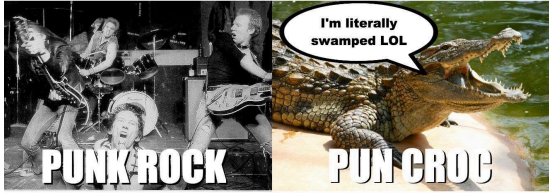


## LEARN THE DIFFERENCE



IT COULD SAVE YOUR LIFE

---

## Who are you?

---

### Information I'd like

- Year
  - Major
  - What languages do we speak?
- 

## What is Language?

Matt Zaslansky - LIGN 101

---

### Today's Plan

- Who are you?
  - What is Language?
  - What are some characteristics of Language?
  - How Universal is Language?
  - What *isn't* language?
  - Does our language shape our thoughts?
- 

## What is Language?

---

## Language

- Language is a system for linking expression to meaning
  - ... but we're probably better off thinking about what makes language, language-like
- 

# What characteristics does Language have?

---

## Three important characteristics of Language

- 1 - Arbitrary relationships between signs and the signified
- 2 - Fully Productive and Creative
- 3 - Speakers have knowledge of grammar and the 'rules' of language

|      |
|------|
| ###  |
| Char |
| #1   |
| ##   |
| rela |
| bet  |
| sign |
| sign |
| ◀ ▶  |

## Language has 'signs' and 'signified' concepts

- The 'sign' is something a human does in communication
  - The 'signified' is something in the world (conceptual or concrete)
  - *What is the link between these things?*
- 

## Non-arbitrary signs

- *The sign is causally linked to the signified*
- Smoke is a sign of fire

- Bleeding from your finger is a sign of a papercut
  - Will's presence is a sign of bad puns
- 

## Iconic (or representational) signs

- *A referential abstraction from the signified*



- Some hand gestures
  - Some sounds meant to 'mimic' the world
- 



## Arbitrary Signs

- *Referential only by agreement and widespread knowledge*
  - Non-representational symbols (or uses of symbols) (h, 🍆, 🍊)
  - Non-iconic gestures
- 



- "Go whatever speed is legal here"
-

🔥 in 'omg he started a 🔥 in his trashcan' is ...

A. Non-Arbitrary, Iconic

B. Arbitrary

---

🔥 in 'omg he started a 🔥 in his trashcan' is ...

A. **Non-Arbitrary, Iconic**

B. Arbitrary

---

🔥 in 'omg that party was 🔥 🔥 🔥 ' is ...

A. Non-Arbitrary, Iconic

B. Arbitrary

---

🔥 in 'omg that party was 🔥 🔥 🔥 ' is ...

A. Non-Arbitrary, Iconic

B. **Arbitrary**

---

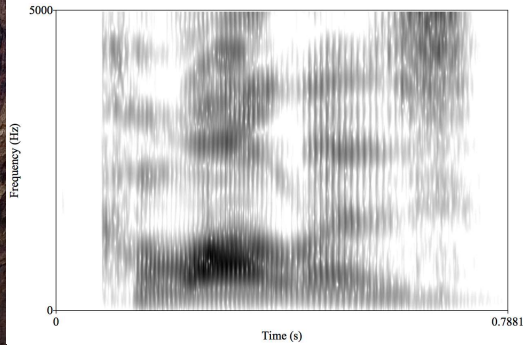
**The most important arbitrary signifiers are...**

---

## **Words!**

- The connection between words and meanings in language is (mostly) arbitrary
-

## The connection between sound and meaning is mostly arbitrary in spoken Language



---

## Signed languages have arbitrary signs too!

(These ASL examples are from Sign with Robert on GIPHY (<https://giphy.com/signwithrobert/>))



---

## Signed languages have arbitrary signs too!



**Signed languages have arbitrary signs too!**



**... but there are often elements of iconicity in signs**



**... but there are often elements of iconicity in signs**



**... but there are often elements of iconicity in signs**



## Spoken language has moments of iconicity too!

---

### Onomatopoeia

*Words which, when spoken, sound kind of like the things they represent*

- (See also the idea of an Ideophone (<http://en.wikipedia.org/wiki/Ideophone>))
- 

