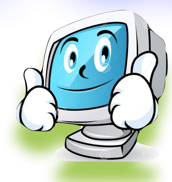
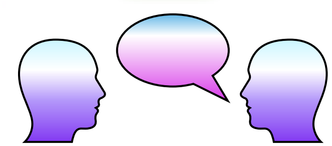
How math helps us better understand language

Linguistics Studies Lecture Series
College of Charleston
February 18, 2021

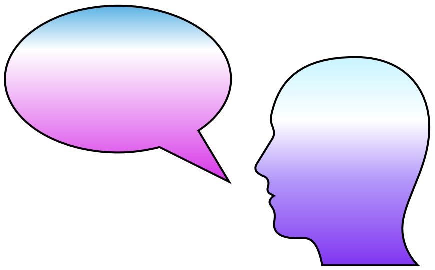
Lisa S. Pearl
Professor
Department of Language Science
SSPB 2219
University of California, Irvine
lpearl@uci.edu

Computation of Language Laboratory
UC Irvine





About language

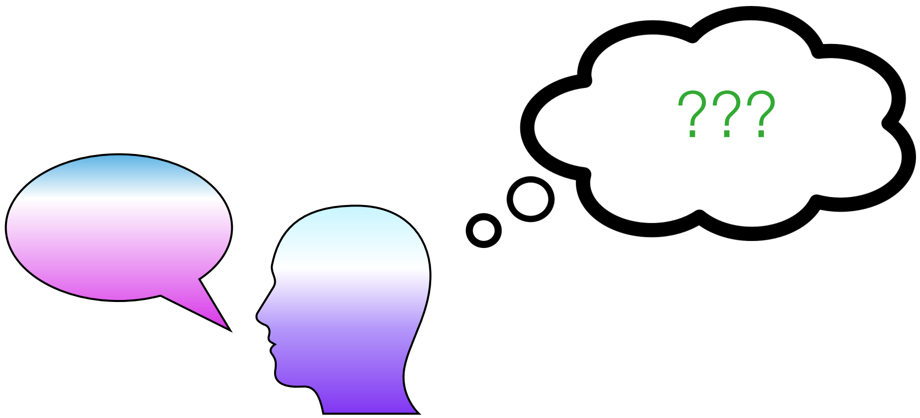


Language is a uniquely human ability.

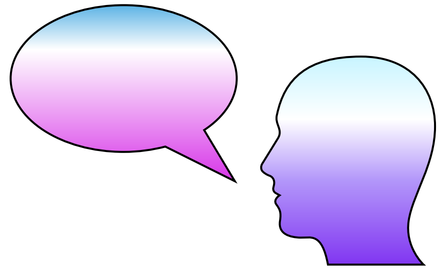
About language

We can investigate many things about language, including

what we know when we know language



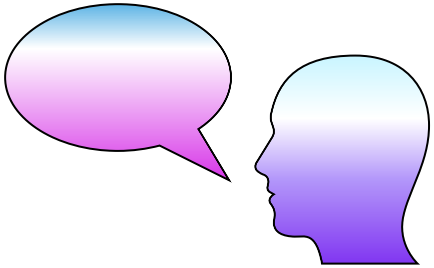
About language



We can investigate many things about language, including

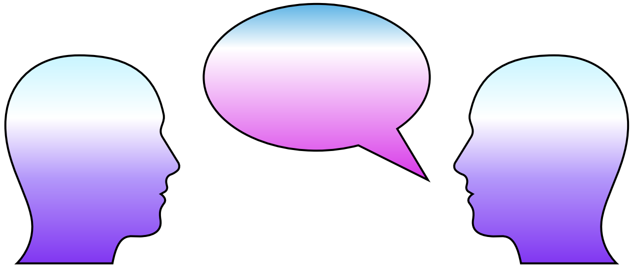
how language knowledge develops

About language

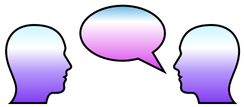


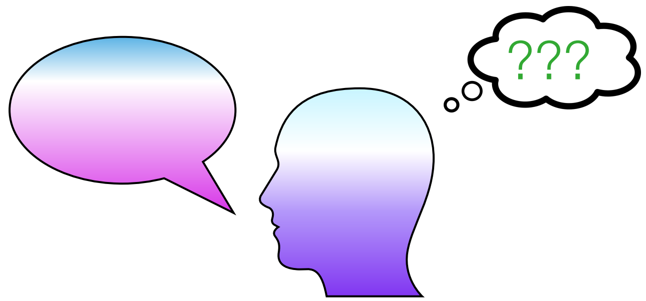
We can investigate many things about language, including

how we use language to
communicate different information



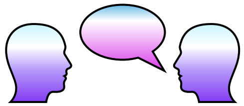
About language

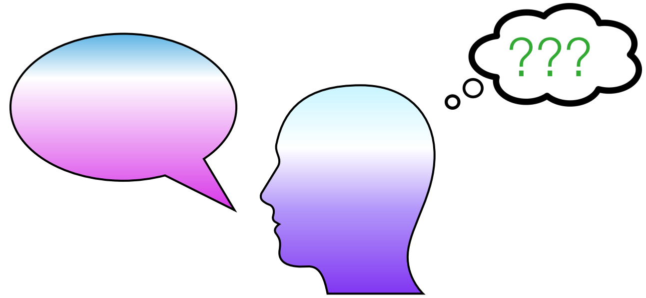




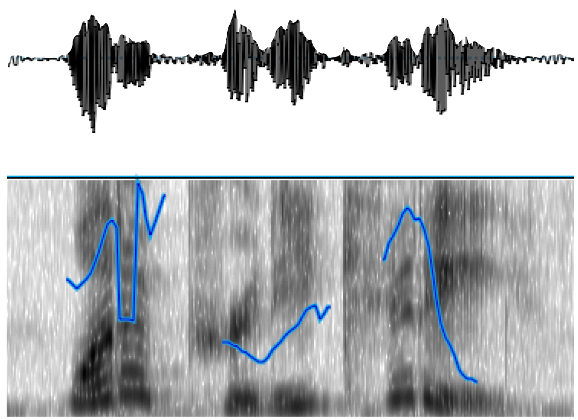
So what kinds of things do we know when we know language?

About language





how to identify words in fluent speech (speech segmentation)

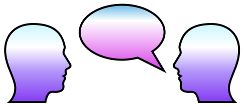


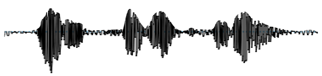
wʌɾəpɹɪɾikɪɾi

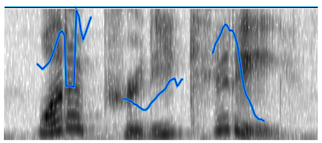
wʌɾ ə pɹɪɾi kɪɾi

what a pretty kitty!

About language

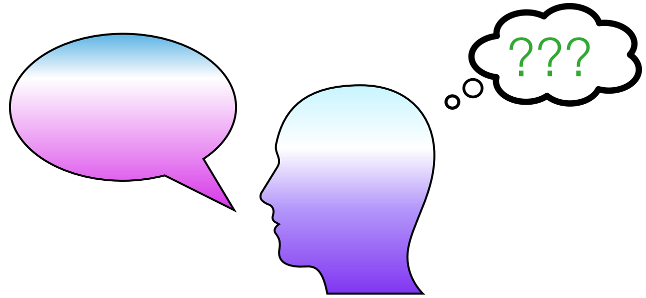






what a pretty kitty!

speech segmentation

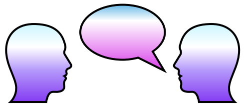


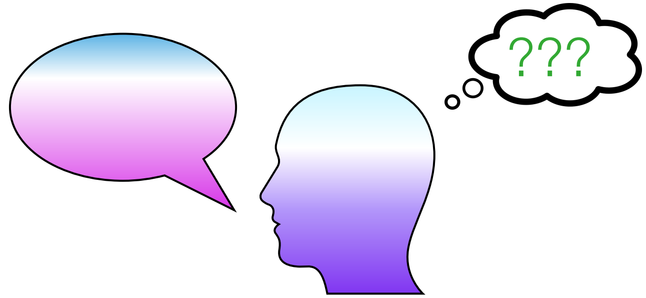
how to pronounce words (phonology)

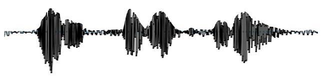
✓ KI tty

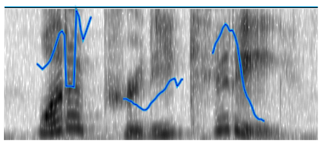
✗ ki TTY

About language









what a pretty kitty!

speech segmentation

KI tty

ki TTY

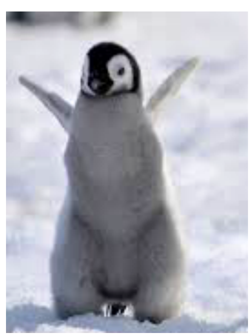
phonology

certain words behave like other words
(syntactic categorization)

Noun

what a pretty ___!

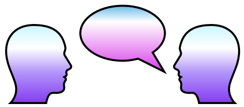
penguin

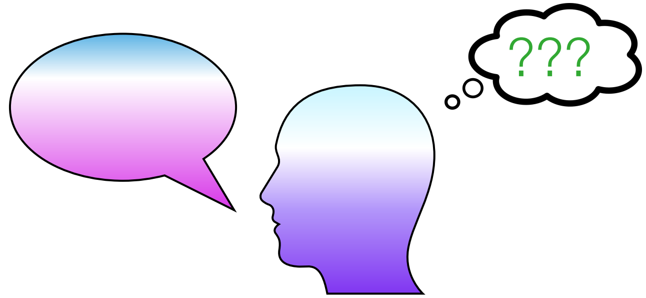


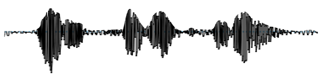
kitty

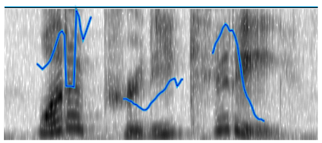
owl

About language









what a pretty kitty!

speech segmentation

✓ KI tty

✗ ki TTY

phonology

penguin **Noun** owl

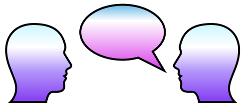
kitty

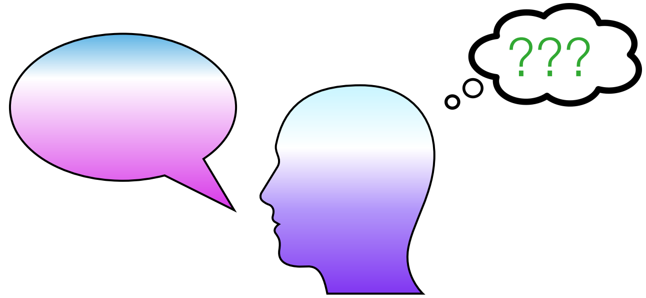
syntactic categorization

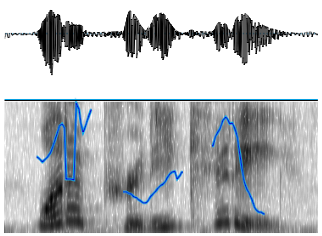
how to interpret words in context
(syntax, semantics)

“Oh look — a pretty kitty!”
“Look — there’s another one!”

About language







what a pretty kitty!

speech segmentation

KI tty

ki TTY

phonology

penguin Noun owl

kitty

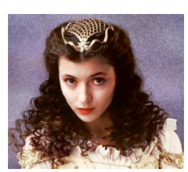
syntactic categorization

"Oh look — a pretty kitty!"
"Look — there's another one!"

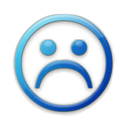
syntax, semantics

how to put words together to ask questions (syntax)

This kitty was bought as a present for someone.

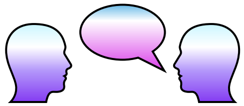


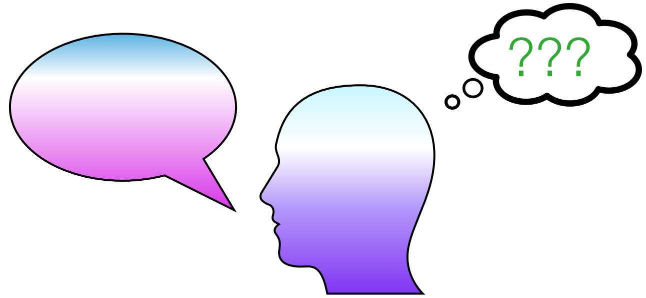
Lily thinks this kitty is pretty.

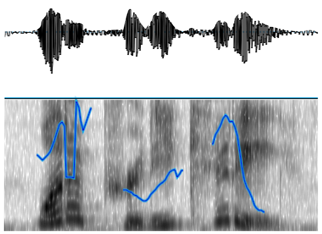


"Who does Lily think the kitty for is pretty?"

About language







what a pretty kitty!

speech segmentation

KI tty

ki TTY

phonology

penguin **Noun** owl

kitty

syntactic categorization

“Oh look — a pretty kitty!”
“Look — there’s another one!”

syntax, semantics

“Who does Lily think the kitty for is pretty?”

syntax

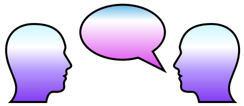
how to identify the right interpretation in context (pragmatics)

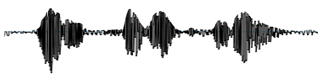
“Every kitty didn’t sit on the stairs”

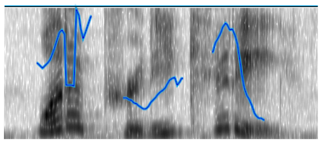
No kitties sat on the stairs.

Not all kitties sat on the stairs.

About language







what a pretty kitty!

speech segmentation

KI tty

ki TTY

phonology

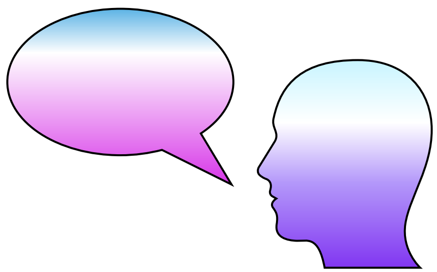
penguin **Noun** owl

kitty

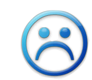
syntactic categorization

"Oh look — a pretty kitty!"
"Look — there's another one!"

syntax, **semantics**



what we know



"Who does Lily think the kitty for is pretty?"

