    - Offer negative reinforcement
    - Not understand
    - Look pained
    - Rephrase the ungrammatical sentence grammatically

## Kids resist instruction...

McNeill (1966)

- Nobody don't like me.
- No, say 'nobody likes me.'
- Nobody don't like me.

[repeats eight times]

- No, now listen carefully; say 'nobody likes me.'
- Oh! Nobody don't likes me.

## Kids resist instruction...

Braime (1971)

- Want other one spoon, daddy.
- You mean, you want the other spoon.
- Yes, I want other one spoon, please Daddy.
- Can you say 'the other spoon'?
- Other...one...spoon
- Say 'other'
- Other
- 'Spoon'
- Spoon
- 'Other spoon'
- Other...spoon. Now give me other one spoon?

## Kids resist instruction...

Cazden (1972) (observation attributed to Jean Berko Gleason)

- My teacher holded the baby rabbits and we patted them.
- Did you say your teacher held the baby rabbits?
- Yes.
- What did you say she did?
- She holded the baby rabbits and we patted them.
- Did you say she held them tightly?
- No, she holded them loosely.

## Negative evidence via feedback?

- Do kids get “implicit” negative evidence?
  - Do adults understand grammatical sentences and not understand ungrammatical ones?
  - Do adults respond positively to grammatical sentences and negatively to ungrammatical ones?

## Approval or comprehension?

Brown & Hanlon (1970):

- Adults understood 42% of the grammatical sentences.
- Adults understood 47% of the ungrammatical ones.
- Adults expressed approval after 45% of the grammatical sentences.
- Adults expressed approval after 45% of the ungrammatical sentences.

This doesn't bode well for comprehension or approval as a source of negative evidence for kids.

## Kids' experience differs

- Parents respond differently
  - Eve & Sarah's parents ask clarification questions after ill-formed *wh*-questions.
  - Adam's parents ask clarification after well-formed *wh*-questions...and after past tense errors.
- How can kids figure out what correlates with grammaticality in *their* situation?

## Kids' experience differs

- Piedmont Carolinas: Heath (1983):  
Trackton adults do not see babies or young children as suitable partners for regular conversation...[U]nless they wish to issue a warning, give a command, provide a recommendation, or engage the child in a teasing exchange, adults rarely address speech specifically to young children.

## Feedback *disappears*

- Adam and Sarah showed almost no reply contingencies after age 4...
- But they still *made* errors after age 4
- And they still stopped making those errors as adults (*learning* didn't cease).

## Three possible types of feedback

- **Complete:** consistent response, indicates unambiguously "grammatical" or "ungrammatical."
- **Partial:** if there is a response, it indicates "grammatical" or "ungrammatical"
- **Noisy:** response given to both grammatical and ungrammatical sentences, but with different/detectible *frequency*.

## Statistics (from Marcus 1993)

Suppose response R occurs 20% of the time for ungrammatical sentences, 12% of the time for grammatical sentences.

Kid gets response R to utterance U, there's a 63% chance (20/32) that U is ungrammatical.  
Guess: ungrammatical, but 38% chance of being wrong.

Kid *doesn't* get response R, 52% chance (88/168) it's grammatical. Guess: grammatical, but 48% chance of being wrong.

## Statistics (from Marcus 1993)

Suppose response R occurs 20% of the time for ungrammatical sentences, 12% of the time for grammatical sentences.

Suppose kid got response R to U, and is 63% confident that U is ungrammatical—ok, but nowhere near good enough to build a grammar.

This is a serious task, a kid's going to want to be *sure*.  
Suppose kid is aiming for 99% confidence (adults make at most 1% speech errors of the relevant kind—pretend this reflects 99% confidence).

## Lacking confidence

- Based on R (20%-12% differential), they'd have to repeat U 446 times (and compile feedback results) to reach a 99% confidence level.
- Based on various studies on noisy feedback, a realistic range might be from 85 times (for a 35%-14% differential) to 679 times (for a 11.3%-6.3% differential).
- This sounds rather unlike what actually happens.