### Onomatopoetic Words in English

Boom

Gurgle

Woof

Snip

Whoosh

Splash

---

### Onomatopoeia differs across languages

---

#### A rooster says...

cock-a-doodle-doo (English)

kukko kiekuu (Finnish)

chicchirichí (Italian)

kuklooku (Urdu)

kukuriku (Hungarian)

Sourced from this awesome site

(<http://www.eleceng.adelaide.edu.au/personal/dabbott/animal.html>)

---

(So, different languages can't even map sounds onto sounds uniformly)

---

## The best evidence for arbitrariness: Words differ across languages

- If the link between signifier and signified were always non-arbitrary, *everybody would have substantially similar words for concepts*
- 

### Spoiler Alert: We do not



---

### Characteristic #1

## Arbitrary relationships between signs and the signified

---

### Characteristic #2

## Languages are fully productive and creative

---

### Let's test that out

- Let's describe an image that you've never seen before
-





---

## **This was not created by a human, so it doesn't follow human categories**

- This was generated by Stable Diffusion ([https://en.wikipedia.org/wiki/Stable\\_Diffusion](https://en.wikipedia.org/wiki/Stable_Diffusion))
  - The prompt was 'beach san diego kitten spaceship silverware motorcycle dinosaur by Thomas Kinkade'
  - This is one of the weirdest natural language processing examples around!
    - **Yet we can still describe the output!**
- 

## **Languages are capable of describing anything (eventually)**

- Might take time
  - Might require some words to be 'circumlocuted' or invented
    - "I don't know a word for this, but I'll use other words to describe what I'm after"
  - Might require incredible skill
    - "They should have sent a poet"
  - But they'll get the job done
- 

## **Languages are also capable of producing an infinite set of unique sentences**

- Any human speaker of a language can produce a sentence that nobody has ever spoken before
  - "I plan to buy a half-empty barrel of Doritos as soon as my penguin gets home with the Crown Jewels."

- “I sure am glad that Michael Bay has agreed to direct a movie adaptation of my favorite childhood cartoon”
- 

## **... and sentences can be infinitely grown and expanded**

- I saw Anna.
  - I saw Anna yesterday.
  - I saw Anna yesterday before I refilled the hummingbird feeder.
  - I saw Anna yesterday afternoon right after writing a bad pun but before I refilled the hummingbird feeder.
  - I saw my neighbor Anna who likes to hug polar bears yesterday afternoon right after writing a bad pun but before I refilled the hummingbird feeder.
- 

## **This particular property is called recursion**

- Sentences can be made to contain elements from other sentences infinitely
  - “My cat saw the guy who went with Kati, Vincent, and Victoria to the park where the New York Knicks talked to Mary...”
  - This can go on literally *forever*
- 



## **Creativity and Productivity is true of all languages**

- We have yet to discover a language which is communicatively 'primitive' or unable to describe reality
  - If a language lacks elements, they will be developed or borrowed over time
  - Languages can differ substantially still
  - A given language may lack words for some concepts
    - ... or may have words that other languages don't have
    - Many words and concepts are culturally based, and may not be shared regardless
  - **Every natural human language that we've found is able to express a rich set of meanings**
- 

## **Characteristic #2**

### **Languages are fully productive and creative**

---

## **Characteristic #3**

### **Speakers have knowledge of grammar and the 'rules' of language**

---

#### **Grammaticality Judgements**

- Some things sound 'like English' and are comprehensible
  - Some things are not English
  - These are also known as 'language intuitions' or 'linguistic intuitions'
- 

#### **You can make grammaticality judgements**

- 'John went to the park with Tamara'
- 'Walk Rick gate Johnny between'

- ‘The moose fiendishly decalcified the hieroglyphics between the Dagestani Giraffes’
  - ‘Somebody set up us the bomb. All your base are belong to us.’
- 