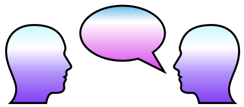
syntax

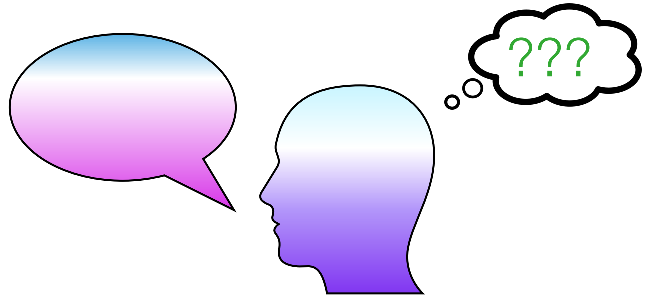
"Every kitty didn't sit on the stairs"

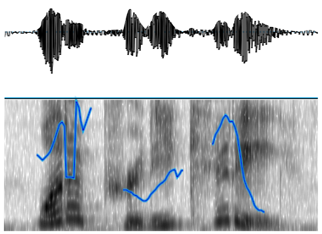
Not all kitties sat on the stairs.

pragmatics

About language







what a pretty kitty!

speech segmentation

✓ KI tty

✗ ki TTY

phonology

penguin **Noun** owl

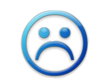
kitty

syntactic categorization

"Oh look — a pretty kitty!"
"Look — there's another one!"

syntax, semantics

How did we develop that knowledge?



"Who does Lily think the kitty for is pretty?"

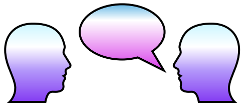
syntax

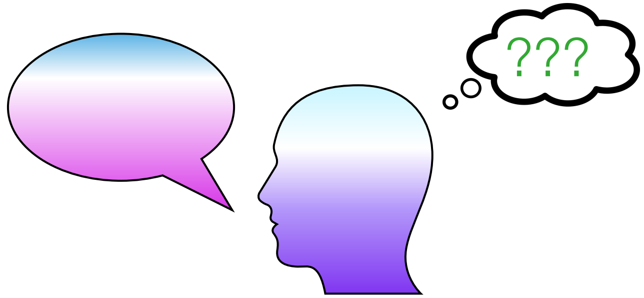
"Every kitty didn't sit on the stairs"

✓ Not all kitties sat on the stairs.

pragmatics

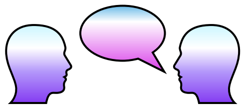
About language

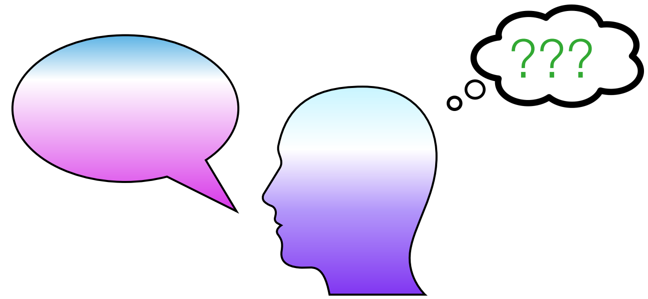




Children are amazing at learning language

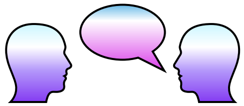
About language

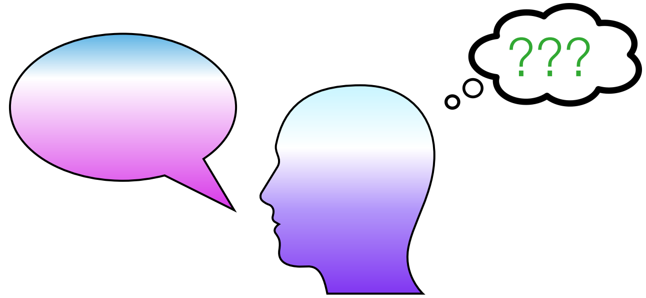




Much of the linguistic system is already known by age 4.

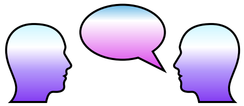
About language

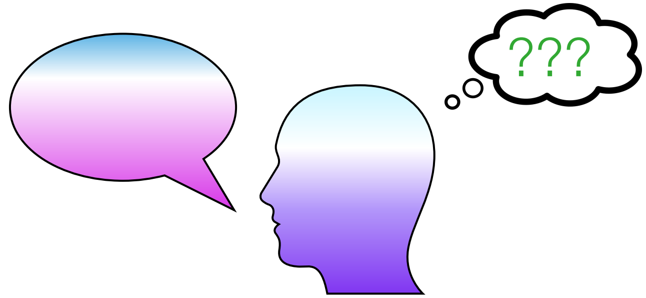




Also, children figure language out mostly without explicit instruction.

About language

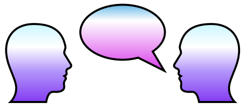


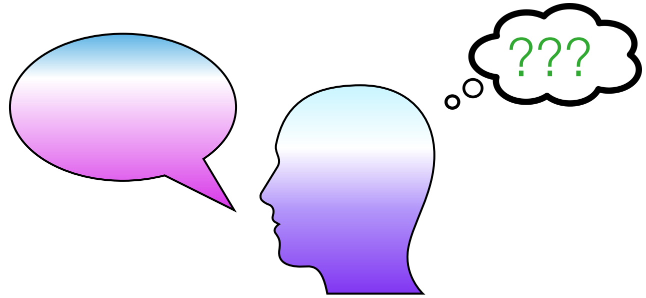


What they're doing: Extracting patterns and making generalizations from the surrounding data mostly just by hearing examples of what's allowed in the language.

What's so hard about that?

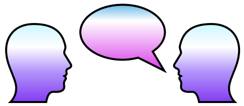
About language

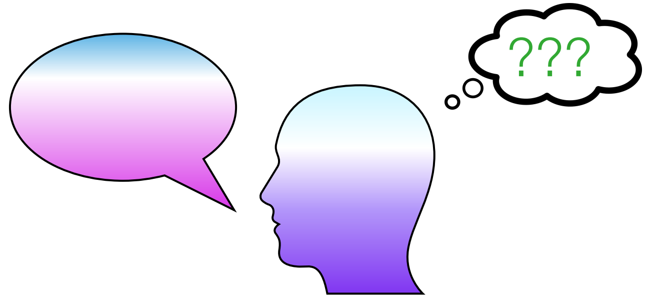




There are often many ways to generalize beyond the input, and most of them aren't right.

About language



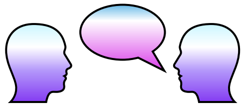


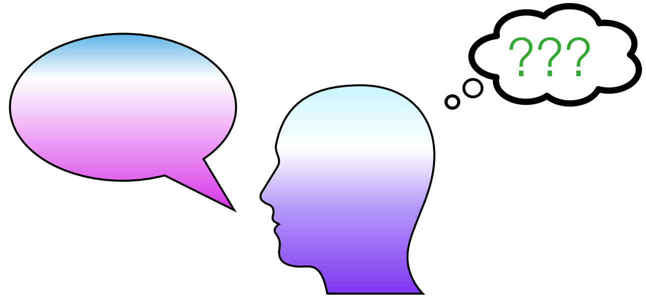
There are often many ways to generalize beyond the input, and most of them aren't right.

???
"birdie"

"What a pretty birdie!"

About language



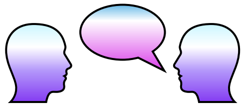


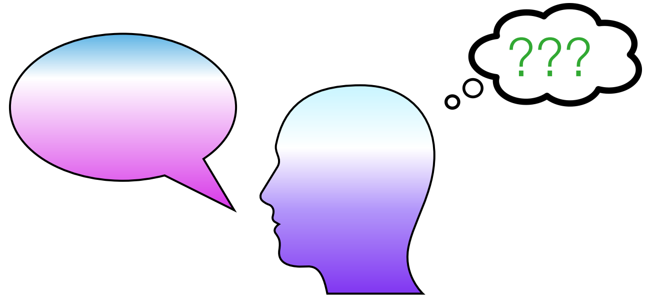
There are often many ways to generalize beyond the input, and most of them aren't right.

???
"birdie"

"Look - a birdie!"

About language



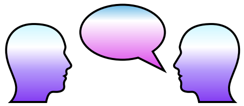


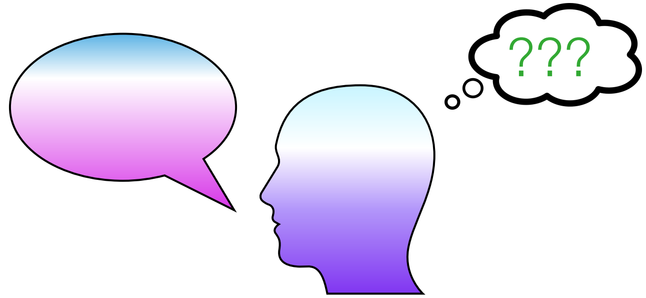
There are often many ways to generalize beyond the input, and most of them aren't right.

???
"birdie"

"Look at that birdie!"

About language



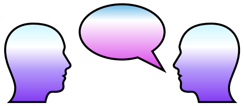


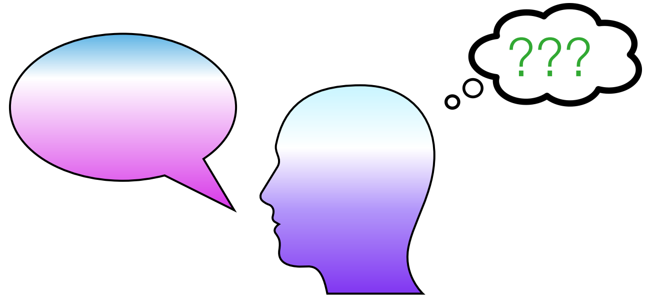
There are often many ways to generalize beyond the input, and most of them aren't right.

???
"birdie"

What generalization to make?

About language

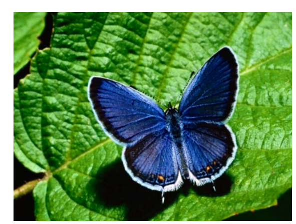




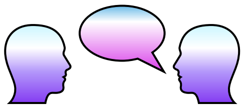
There are often many ways to generalize beyond the input, and most of them aren't right.

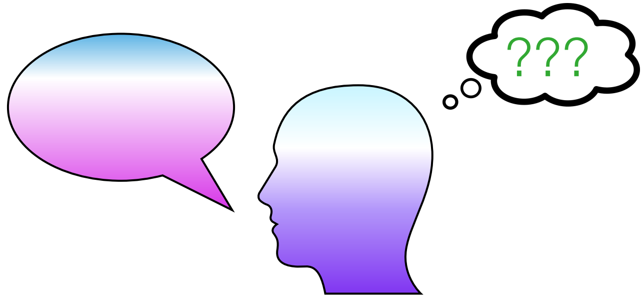
???
"birdie"

= blue creature



About language



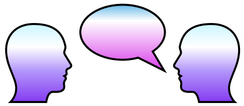


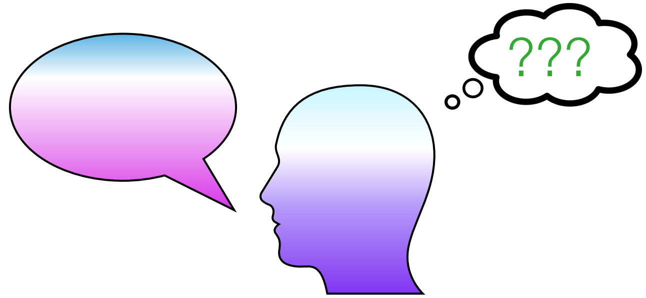
There are often many ways to generalize beyond the input, and most of them aren't right.

???
"birdie"

= creature on branch

About language



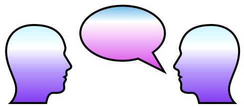


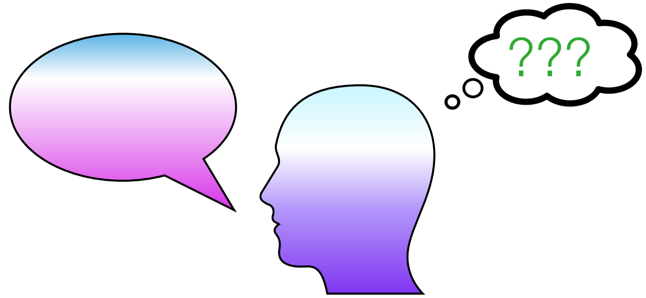
There are often many ways to generalize beyond the input, and most of them aren't right.

???
"birdie"

= [whatever makes something a bird]

About language





These kinds of induction problems are everywhere in cognitive development, including language development.

phonology

syntax, semantics

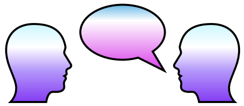
syntactic categorization

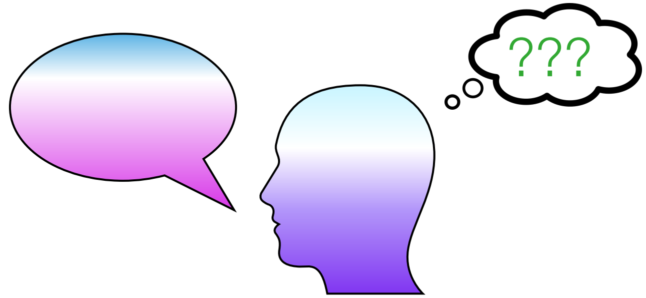
speech segmentation

syntax

pragmatics

About language





These kinds of induction problems are everywhere in cognitive development, including language development.

phonology

syntax, semantics

speech segmentation

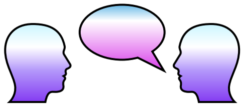
syntactic categorization

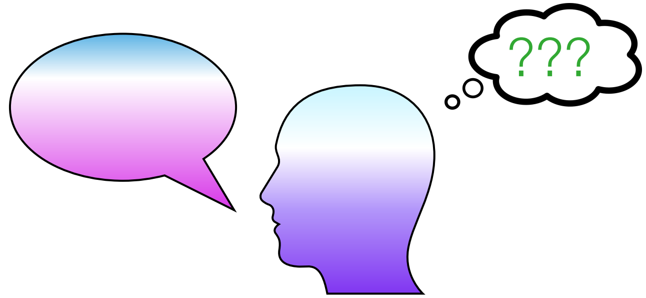
pragmatics

syntax

Language development =
Solving a lot of induction problems.

About language





It’s amazing how good kids are at this.

phonology

syntax, semantics

syntactic categorization

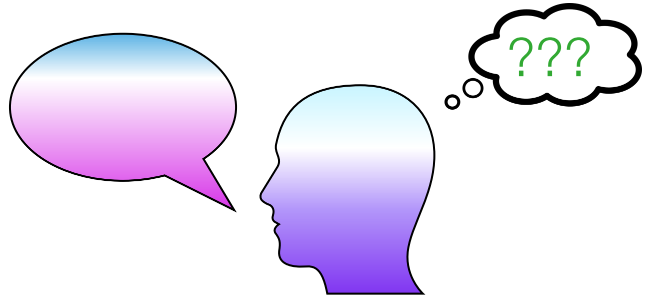
speech segmentation

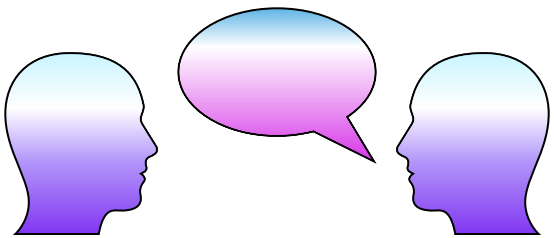
pragmatics

syntax

Language development =
Solving a lot of induction problems.