## In a way, it's moot anyway...

- One of the striking things about child language is how few errors they actually make.
- For negative feedback to work, the kids have to make the errors (so that it can get the negative response).
- But they don't make enough relevant kinds of errors to determine the complex grammar.

## Yes-no questions

- 17) The man is here.  
18) Is the man here?

**Hypothesis 1:** Move the first *is* (or modal, auxiliary) to the front.

**Hypothesis 2:** Move the first *is* after the subject noun phrase to the front.

- 19) The man who is here is eating dinner.

## Yes-no questions

- 19) The man who is here is eating dinner.  
20) \*Is the man who here is eating dinner? (\*H1)  
21) Is the man who is here eating dinner? (√H2)

No kid's ever said (20) to mean (21), which would have been necessary to distinguish hypotheses 1 and 2... Why not?

It seems that kids don't even *entertain* Hypothesis 1. And that's fine, because it seems like Hypothesis 1 is a kind of rule not found in *any* adult language.

## Abstract principles

- **Principle C:** Nothing coreferential can c-command a proper name.  
  
\*He<sub>i</sub> believes John<sub>i</sub>'s teacher.  
His<sub>i</sub> teacher believes John<sub>i</sub>.
- Study of adult grammar reveals that **c-command** is the appropriate abstract notion, defined on syntactic structures. But how do kids learn about c-command? You can't *hear* c-command.
- What's more, study of adult grammar reveals that Principle C holds in every language!

## So, we've got...

- Kids don't make as many mistakes as would be needed for hypothesis testing.
- Kids seem to receive no relevant negative evidence while learning language anyway.
- Kids learn *fast*.
- Kids become adults with all the grammatical knowledge pertaining thereto (uniform, highly complex)
- Kids come to know abstract principles (like Principle C) without access to evidence determining them. In many cases, these principles are observed in all human languages. "Poverty of the stimulus"

## Having language = being human

- A linguistic capacity is part of being human.
- Like having two arms, ten fingers, a vision system, humans have a language faculty.
- Specification of having arms instead of wings, etc., is somehow encoded genetically.
- Structure of the language faculty is predetermined, like the structure of the vision system is.
- The language faculty (tightly) constrains what kinds of languages a child can learn.
- =“Universal Grammar” (UG).

## Universal Grammar

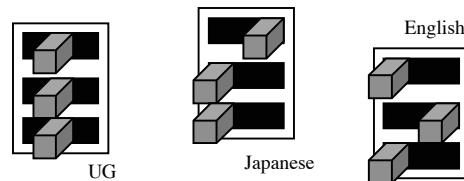
- UG tightly constrains the learning process.
- Study of Syntax, phonology, etc., is generally trying to uncover properties of Language, to specify what *kind* of languages a child can learn, to see what kinds of restrictions UG places on language.
- But kids don't just enter the world speaking like adults—there's *development*.
- And, adults don't all end up speaking the same language—there is *learning*.

## Learnability

- The Principles & Parameters model is designed to address the learnability problem we faced:
  - Languages are very complex.
  - Languages differ (*something* has to be learned).
  - Children get insufficient and variable evidence to deduce the uniform rules of grammar they end up with.
  - Children have adult-like grammars relatively quickly.

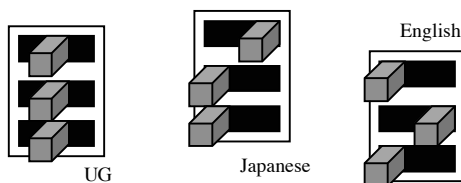
## Principles and Parameters

- The proposed solution to the apparent paradox is to suppose that *to a large extent all human languages are the same*. The grammatical systems obey the same principles in all human languages.



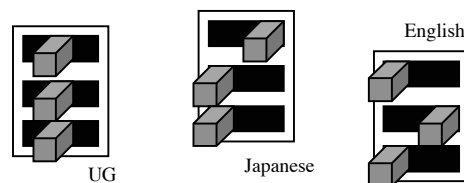
## Principles and Parameters

- Languages differ, but only in highly limited ways.
  - In the order between the verb and the object.
  - In whether the verb raises to tense
  - ...