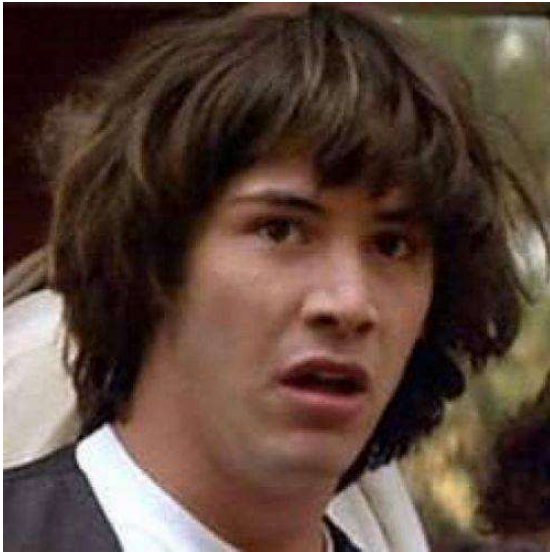
## Grammaticality judgements are about language, not individual words

- ‘Gleeble’ - A small fish
- ‘To Garflabble’ - To perform an elaborate mating display
- ‘Flundubble’ - A large octopus-like creature

|       |
|-------|
| ‘Gle  |
| sm    |
| ‘Gle  |
| sm    |
| ‘To   |
| Gar   |
| To p  |
| an    |
| ma    |
| dis   |
| ‘Flu  |
| A la  |
| oct   |
| cre   |
| ###   |
| sen   |
| ‘Gle  |
| gar   |
| flun  |
| all t |
| gra   |
| A) \  |
| B) I  |
| ◀ ▶   |

**(and by the way, how do we know the plural of 'gleeble'?)**

---



- (Nah, it's just phonology)
- 

**Is the following sequence grammatical?**

---

**How about this one?**

---

**Judgements can be more subtle**

- ?There looks to be a problem outside
  - ?There sounds to be somebody in there
  - ?There smells to be something rotten in there
  - ?There feels to be something in the bag
  - ?There tastes to be an issue with the cookies
- 

**We as speakers have strong feelings about the 'correctness' and 'rules' of language**

- ... but not so much for other kinds of basic communicative gestures
  - So, this is yet another element of Language
-

## Characteristic #3

# Speakers have knowledge of grammar and the 'rules' of language

---

## There are other characteristics that are important in human Language

- Language is symmetrical (e.g. we can send and receive)
- Language can talk about situations not in the 'here and now'
- Language is made of discrete, isolatable units (like words or sounds)
- Language has slang, taboo words
- Language accomplishes certain functions (e.g. asking questions, issuing commands)
- All languages can be translated into any other
- But the three we'll focus on are...

###

imp  
cha  
of L

- 1  
rela  
bet  
sig  
sig

- 2  
Pro  
and

- 3  
hav  
kno  
gra  
the  
lan

- All  
lan  
sha  
pro

## Given that human languages seem to share some properties

- ... and we're yet to find a group of humans without language
  - Some theorists have argued that this is evidence for...
- 

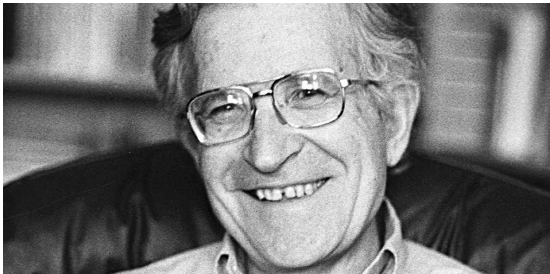
# Universal Grammar

---

## Universal Grammar

- An idea initially posited by Noam Chomsky
- 

## A Note on the Noam



## Universal Grammar

An idea initially posited by Noam Chomsky

- UG holds that there's some genetic basis to human language
  - Some properties of human language are innate
    - Not just the ones to do with physical structures
  - There are a set of constraints governing language that are 'built in' to our physiology
-