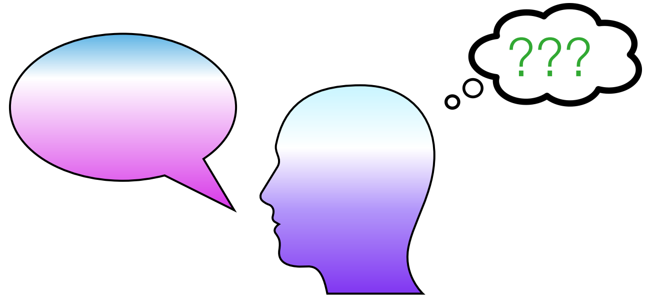
About language

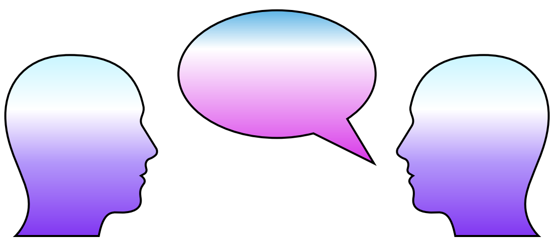




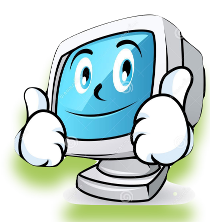
How do we use language knowledge to communicate?

About language



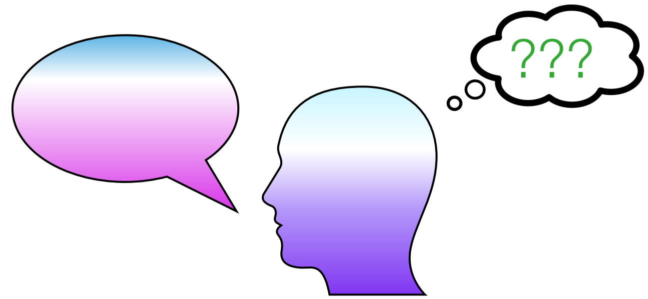


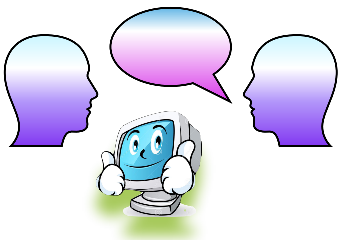
How do we use language knowledge to communicate?



And can we help machines do it too?

About language



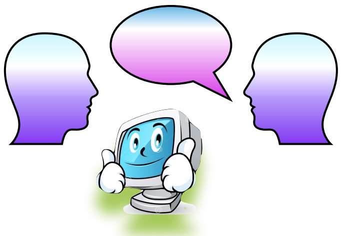


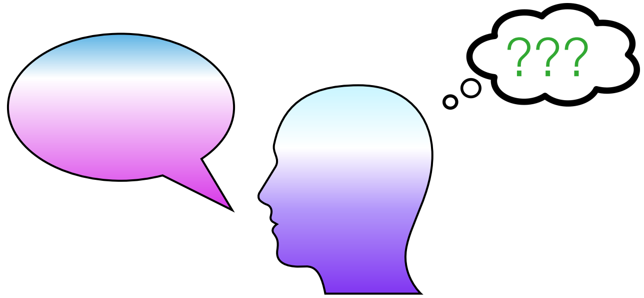
Exclamation

Yes/no question

"C'mon — don't you think this is awesome?"

About language





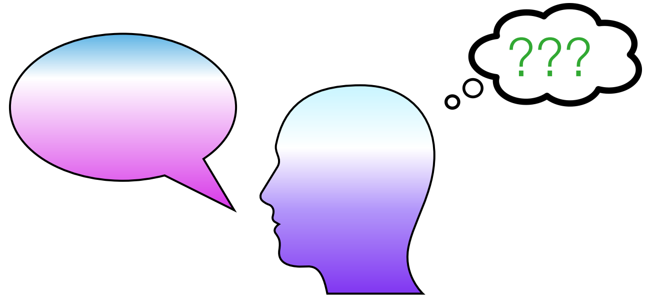
"C'mon — don't you think this is awesome?"

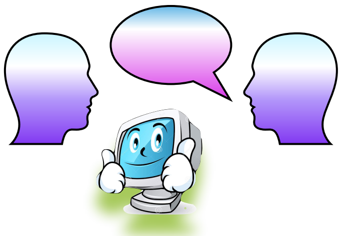
addressee

mental state

good++++

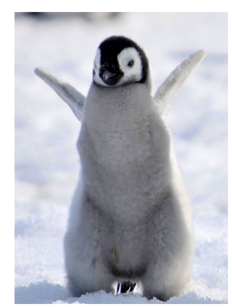
About language



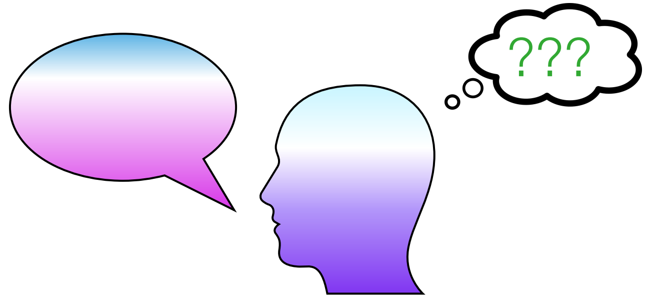


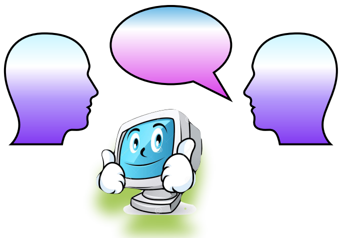
“C’mon — don’t you think this is awesome?”

something salient
in the context



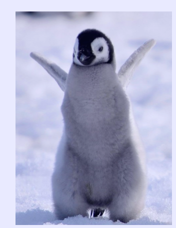
About language



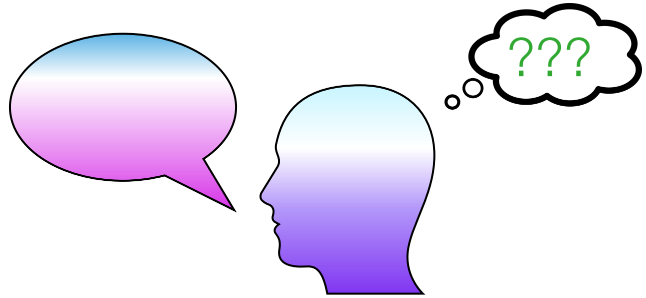


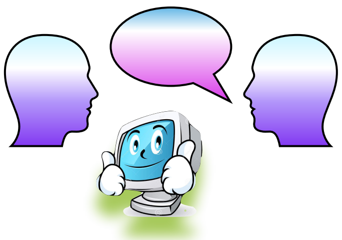
"C'mon — don't you think this is awesome?"

core information



About language

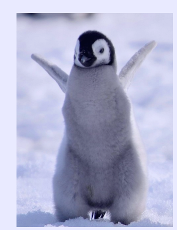




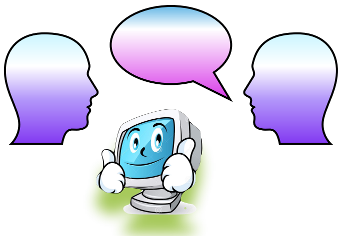
But more subtle information is communicated too.

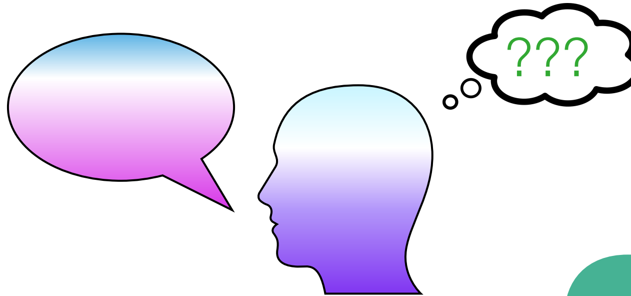
“C’mon — don’t you think this is awesome?”

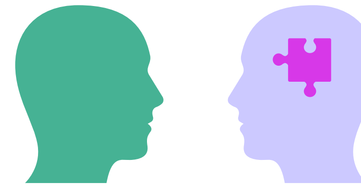
core information



About language





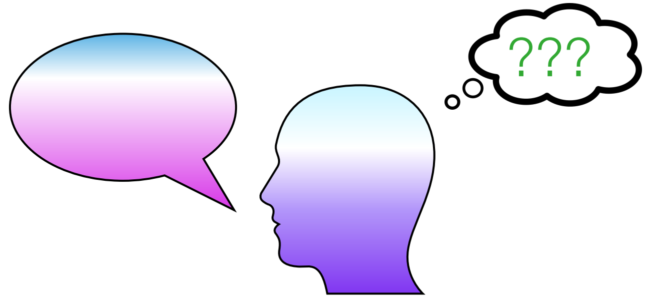


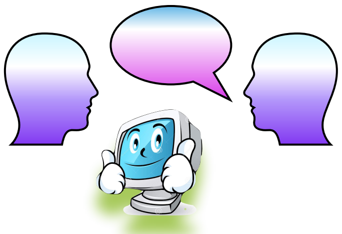
The speaker likely has a persuasive intention.

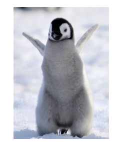
"C'mon — don't you think this is awesome?"

core information

About language

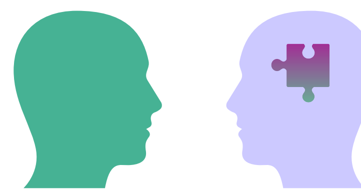




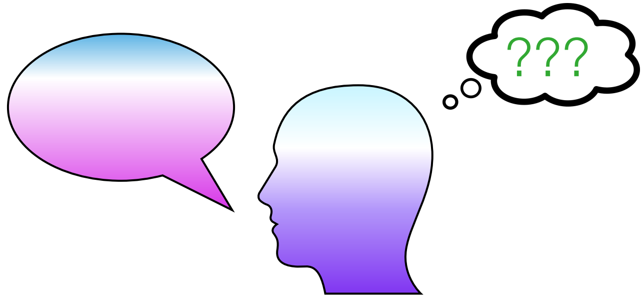


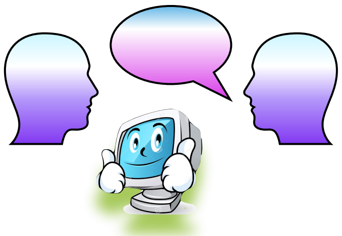
"C'mon — don't you think this is awesome?"

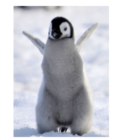
If the speaker actually doesn't like penguins, he could be intending to ingratiate himself with the addressee (using deception).



About language

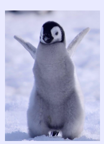






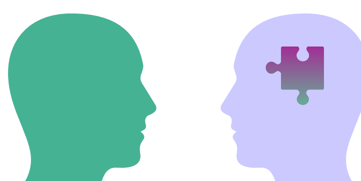
“C’mon — don’t you think this is awesome?”

core information



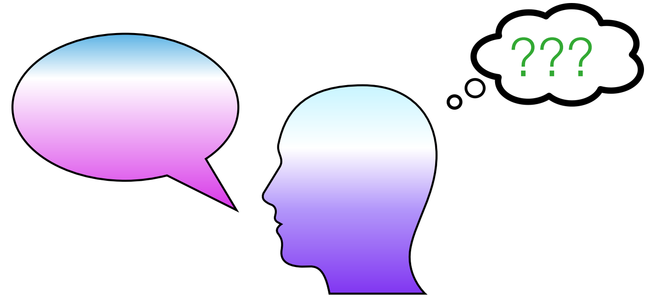
At face value, the speaker seems to have a good feeling about penguins (positive sentiment).

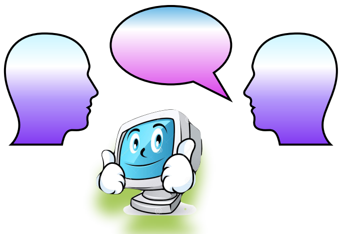
subtle information



intentions

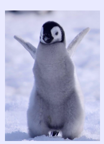
About language





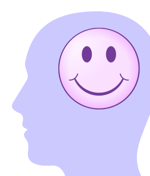
"C'mon — don't you think this is awesome?"

core information

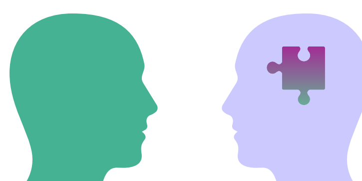


The casual style of speaking suggests familiarity with the addressee, and may indicate something about the speaker's identity.