## Principles and Parameters

- This reduces the task for the child immensely—all that the kid needs to do is to determine from the input which setting each of the parameters needs to have for the language in his/her environment.



## The standard picture

- The way this is usually drawn schematically is like this. The Primary Linguistic Data (PLD) serves as input to a Language Acquisition Device (LAD), which makes use of this information to produce a grammar of the language being learned.



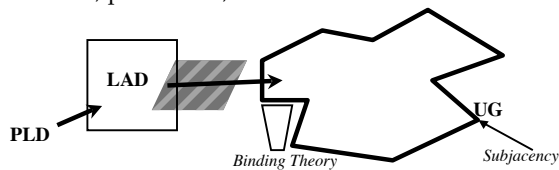
## The standard picture

- This isolates the innately specified language faculty into a single component in the picture. The LAD contains (a specification for) all of the principles and the parameters, and has a procedure for going from PLD to parameter settings.



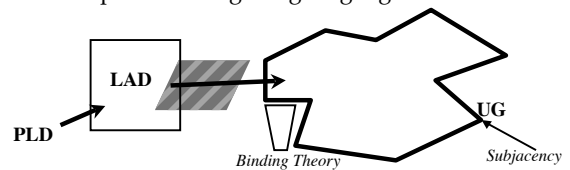
## Modeling human language capacity

- We may be able to avoid confusion later, though, if we differentiate the innately provided system into its conceptual components.
- This is my rendition of a way to think about UG, parameters, and LAD.



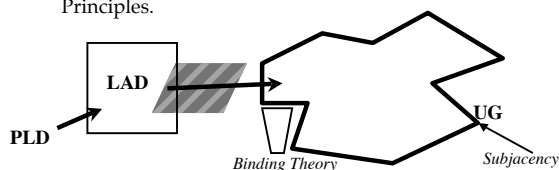
## Modeling human language capacity

- UG provides the parameters and contains the grammatical system (including the principles, like Subjacency, Binding Theory, etc.) that makes use of them.
- LAD sets the parameters based on the PLD. Responsible for getting language to kids.



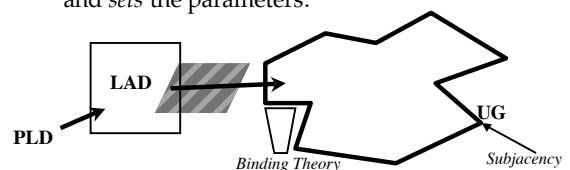
## Modeling human language capacity

- The idea behind this diagram is that UG is something like the shape of language knowledge.
  - Knowledge of language can only take a certain, innately pre-specified "shape".
  - A system with this "shape" has certain properties, among them Binding Theory, Subjacency, ... the Principles.



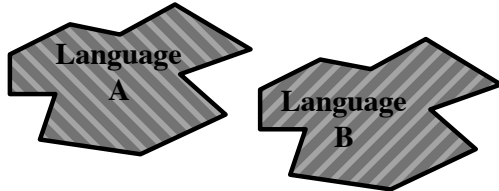
## Modeling human language capacity

- The Parameters are different ways in which stored knowledge can conform to the "shape" of UG.
- The LAD is a system which analyzes the PLD and sets the parameters.



## Principles and Parameters

- So two languages which differ with respect to one parameter setting might be represented kind of like this.
  - This is of course a cartoon view of things, but perhaps it might be useful later.



## Principles and Parameters

- So what *are* the Principles and Parameters?
- **Good question!** —and that’s what theoretical linguistics is all about.
- Since 1981, many principles and parameters have been proposed. As our understanding of language grows, new evidence comes to light, and previous proposals are discarded in favor of better motivated ones. It’s hard to keep a current tally of “the principles we know of” because of the active nature of the field.