## UG is not facts about your individual language

- A child of two Mandarin Chinese speakers, raised by English speakers, will speak English
  - Every language is *not* the same, and linguistic grammar varies substantially
  - *UG is about biological predisposition to Language, not to A language*
- 

## Evidence for UG

- No cultures, people, or societies have been found with no language
  - No other species has been found with human-like language
  - Children acquire language impressively well
    - Even in cultures where people don't speak to children
  - Signed languages show similar properties to spoken languages
    - Even when they arise in isolation
  - Some things appear to be universal
    - The big one is Recursion
- 

## Evidence against UG

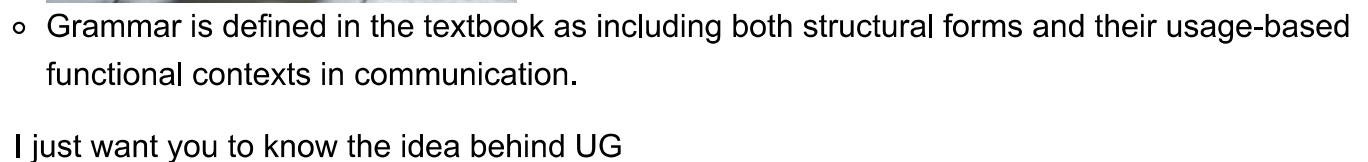
- There doesn't seem to be good biological evidence for it
  - Many properties of human language can be derived from functional facts about communication
  - Recursion might not be universal
    - Dan Everett (and others) claim that speakers of Pirahã lack recursion
    - Pirahã is a pretty neat language
  - Even Chomsky doesn't support this view anymore (<https://dlc.hypotheses.org/1269>)
- 

## Are elements of Linguistic structure innate?

- The predominant view among professional linguists for a long time was aligned with Chomsky in being a nativist approach, meaning that most people in the field assumed that our biological capacity for language is very specifically linguistic
  - Nativist/innatist arguments have gone from 'unquestioned' to 'questionable' in the field



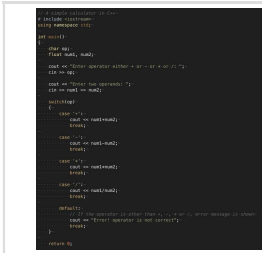
- Edited by Carol Genetti



## What *isn't* Language

- Semaphore (<https://www.omniglot.com/writing/semaphore.htm>), and Morse Code
  - These are just re-coding an existing language
- Writing systems
  - They're closely associated with an existing spoken or signed human language

- Languages like COBOL, C++, Java, Python, Rust



# Are programming languages languages?

- 1 - Arbitrary relationships between signs and the signified
- 2 - Fully Productive and Creative
- 3 - Speakers have knowledge of grammar and the ‘rules’ of language

# Are programming languages languages?

- 1 - Arbitrary relationships between signs and the signified
- 2 - Fully Productive and Creative
- 3 - Speakers have knowledge of grammar and the ‘rules’ of language

###

pro

lan

lan

1 -

rela

bet

sig

the

2 -

Pro

and

3 -  
hav  
knc  
gra  
the  
lan

## Are programming languages languages?

1 - Arbitrary relationships between signs and the signified

2 - Fully Productive and Creative

3 - Speakers have knowledge of grammar and the 'rules' of language

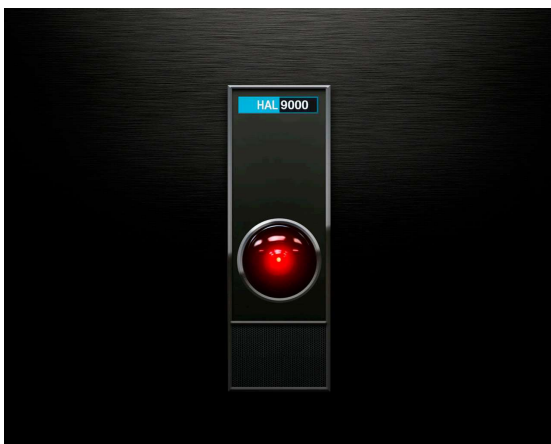
---

So, no.