subtle information

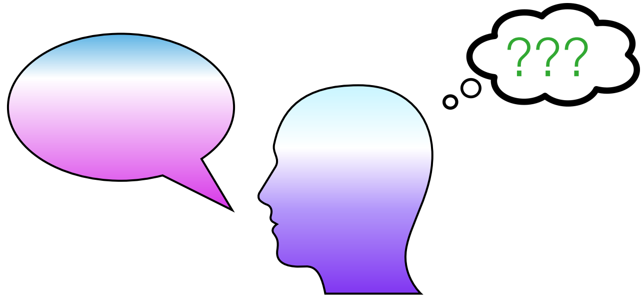


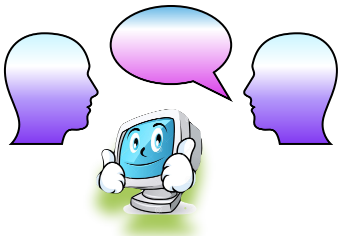
emotions/attitudes



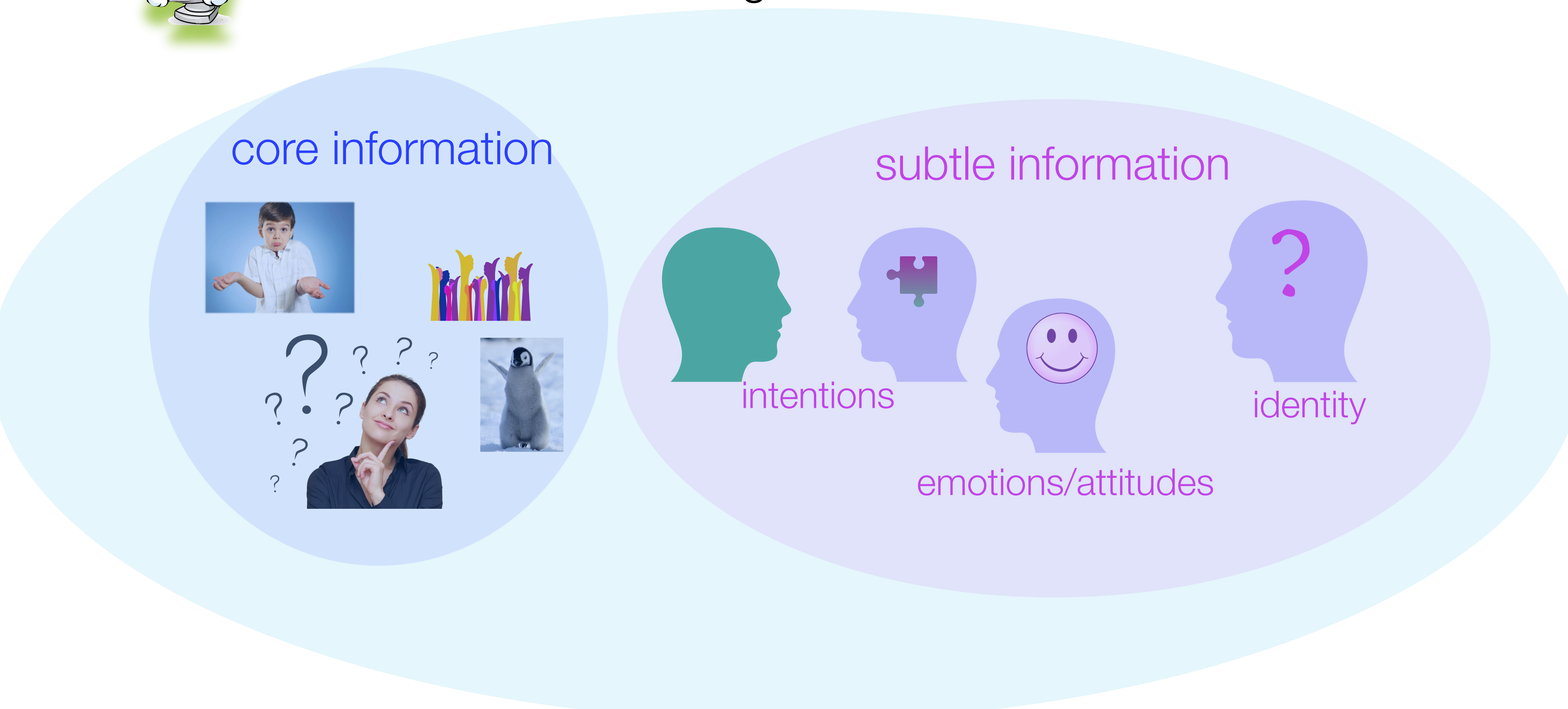
intentions

About language

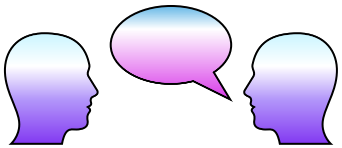


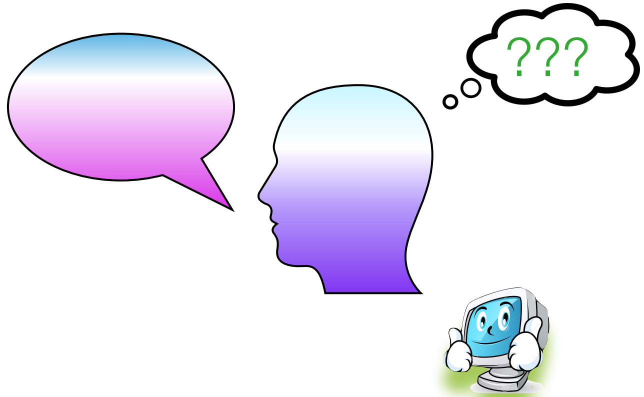


So, our knowledge of language use involves communicating both core and subtle information.

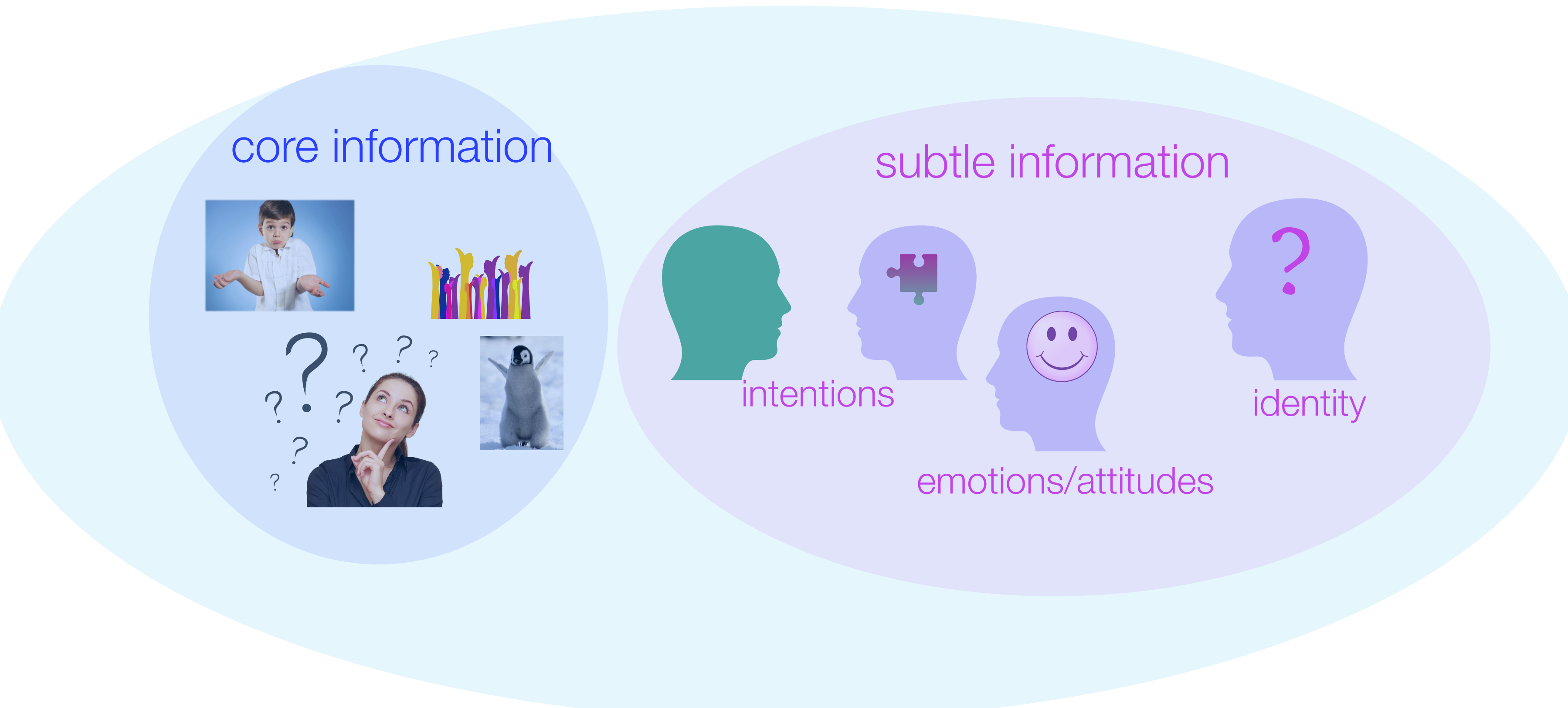


About language

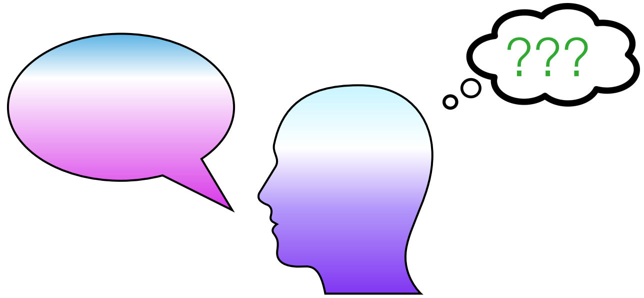




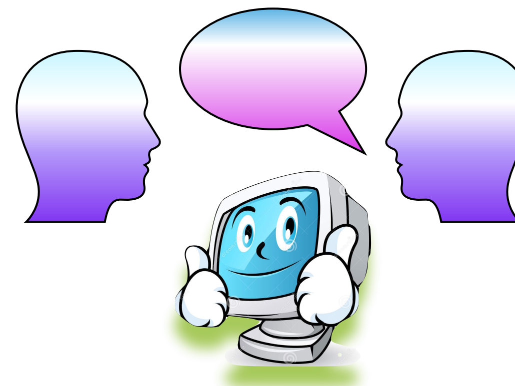
This is what we'd like machines to be able to do just as well as humans.



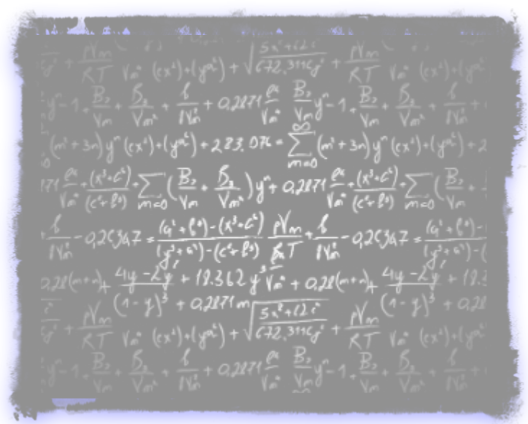
About language



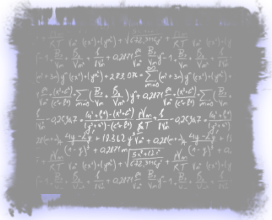
My research focuses on understanding development and use.



And I use math to do it!



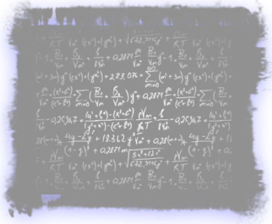
About math



Quantitative techniques = techniques that rely on math

One main part: Counting things

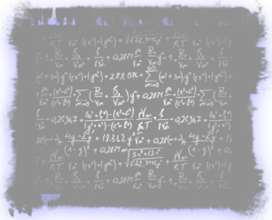
About math



Quantitative techniques = counting things

(sometimes we count a lot of things)

About math



Quantitative techniques = counting things

Another part: principled reasoning
based on those counts

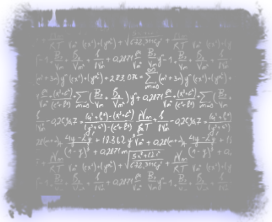
Bayesian inference

$$p(Generalization|Data) \propto p(Generalization) \cdot p(Data|Generalization)$$

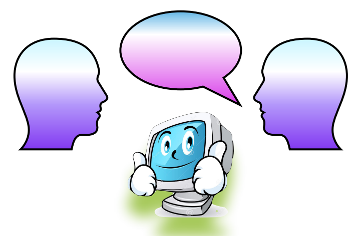
Tolerance & Sufficiency Principles

$$exceptions <= \frac{\#\ items}{ln(\#\ items)}$$

About math

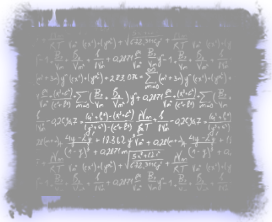


Quantitative techniques = counting things + principled reasoning

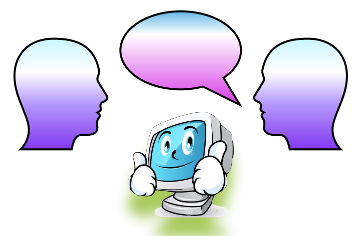


Then we use these quantitative techniques to help us understand how little humans develop language knowledge and how adult humans and machines use language knowledge

About math

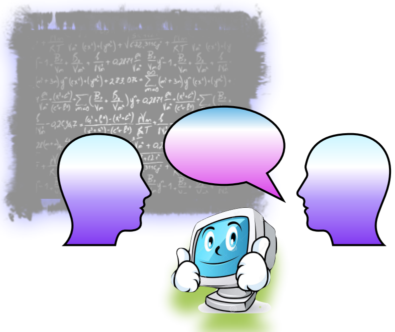


Quantitative techniques = counting things + principled reasoning

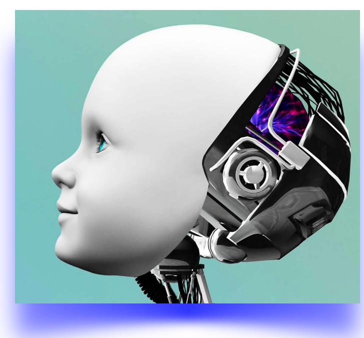


But what do we count and reason over? How do we connect that information to language development and use?

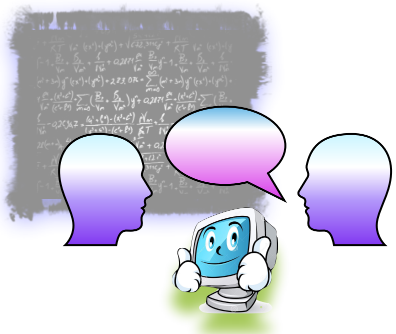
Quantitative techniques for language development

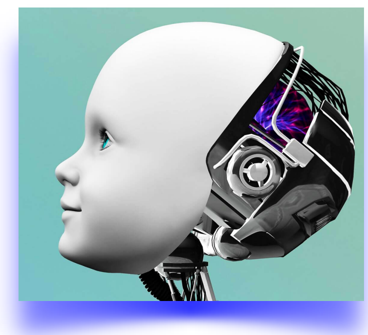


To understand language development, we're typically using computational cognitive modeling to encode a child's learning process very precisely.



Quantitative techniques for language development

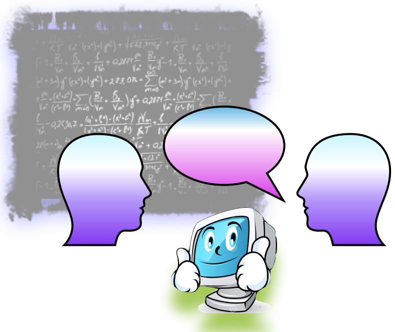


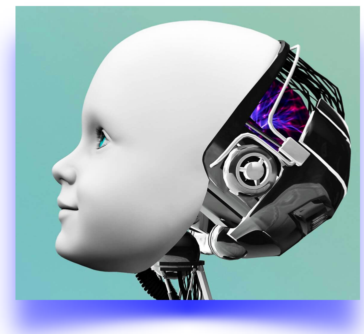


We think the child is learning by counting different parts of her input and reasoning over those counts in a sensible way.

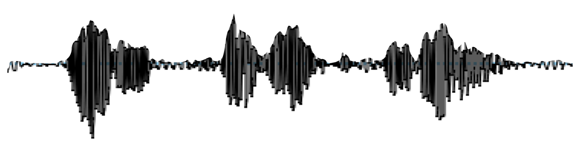
So, the model will count those same things and learn about language by doing principled reasoning over those counts.