## Principles and Parameters

- Some of the (proposed) Parameters that have received a fair amount of press are:
  - Bounding nodes for Subjacency
  - Binding domain for anaphors and pronouns
  - Verb-object order
  - Overt verb movement (V moves to tense)
  - Allowability of null subject (*pro*) in tensed clauses
- We’ll look at each of them in due course...

## Verb-object order

The parameter for verb-object order (more generally, the “head parameter” setting out the order between X’-theoretic head and complement) comes out as:

- Japanese: Head-final (X follows complement)
- English: Head-initial (X precedes complement).

Figuring out which type the target language is is often fairly straightforward. Kids can hear evidence for this quite easily. (Not *trivial*, though—consider German SOV-V2)

## Principle A

22) Sam believes [that Harry overestimates himself]



23) Sam-wa [Harry-ga zibun-o tunet-ta to] it-ta  
 Sam-top Harry-nom self-acc pinch-past-that say-past  
 ‘Sam said that Harry pinched him(self).’



## Principle A

- **Principle A.** A reflexive pronoun must have a higher antecedent in its binding domain.
- *Parameter: Binding Domain*
  - Option (a): domain = smallest clause containing the reflexive pronoun
  - Option (b): domain = utterance containing the reflexive pronoun

## But how can you set this parameter?

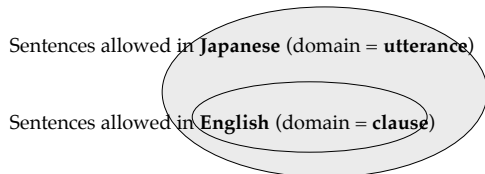
- Every sentence a kid learning English hears is consistent with *both* values of the parameter!
- If a kid learning English decided to opt for the “utterance” version of the domain parameter, nothing would ever tell the kid s/he had made a mistake.
- S/he would end up with non-English intuitions.

## But how can you set this parameter?

- A kid learning Japanese can tell right away that their domain is the sentence, since they’ll hear sentences where *zibun* refers to an antecedent outside the clause.

## But how can you set this parameter?

- The set of sentences allowed in English is a *subset* of the set of sentences allowed in Japanese. If you started assuming the English value, you could learn the Japanese value, but not vice-versa.



## Subset principle/defaults

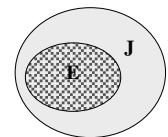
- Leads to: The acquisition device selects the most restrictive parametric value consistent with experience. (*Subset principle*)
- That is, for the Principle A domain parameter, you (a LAD) *start* assuming you’re learning English and switch to Japanese only if presented with evidence.

## What it takes to set a parameter

- 
- Binding domain parameter
    - Option (a): Binding domain is clause.
    - Option (b): Binding domain is utterance.
  - *English* = option a, *Japanese* = option b.

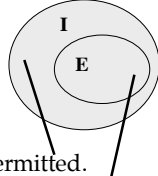
## What it takes to set a parameter

- Binding domain parameter
  - Kids should start under the assumption that the parameter has the English setting.
  - If they hear only English sentences, they will stick with that setting.
  - If they hear Japanese sentences, they will have evidence to move to the Japanese setting.



## What it takes to set a parameter

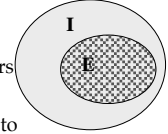
Very sensible. Now, let's consider another parameter of variation across languages.



- Null subject parameter
  - Option (a): Null subjects are permitted.
  - Option (b): Null subjects are not permitted.
- *Italian* = option a, *English* = option b.

## What it takes to set a parameter

- The Subset principle says that kids should start with the English setting and *learn* Italian if the evidence appears



- But even *English* kids are well-known to drop subjects early on in acquisition. As if had the Italian setting for this parameter.

## Moreover...

- English kids hear *looks good* and *seems ok* and *stop that right now*. Why don't they end up speaking Italian? If they mis-set the parameter, how could they ever recover?
- Italian kids hear subjectless sentences—why don't they interpret them as imperatives or fragments (so as not to have to change the parameter from the default)?

## Triggers

- It seems like *actual occurrence of null subjects* isn't a very good clue as to whether a subject is a null subject language or not.
- Are there better clues? If a strapping young LAD were trying to set the null subject parameter, what should it look for?