- (Current) Computer Programming languages are not Language

---

(Sorry, Hal)



## Is 'body language' language?

---

- <https://www.youtube.com/shorts/JCa3JLAu5-s> (<https://www.youtube.com/shorts/JCa3JLAu5-s>)

---

## What is Princess Emily communicating non-verbally?

- A. A - Curiosity
  - B. B - Happiness
  - C. C - Disgust
  - D. D - Sadness
- 

## What is Princess Emily communicating non-verbally?

- A. A - Curiosity
  - B. B - Happiness
  - C. **C - Disgust**
  - D. D - Sadness
- 

|       |
|-------|
| ###   |
| lan   |
| lan   |
| 1 -   |
| rela  |
| bet   |
| sig   |
| the   |
| 2 -   |
| Pro   |
| and   |
| 3 -   |
| hav   |
| kno   |
| gra   |
| the   |
| lan   |
| ◀   ▶ |

|  |
|--|
| ###<br>lan<br>lan                      |
| 1 -<br>rela<br>bet<br>sig<br>the       |
| 2 -<br>Pro<br>and                      |
| 3 -<br>hav<br>kno<br>gra<br>the<br>lan |
| ◀   ▶                                  |

|                                  |
|----------------------------------|
| ###<br>lan<br>lan                |
| 1 -<br>rela<br>bet<br>sig<br>the |
| 2 -<br>Pro<br>and                |

3 -  
hav  
knc  
gra  
the  
lan

(Maybe? But it's not grammar in the way natural languages exhibit!)


## Do any non-humans have language?

---

### Yes, but only one so far

---

#### Large Language Models can produce human language

- Large Transformer-based Neural language models (like ChatGPT or OpenAssistant) are now capable of producing robust, grammatical, and used-as-humans-would language
  - They have 'passed the bar' for many linguists as competent language users
    - Most of the historical problems language models faced are now 'solved'
    - *This is different than saying 'These are true AI'*
  -  They are much less efficient language learners than children  
(<https://arxiv.org/abs/2208.07998>)
    - ... and work is ongoing to see whether they can do more with less
- 

#### Aside: ChatGPT has no plausible 'universal grammar'

- The code doesn't include grammatical facts, a 'language organ', or innate grammatical patterns
- English models have accidentally learned some other languages
  - (e.g. it appears to be able to use one languages' patterns to generalize)

- Some UG Folks are still fighting (<https://www.nytimes.com/2023/03/08/opinion/noam-chomsky-chatgpt-ai.html>). (also [this](https://mronline.org/2023/04/24/chatgpt-and-human-intelligence-noam-chomsky-responds-to-critics/) (<https://mronline.org/2023/04/24/chatgpt-and-human-intelligence-noam-chomsky-responds-to-critics/>))
- 

## **There is finally another kind of thing which can do human language**

- ... and it's a very large matrix of numbers.
    - Weird.
- 

## **What about living creatures?**

---

### **Lots of animals communicate**

- Mating calls
  - Alert calls
  - Territorial displays
- 

### **Mimicry**

"I can talk, can you fly?"; A...



## Nim Chimpsky



---

## Nim Chimpsky

- A chimpanzee raised in a human family to learn ASL by Herbert Terrace
- Learned 125 signs (maybe fewer)
- Combined signs to produce 2-4 word utterances

---

## Nim Chimpsky didn't seem to have full-on language

- Never got past 2-4 sign utterances
  - 'Nim Banana Me Eat'
- Only 7% of utterances served to advance the conversation
- 39% of utterances were repetitions of what had just been signed
- Entirely 'pragmatic' use of language
- Terrace argued that Nim never learned anything
- There were some methodological... oddities

---

## Nim led a pretty sad life

- Became rather strong and violent, as male chimps do.