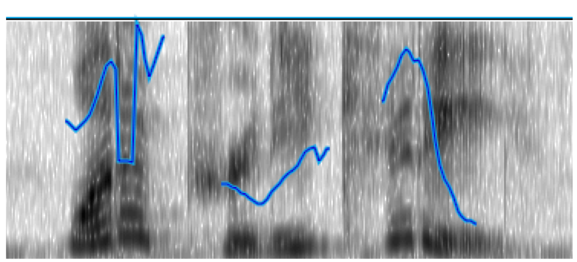
Quantitative techniques for language development





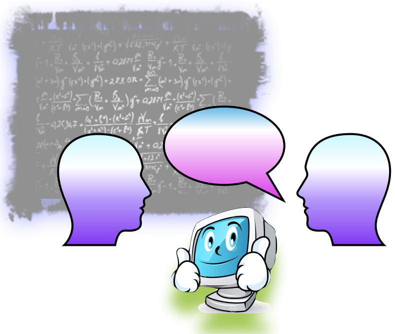
Let’s think about this for speech segmentation.

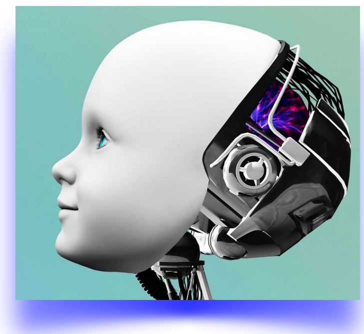




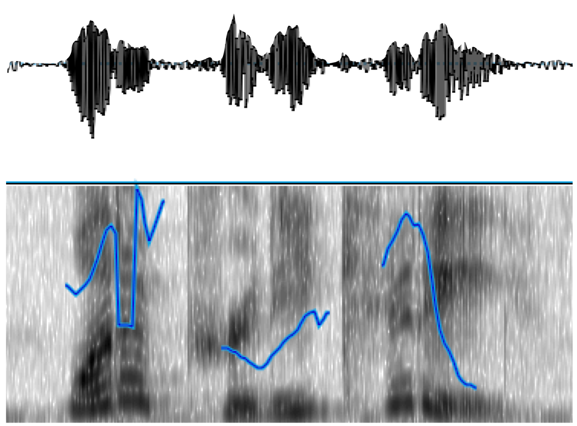
whataprettykitty!

Quantitative techniques for language development



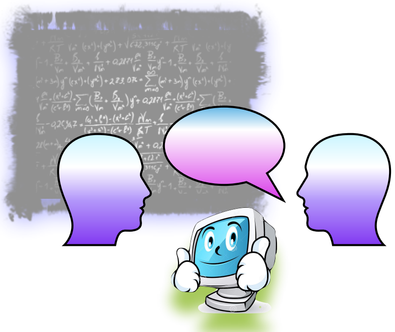


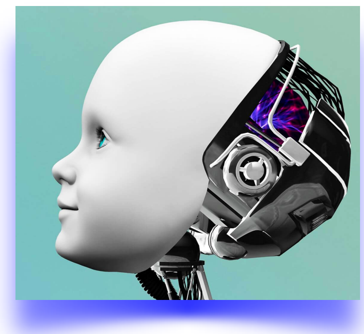
Let's think about this for speech segmentation.



what a pretty kitty!

Quantitative techniques for language development





These are the kinds of utterances infants hear.

what a pretty kitty!

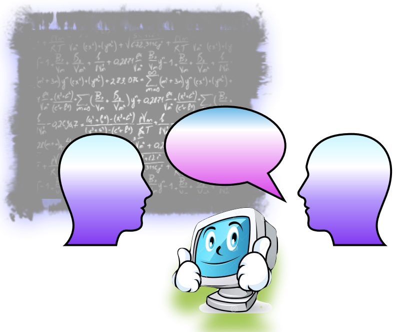
what a cute penguin!

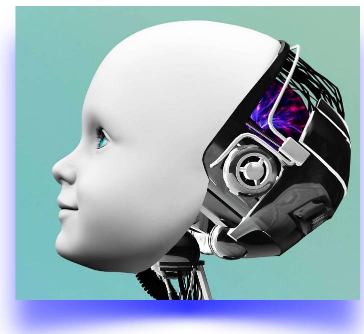
look at the pretty birdie!

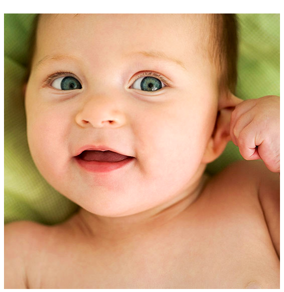
the kitty is very cute!

do you see the kitty?

Quantitative techniques for language development







what a pretty kitty!

what a cute penguin!

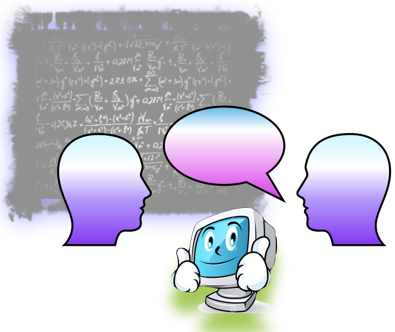
look at the pretty birdie!

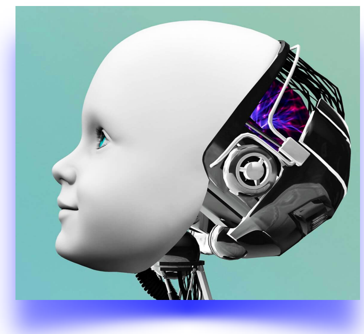
the kitty is very cute!

do you see the kitty?

But they’re more like this before infants know where the words are.

Quantitative techniques for language development





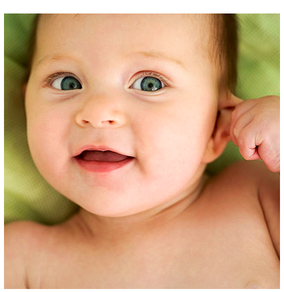
whataprettykitty

whatacutepenguin

lookattheprettybirdie

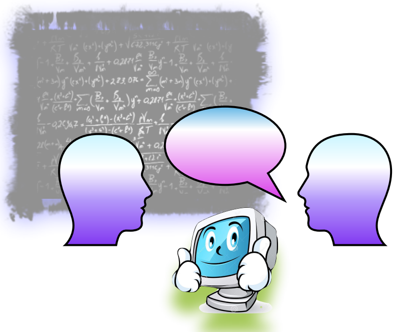
thekittyisverycute

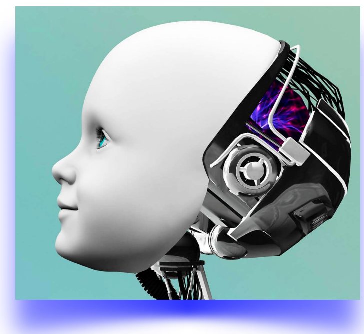
doyouseethekitty



But they’re more like this before infants know where the words are.

Quantitative techniques for language development





One idea is that children initially perceive syllables.

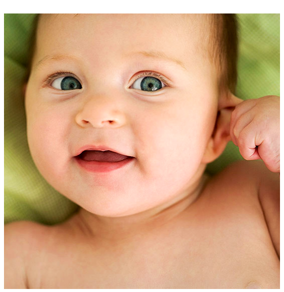
whataprettykitty

whatacutepenguin

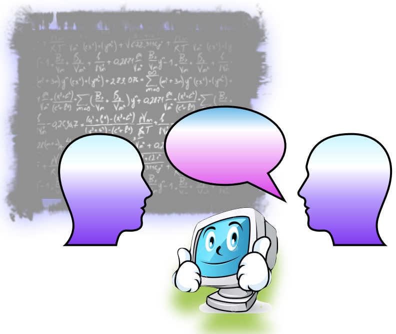
lookattheprettybirdie

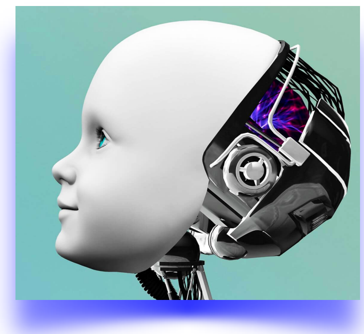
thekittyisverycute

doyouseethekitty



Quantitative techniques for language development

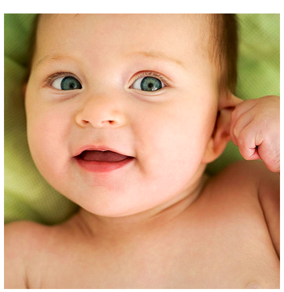




One idea is that children initially perceive syllables.

what a pre tty ki tty

what a cute pen guin

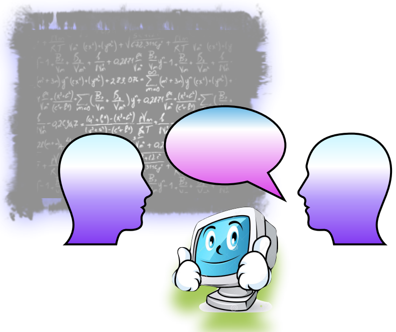


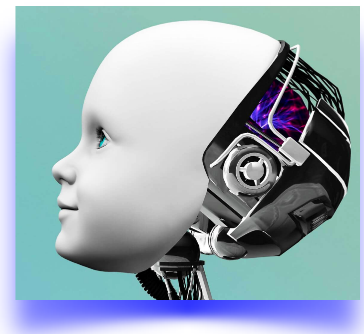
look at the pre tty bir die

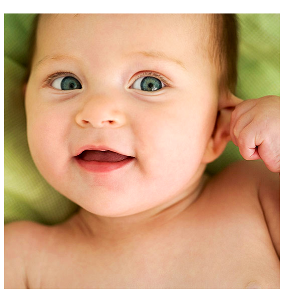
the ki tty is ve ry cute

do you see the ki tty

Quantitative techniques for language development







what a pre tty ki tty

what a cute pen guin

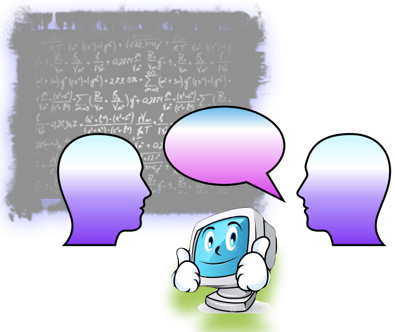
look at the pre tty bir die

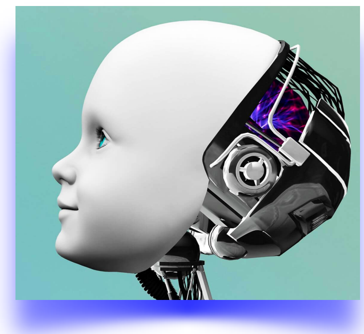
the ki tty is ve ry cute

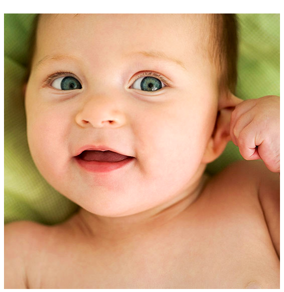
do you see the ki tty

One learning theory is that infants count syllables and where they occur.

Quantitative techniques for language development







what a pre tty ki tty

what a cute pen guin

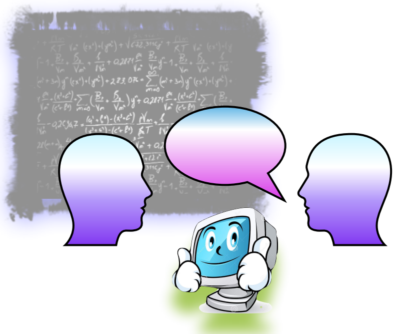
look at the pre tty bir die

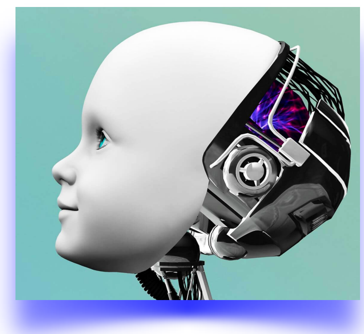
the ki tty is ve ry cute

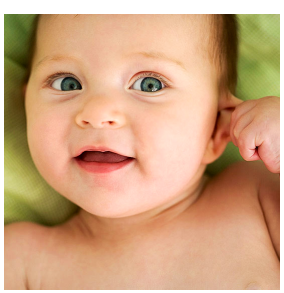
do you see the ki tty

Then, infants can reason about those counts to figure out where words are.

Quantitative techniques for language development







what a pre tty kitty

what a cute pen guin

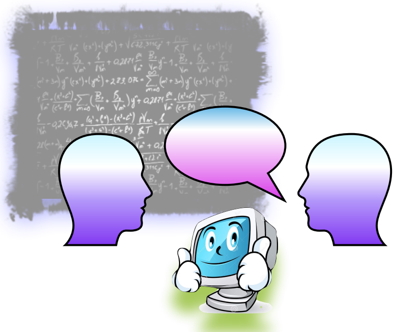
look at the pre tty bir die

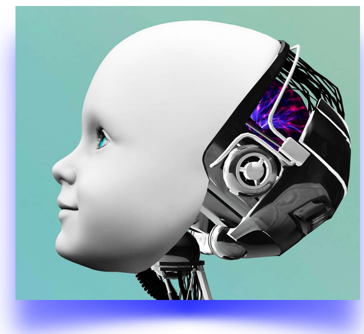
the kitty is ve ry cute

do you see the kitty

Then, infants can reason about those counts to figure out where words are.

Quantitative techniques for language development





what a pre tty kitty

what a cute pen guin

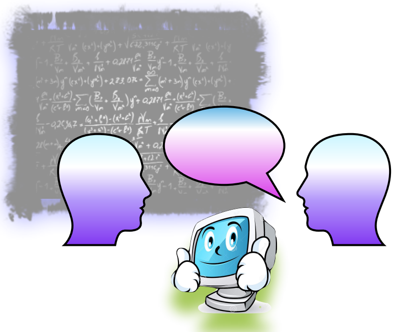
look at the pre tty bir die

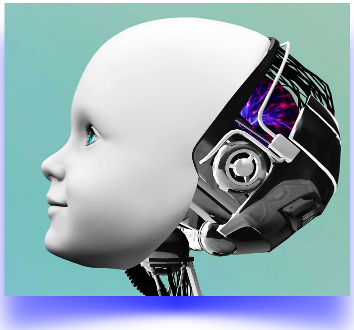
the kitty is ve ry cute

do you see the kitty

So this is what our computational cognitive model can do.

Quantitative techniques for language development





what a pre tty kitty

what a cute pen guin

look at the pre tty bir die

the kitty is ve ry cute

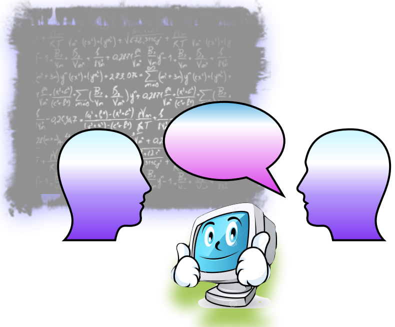
do you see the kitty

Example model from my research group: Use Bayesian inference to reason about counts of syllables in a child's input.