## Triggers

- Turns out: Only true subject-drop languages allow null subjects in tensed embedded clauses.
- 24) \*John knows that [— must go]. (English)
- 25) Juan sabe que [— debe ir]. (Spanish)  
'Juan knows that [he] must go.'
- Perhaps the LAD "knows" this and looks for exactly this evidence. Null subjects in embedded tensed clauses would be a *trigger* for the (positive setting of the) null subject parameter.

## Triggers

- A potential problem with the proposed subject-drop trigger is that it requires *complex* sentences—you need to look at an embedded sentence to check for the trigger.
- Such sentences might be too complicated for kids to process.
- **Degree-1 learnability:** Triggers need look no lower than 1 level of embedding.
- **Degree-0 learnability:** Triggers need look only at main clauses.

## Triggers

- Many who work on learnability have adopted the hypothesis that triggers need to be degree-0 learnable.

*IP and TP are often used interchangeably*

- **Subjacency.**  $*[wh [\alpha \dots [\beta \dots t \dots ]]]$  where  $\alpha$  and  $\beta$  are bounding nodes.

Bounding node parameter for IP:

- Option (a): IP is a bounding node (English).
- Option (b): IP is **not** a bounding node (French, Italian).

## Triggers

- Thus, a kid learning French couldn't choose option (b) by hearing this...

28) *Viola* un liste de gens... 'there is a list of people...'

[à qui on n'a pas encore trouvé [quoi envoyer *t t*]]  
to whom one has not yet found [what to send]]

- ...since that's a degree-2 trigger. But...

## Triggers



29) *Combien as-* [IP tu vu [NP t de personnes]]?  
How-many have you seen of people  
'How many people did you see?'

- If IP were a bounding node, this should be ungrammatical in French, so this can serve as (degree-0) evidence for option (b).

## Triggers

- Principles are part of UG
- Parameters are defined by UG
- Triggers for parameter settings are defined as part of the LAD.

## Navigating grammar spaces

- Regardless of the technical details, the idea is that in the space of possible grammars, there is a restricted set that correspond to possible *human* grammars.
- Kids must in some sense navigate that space until they reach the grammar that they're hearing in the input data.

## Learnability

- So how do they do it?
- Where do they start?
- What kind of evidence do they need?
- How much evidence do they need?
- Research on *learnability* in language acquisition has concentrated on these issues.

## Are we there yet?

- There are a lot of grammars to choose from, even if UG limits them to some finite number.
- Kids have to try out many different grammars to see how well they fit what they're hearing.
- We don't want to require that kids remember everything they've ever heard, and sit there and test their current grammar against the whole corpus of utterances—that's a lot to remember.

## Are we there yet?

- We also want the kid, when they get to the right grammar, to stay there.
- Error-driven learning
  - Most theories of learnability rely on a kind of error-detection.
  - The kid hears something, it's not generable by their grammar, so they have to switch their hypothesis, to move to a new grammar.

## Plasticity

- Yet, particularly as the navigation progresses, we want them to be zeroing in on the right grammar.
- Finding an error doesn't mean that you (as a kid) should jump to some random other grammar in the space.
- Generally, you want to move to a nearby grammar that improves your ability to generate the utterance you heard—move in baby steps.

## Triggers

- Gibson & Wexler (1994) looked at learning word order in terms of three parameters (head, spec, V2).
- Their triggering learning algorithm says if you hear something you can't produce, try switching one parameter and see if it helps. If so, that's your new grammar. Otherwise, stick with the old grammar and hope you'll get a better example.

## Local maxima

- A problem they encountered is that there are certain places in the grammar space where you end up more than one switch away from a grammar that will produce what you hear.
- This is locally as good as it gets—nothing next to it in the grammar space is better—yet if you consider the whole grammar space, there is a better fit somewhere else, you just can't get there with baby steps.

## Local maxima

- This is a point where any move you make is worse, so a conservative algorithm will never get you to the best place. Something a working learning algorithm needs to avoid. (And kids, after all, make it).