- Later given back to a research facility and used for medical research
- Still used signs (e.g. 'play', hug') from time to time
- Died at an animal rescue ranch of a heart attack
- Featured in the (sad) documentary Project Nim (<https://www.imdb.com/title/tt1814836/>).

## **The Honeybee 'Waggle Dance'**

<https://www.youtube.com/watch?v=bFDGPgXtK-U> (<https://www.youtube.com/watch?v=bFDGPgXtK-U>)

---

## **Bat Vocalizations**

- According to Prat et al. 2016 (<https://www.nature.com/articles/srep39419>), bat vocalizations have detectable information about...
  - The speaker
  - Behavior context (e.g. feeding, mating, perching, sleep)
  - The desired listener
- 

## **Prairie Dogs**



## **Prairie Dog Language?**

- Dr. Con Slobodchikoff (<http://jan.ucc.nau.edu/~cns3/>) argues that prairie dogs have...
- Different calls for different species of predators
- Calls containing descriptive information about location, size, shape, color of predators
- Prairie dogs will alarm differently on the basis of your shirt

- Yes, prairie dogs are judging your fit
  - Some signs of new calls for shapes
  - A great video on this (<https://www.youtube.com/watch?v=y1kXCh496U0>)
- 

## **But is it language?**

- We've only seen one domain of 'language'
    - A rich vocabulary of various forms of 'Oh crap! Hide!'
  - There's no sign of non-pragmatic language use
    - e.g. "Hey Herbie, remember all those grubs we ate last month? Yeah, that was great"
  - So, we'd need more data to call this 'Language'
    - ... and I need to see this work replicated!
- 

## **Whale whale whale**

- We've just discovered that whales produce vowels as a part of their clicks
  - This points us towards new approaches of understanding whale language!
- 

## **So, do non-human animals have Language?**

- Nobody's truly doing human language yet
  - There's evidence of complex information being encoded
  - ... but we're still gonna need more evidence before it's 'Language'
- 

## **So, we know what language is.**

- We know it's a universal thing, although with variation
  - And we know that animals don't seem to do it (yet?)
  - Now, for the (potentially) scarier side of Language
- 

## **We think in language, regularly**

- We use linguistic concepts in categorizing the world

- We discuss the world using our most available set of terms
  - We lack terms for some more abstract feelings
  - So, language is a part of our thinking process
  - This leads many to ask...
- 

## Does the language we speak determine how we think?

---

### The Sapir-Whorf Hypothesis

- Also known as 'Linguistic Relativity'
  - 'The language you speak influences how you think about the world'
  - This is fascinating and terrifying
- 

### There are many possible forms

- **No effect:** 'The language we speak has *zero* effect on our thought or cognition'
  - **Weak:** 'The language we speak has some effect on our thought and cognition'
  - **Strong:** 'The language we speak guides our thought and cognition'
  - **Determinism:** 'The language we speak *determines* our thought and cognition'
- 

### The debate is now mostly about effect size and domain

- It's clear that there's *some* effect
    - See Dr. Lera Boroditsky at UCSD make a great case  
([https://www.ted.com/talks/lera\\_boroditsky\\_how\\_language\\_shapes\\_the\\_way\\_we\\_think](https://www.ted.com/talks/lera_boroditsky_how_language_shapes_the_way_we_think)).
  - But it's also clear that it's not deterministic
    - We can recognize and discuss objects, events, or colors which we don't have words for
  - The debate is now focused on the size and nature of the effect
    - ... and whether/how much it *actually* matters in our lives
-

## **This is an amazing area of research**

- ... and it's one I wish we could spend more time on
- 

## **For next time...**

- Watch Dr. Boroditsky's TED Talk  
([https://www.ted.com/talks/lera\\_boroditsky\\_how\\_language\\_shapes\\_the\\_way\\_we\\_think](https://www.ted.com/talks/lera_boroditsky_how_language_shapes_the_way_we_think))
    - Not required, but worth your time
- 

Thank you!