Bayesian inference

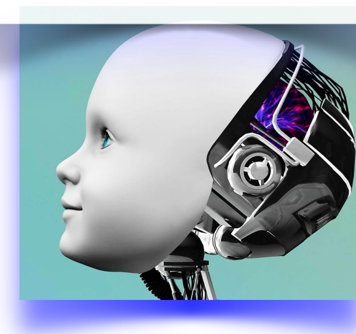
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Bayesian inference



This learning strategy involves the child imagining what collection of words (a lexicon) could be used to create the utterances she hears.

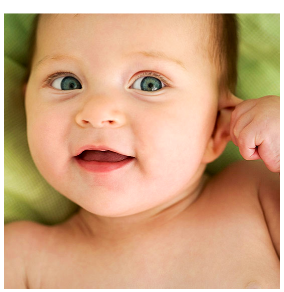
what a pre tty ki tty

what a cute pen guin

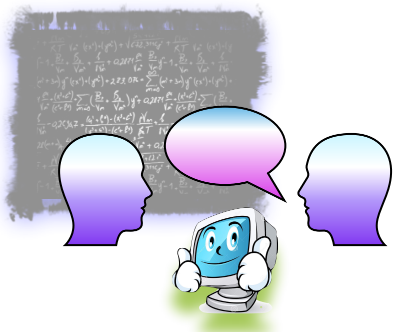
look at the pre tty bir die

the ki tty is ve ry cute

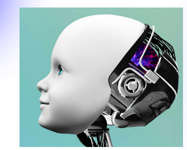
do you see the ki tty



Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018



Bayesian inference

Some possible lexicons a child might consider for the first utterance:

what a pre tty ki tty

whata
pre
tty
ki

whata
pretty
kitty

what
a
pretty
kitty

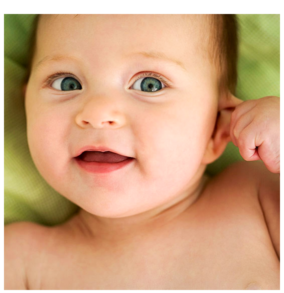
what
a
pre
tty
ki

whatapre
tty
ki

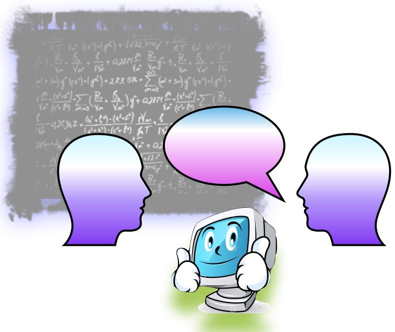
whatapre
tty
kitty

whatapretty
kitty

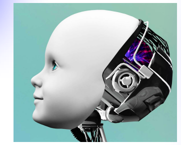
whataprettykitty



Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018



Bayesian inference

Each possible lexicon could be used to generate the observed utterance.

what a pre tty ki tty

whata
pre
tty
ki

whata
pretty
kitty

what
a
pretty
kitty

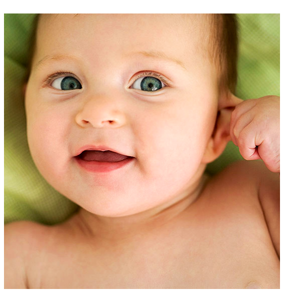
what
a
pre
tty
ki

whatapre
tty
ki

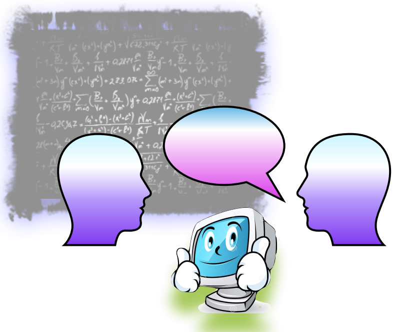
whatapre
tty
kitty

whatapretty
kitty

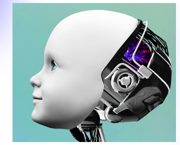
whataprettykitty



Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018



Bayesian inference

Each possible lexicon could be used to generate the observed utterance.

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

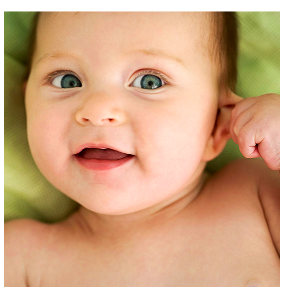
what a pre tty ki

whatapre tty ki

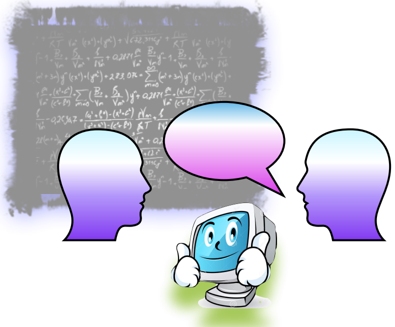
whatapre tty kitty

whatapretty kitty

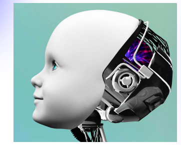
whataprettykitty



Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018



Bayesian inference

Each possible lexicon could be used to generate the observed utterance.

what a pre tty ki tty

whata pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

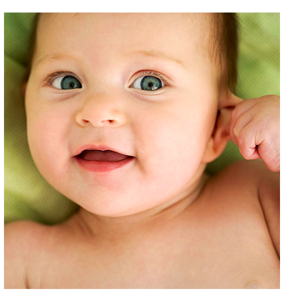
what a pre tty ki

whatapre tty ki

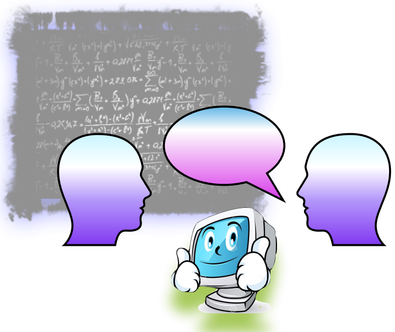
whatapre tty kitty

whatapretty kitty

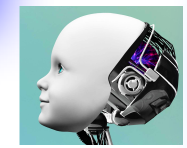
whataprettykitty



Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018



Bayesian inference

Each possible lexicon could be used to generate the observed utterance.

what a pre tty ki tty

whata pretty kitty

whata
pre
tty
ki

whata
pretty
kitty

what
a
pretty
kitty

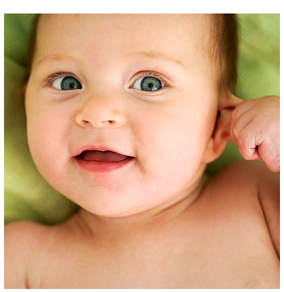
what
a
pre
tty
ki

whatapre
tty
ki

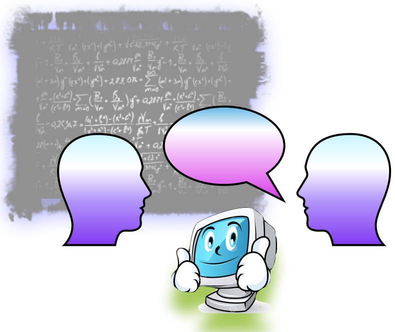
whatapre
tty
kitty

whatapretty
kitty

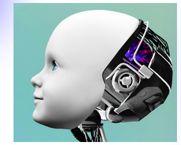
whataprettykitty



Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018



Bayesian inference

Each possible lexicon could be used to generate the observed utterance.

what a pre tty ki tty

whatapretty kitty

whata
pre
tty
ki

whata
pretty
kitty

what
a
pretty
kitty

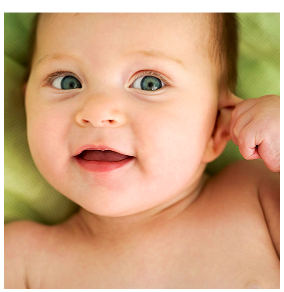
what
a
pre
tty
ki

whatapre
tty
ki

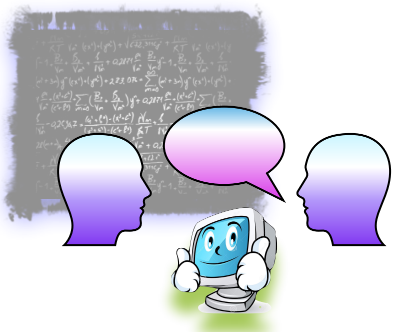
whatapre
tty
kitty

whatapretty
kitty

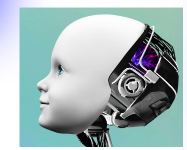
whataprettykitty



Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018



Bayesian inference

The child tries to identify the most probable one.

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

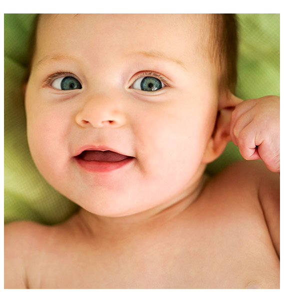
what a pre tty ki

whatapre tty ki

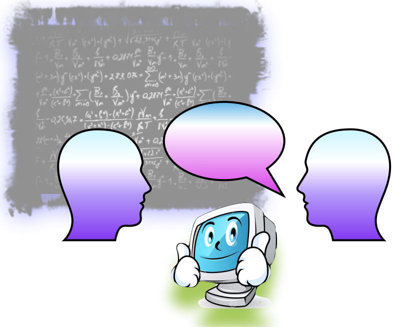
whatapre tty kitty

whatapretty kitty

whataprettykitty



Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

what a pre tty ki

whatapre tty ki

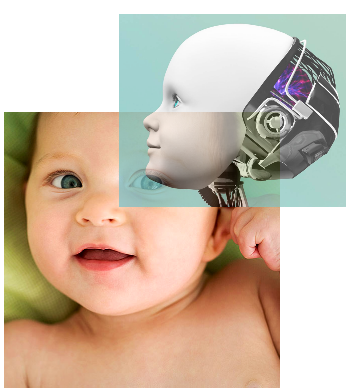
whatapre tty kitty

whatapretty kitty

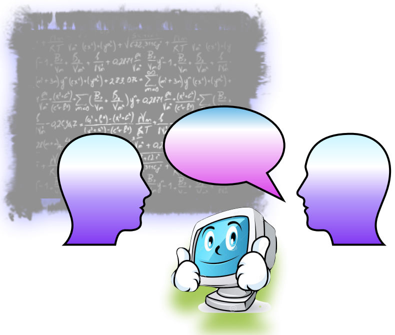
whataprettykitty

The modeled child in the computational cognitive model chooses the most probable one using Bayesian inference.

Bayesian inference



Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

what a pre tty ki

whatapre tty ki

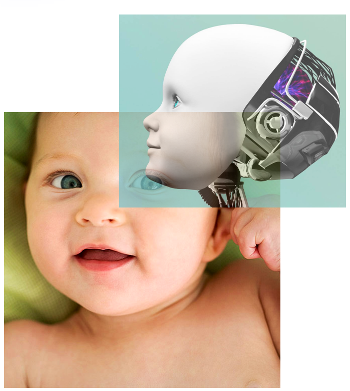
whatapre tty kitty

whatapretty kitty

whataprettykitty

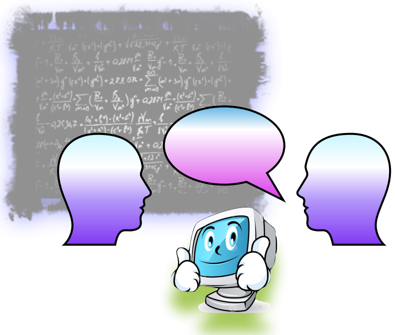
Bayesian inference

$$p(lexicon \mid utterances) \propto p(lexicon) \cdot p(utterances \mid lexicon)$$



We want to figure out the lexicon with the highest probability, given the utterances the child encounters (and the syllables in them).

Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

what a pre tty ki

whatapre tty ki

whatapre tty kitty

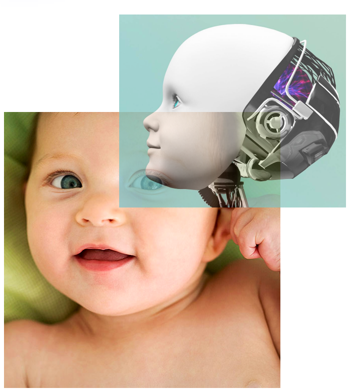
whatapretty kitty

whataprettykitty

Bayesian inference

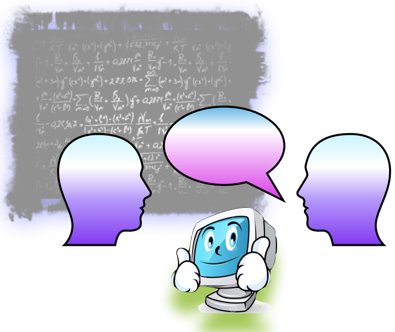
$$p(lexicon \mid utterances) \propto$$

$$p(lexicon) \cdot p(utterances \mid lexicon)$$



The model reasons about this by considering two other probabilities.

Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

what a pre tty ki

whatapre tty ki

whatapre tty kitty

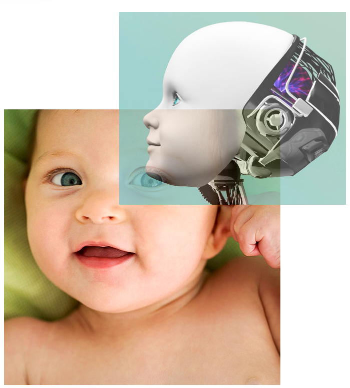
whatapretty kitty

whataprettykitty

Bayesian inference

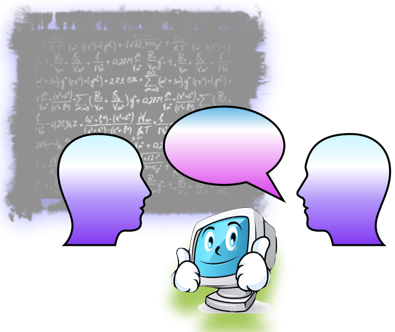
$$p(lexicon \mid utterances) \propto$$

$$p(lexicon) \cdot p(utterances \mid lexicon)$$



The prior probability of the lexicon captures any preferences the modeled child has (which we think real children would have too).

Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

what a pre tty ki

whatapre tty ki

whatapre tty kitty

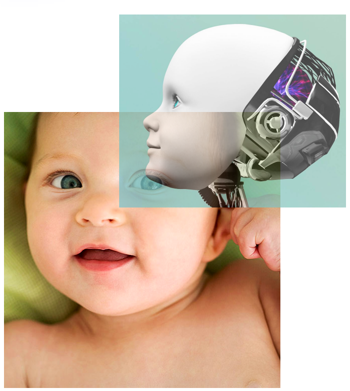
whatapretty kitty

whataprettykitty

Bayesian inference

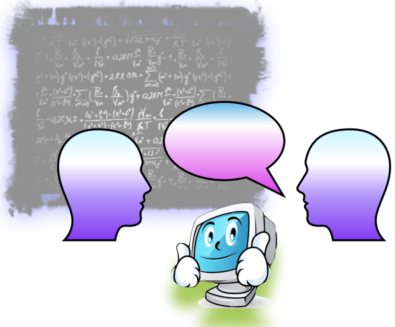
$$p(lexicon \mid utterances) \propto$$

$$p(lexicon) \cdot p(utterances \mid lexicon)$$



One preference: Prefer lexicons with fewer words.

Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

what a pre tty ki

whatapre tty ki

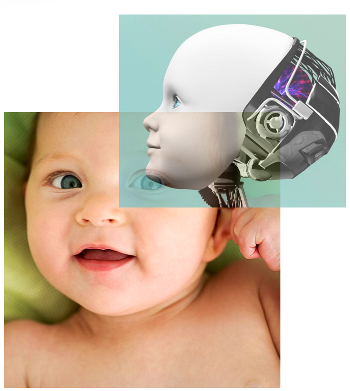
whatapre tty kitty

whatapretty kitty

whataprettykitty

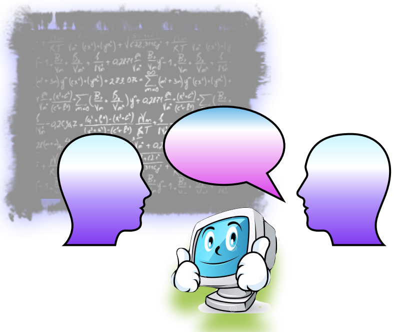
Bayesian inference

$$p(lexicon \mid utterances) \propto p(lexicon) \cdot p(utterances \mid lexicon)$$



Another preference: Prefer lexicons with shorter words.

Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

what a pre tty ki

whatapre tty ki

whatapre tty kitty

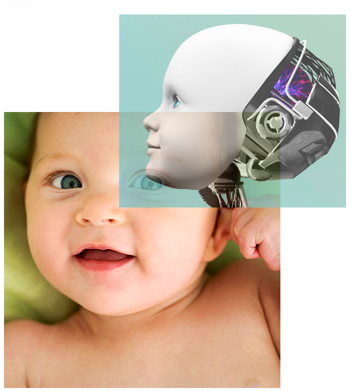
whatapretty kitty

whataprettykitty

Bayesian inference

$$p(lexicon \mid utterances) \propto$$

$$p(lexicon) \cdot p(utterances \mid lexicon)$$

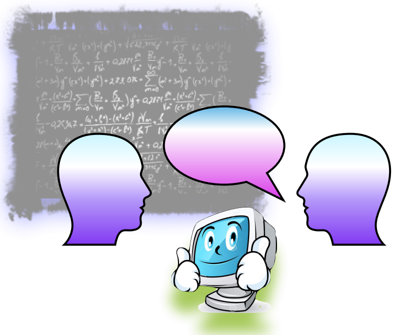


These preferences compete with each other.

shorter words

fewer words

Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

what a pre tty ki

whatapre tty ki

whatapre tty kitty

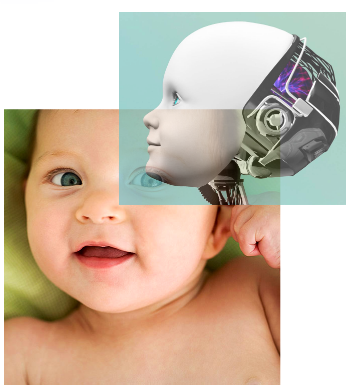
whatapretty kitty

whataprettykitty

Bayesian inference

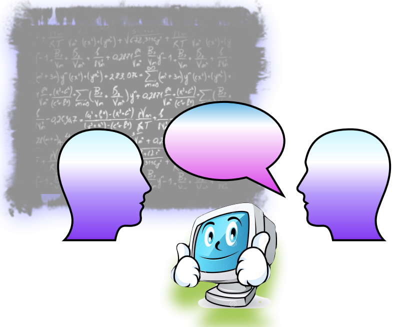
$$p(lexicon\,|\,utterances) \propto$$

$$p(lexicon) \cdot p(utterances\,|\,lexicon)$$



The modeled child uses Bayesian inference to reason about the lexicon that's the best balance between these preferences…

Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

whata pre tty ki

whata pretty kitty

what a pretty kitty

what a pre tty ki

whatapre tty ki

whatapre tty kitty

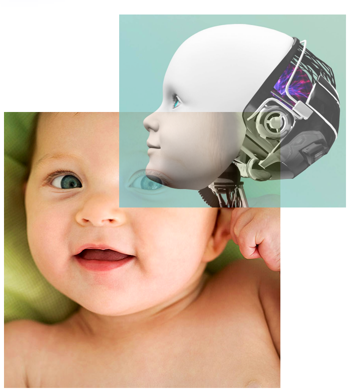
whatapretty kitty

whataprettykitty

Bayesian inference

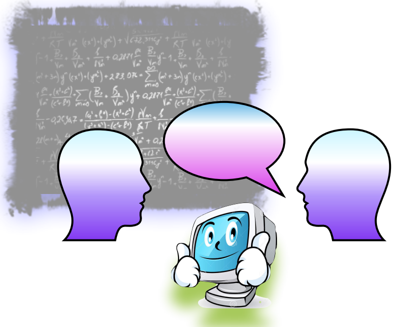
$$p(lexicon \mid utterances) \propto$$

$$p(lexicon) \cdot p(utterances \mid lexicon)$$



…that also has a high probability of generating the observed utterances = a high likelihood.

Quantitative techniques for language development



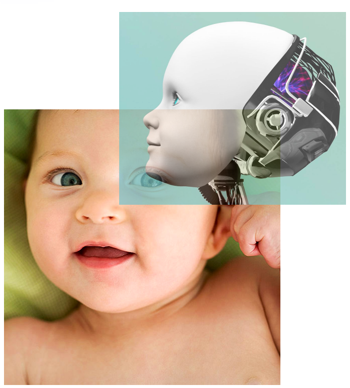
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

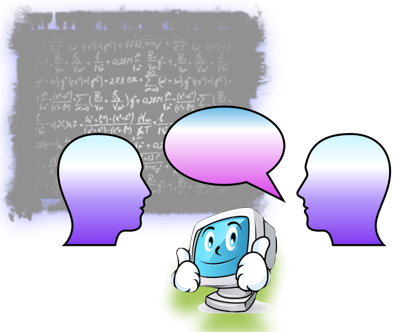
what
a
pretty
kitty

Bayesian inference

We can then see if the lexicon the modeled child identifies has the right kind of things in it (like real words).



Quantitative techniques for language development



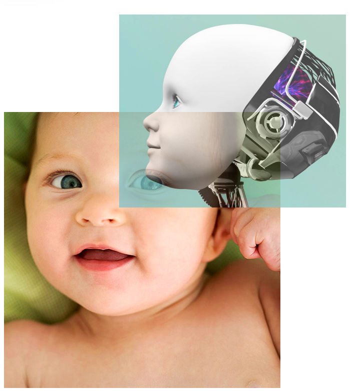
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

what a pre tty ki tty

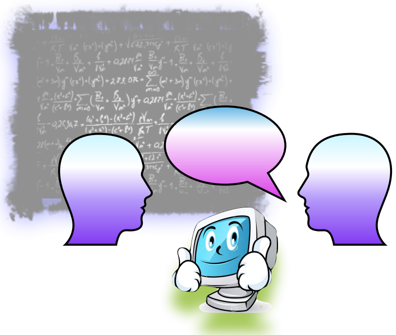
what
a
pretty
kitty

Bayesian inference

If so, then the computational cognitive model has captured (using this math) how a child could identify words in fluent speech.



Quantitative techniques for language development

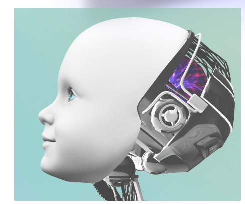


Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Bayesian inference

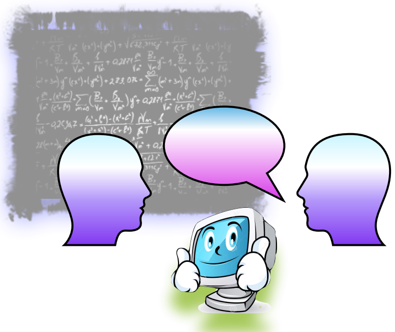
what a pre tty ki tty

what
a
pretty
kitty



It turns out this learning strategy for speech segmentation is useful.

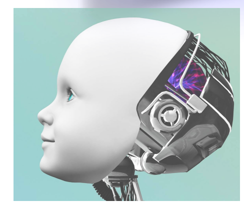
Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Bayesian inference

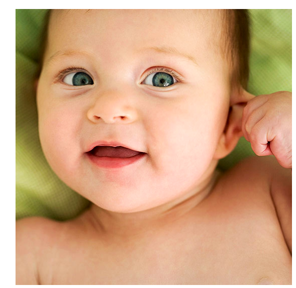
what a pre tty ki tty



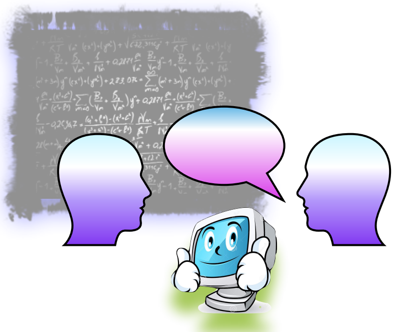
useful

what
a
prettykitty

It can segment realistic English input to children fairly well…though the inferred lexicons aren't perfect.



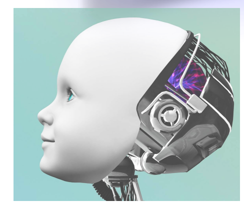
Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Bayesian inference

what a pre tty ki tty

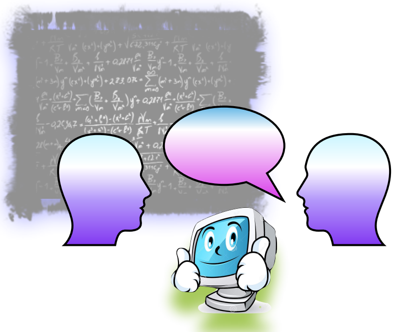


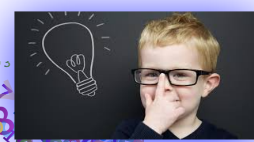
useful ✓

what
a
prettykitty

But it turns out these imperfect lexicons are very useful for subsequent stages of language development, like learning what a word form refers to.

Quantitative techniques for language development



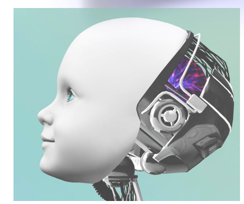


Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Bayesian inference

useful ✓

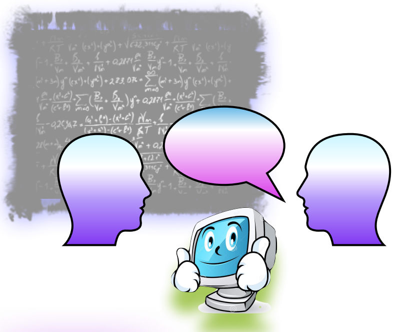
what a pre tty ki tty



what
a
prettykitty

This was true for idealized modeled children, with perfect memory and perfect processing abilities.

Quantitative techniques for language development



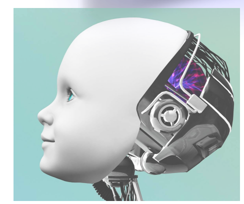
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Bayesian inference

useful ✓

what a pre tty ki tty

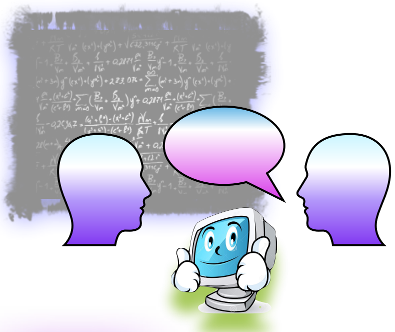
what
a
prettykitty



But what about modeled children with more realistic constraints on their memory and processing abilities?

Is this segmentation strategy useable by children, who have these kinds of cognitive limitations?

Quantitative techniques for language development



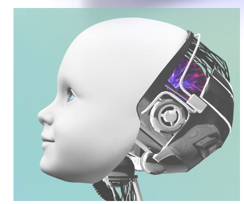
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Bayesian inference

what a pre tty ki tty

useful

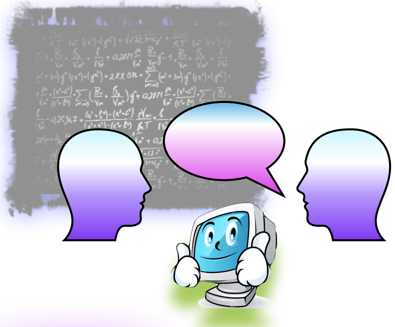
what
a
prettykitty



useable

Yes! Modeled children with cognitive constraints on their memory and processing abilities can still use this strategy to segment English quite well.

Quantitative techniques for language development

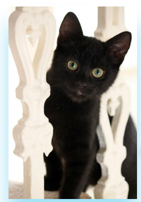


Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

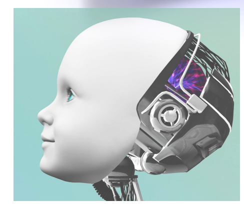
Bayesian inference

what a pre tty ki tty

useful



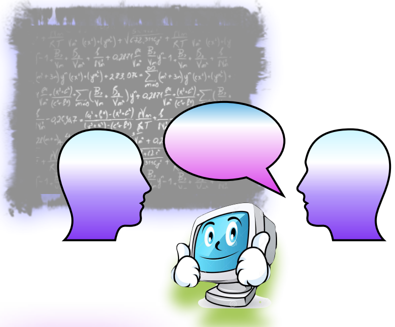
what
a
prettykitty



useable

Does it work for different languages (besides English)?

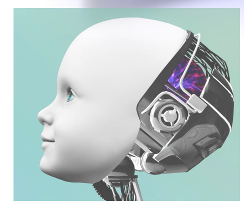
Quantitative techniques for language development



Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Bayesian inference

what a pre tty ki tty



useful ✓

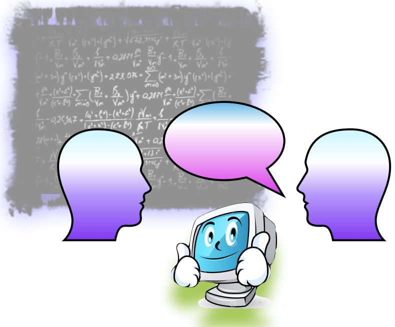
what
a
prettykitty

useable ✓

Yes! It segments well for languages with different properties: Spanish, Italian, German, Hungarian, Japanese, Farsi

different languages ✓

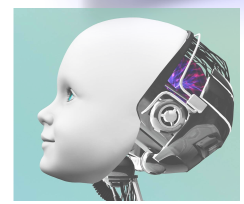
Quantitative techniques for language development

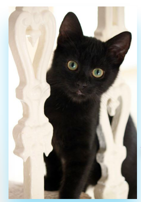


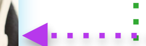
Phillips & Pearl 2012, 2014a, 2014b, 2015a, 2015b, Pearl & Phillips 2018

Bayesian inference

what a pre tty ki tty







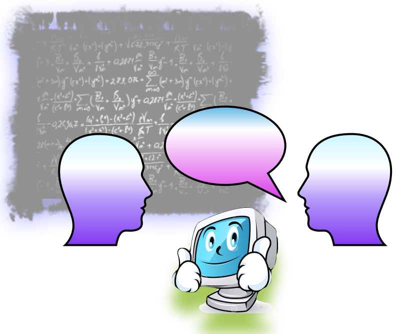
what
a
prettykitty

language

We were able to discover how good this segmentation strategy is by using computational cognitive modeling.

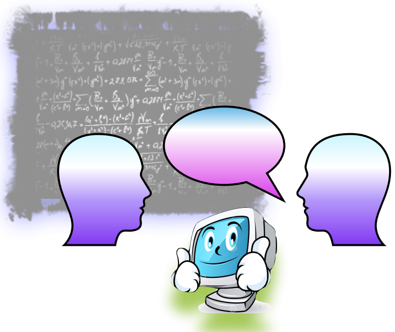
math

Quantitative techniques for language development



Some general questions about language development that my research has tried to answer this way

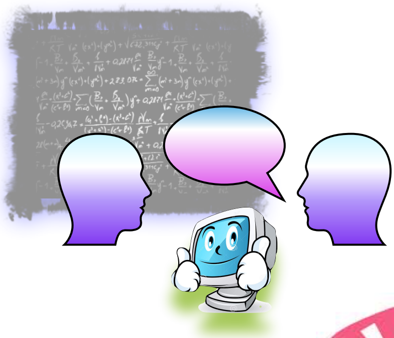
Quantitative techniques for language development



Which learning strategies could children be using?

(Bayes & Pearl under review, Pearl 2021, Forsythe & Pearl 2019, Bates & Pearl 2019, Nguyen & Pearl 2019, Phillips & Pearl 2018, Pearl 2017, Bar-Sever & Pearl 2016, Phillips & Pearl 2015a, 2015b, 2014a, 2014b, 2012; Pearl 2014, Pearl et al. 2011, Pearl et al. 2010)

Quantitative techniques for language development

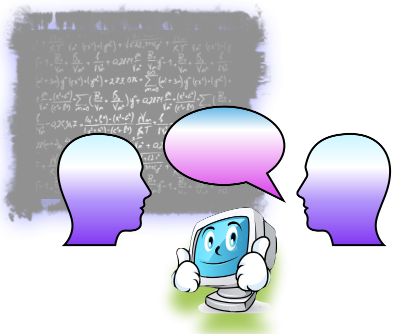


Which learning strategies could children be using?

Which learning biases are necessary?

(Pearl 2021, Pearl & Sprouse 2019, Nguyen & Pearl 2019, Pearl, Ho, & Detrano 2017, 2014; Pearl & Mis 2016, Pearl & Sprouse 2015, 2013a, 2013b, Pearl & Mis 2011, Pearl & Lidz 2009, Pearl 2008, Pearl & Weinberg 2007)

Quantitative techniques for language development



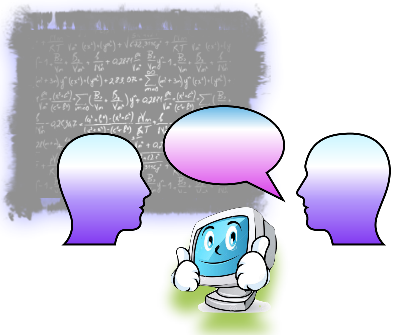
Which learning strategies could children be using?

Which learning biases are necessary?

Which knowledge representations are learnable — and which aren't?

(Pearl & Sprouse in press, Bates & Pearl 2019, Pearl, Ho, & Detrano 2017, 2014; Pearl 2017, Pearl 2011, Pearl 2009)

Quantitative techniques for language development



Which learning strategies could children be using?

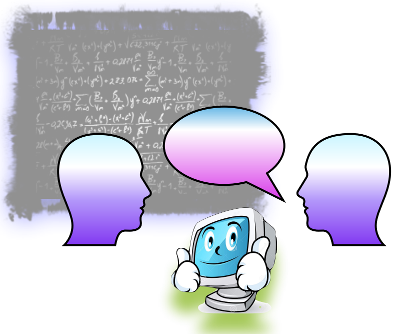
Which learning biases are necessary?

Which knowledge representations are learnable — and which aren't?

When do children learn different aspects of the linguistic system?

(Nguyen & Pearl under review, Bates, Pearl, & Braunwald 2018, Savinelli, Scontras, & Pearl 2018, Bar-Sever, Lee, Scontras, & Pearl 2018, Savinelli, Scontras, & Pearl 2017, Nguyen & Pearl 2018, Caponigro, Pearl et al. 2012, Caponigro, Pearl et al. 2011)

Quantitative techniques for language development



Which learning strategies could children be using?

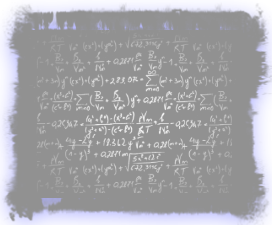
Which learning biases are necessary?

Which knowledge representations are learnable — and which aren't?

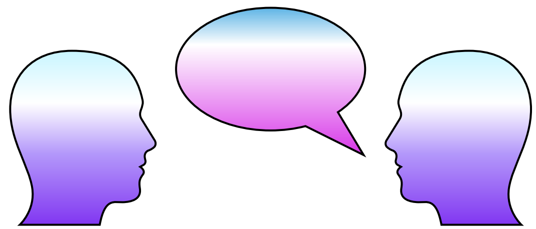
When do children learn different aspects of the linguistic system?

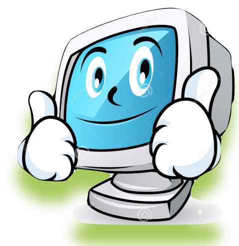
What factors affect children's observable behavior?

(Scontras & Pearl under review, Nguyen & Pearl under review, Forsythe & Pearl 2019, Nguyen & Pearl 2019, Nguyen & Pearl 2018, Savinelli, Scontras, & Pearl 2018, Nguyen & Pearl 2017, Savinelli, Scontras, & Pearl 2017)



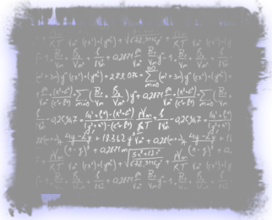
Quantitative techniques for language use

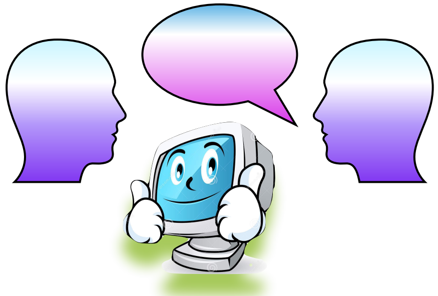




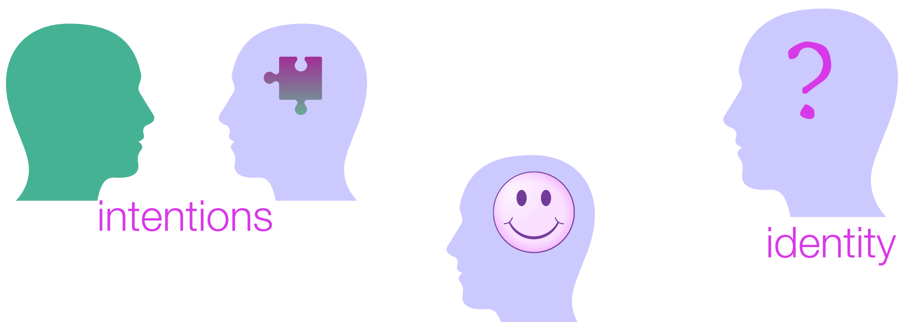
What about using math to help us understand how humans use language and how machines could learn to do the same thing?

Quantitative techniques for language use





Let’s focus on subtle information that can be expressed in language.

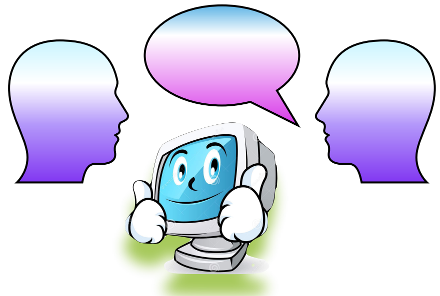


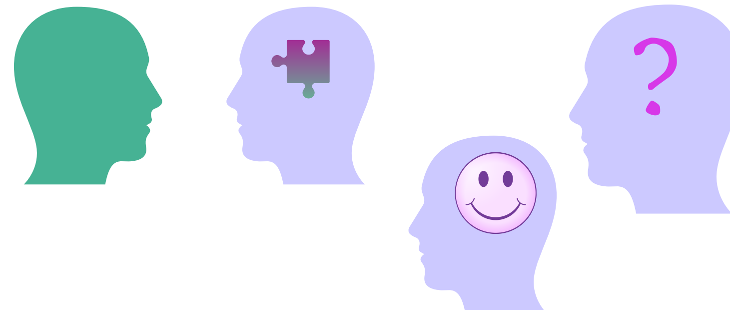
intentions

identity

emotions/attitudes

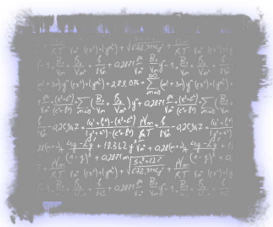
Quantitative techniques for language use





Math is again at the heart of the techniques researchers use.

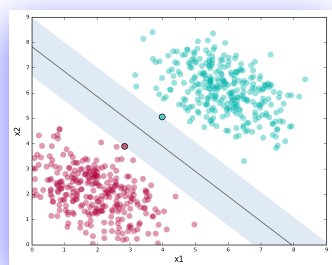
computational tools



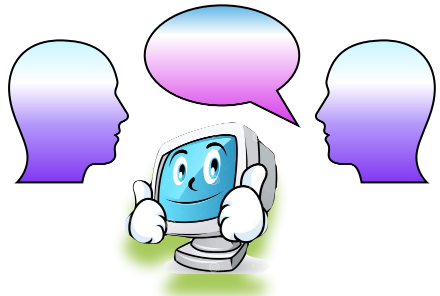
mathematical knowledge

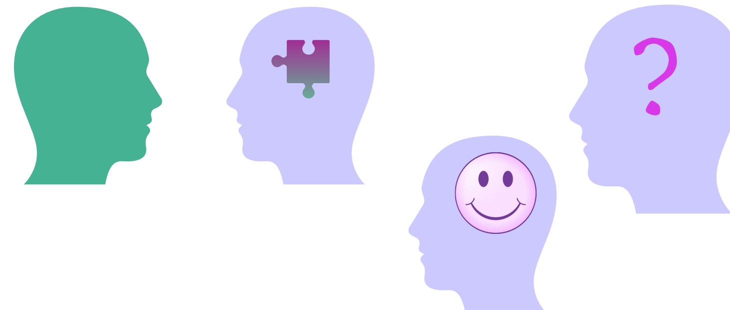
$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

machine learning techniques



Quantitative techniques for language use

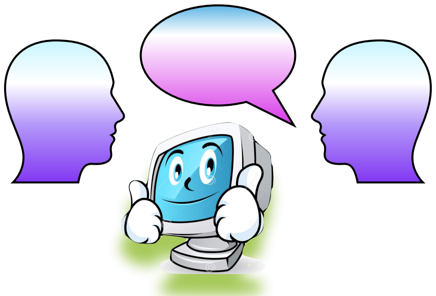


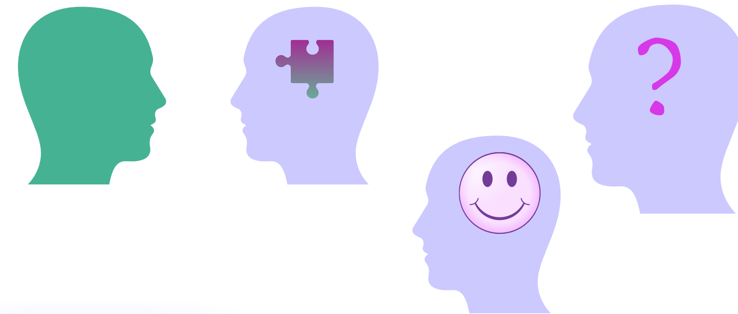


$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

So, we’re counting things and reasoning about those counts in principled ways.

Quantitative techniques for language use

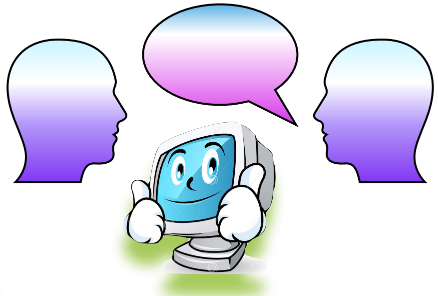


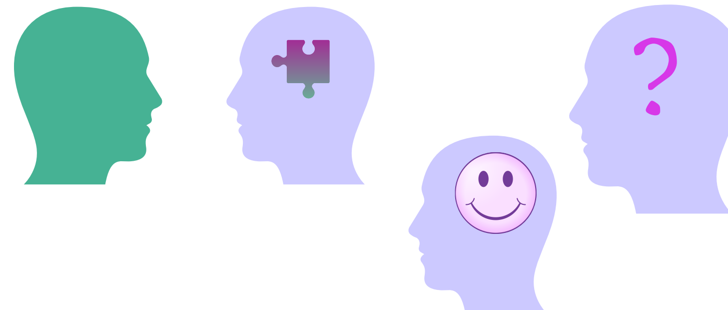


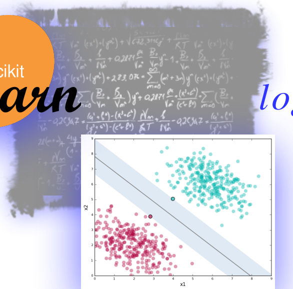
$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

But what are we counting?

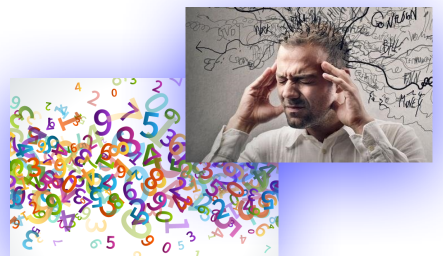
Quantitative techniques for language use





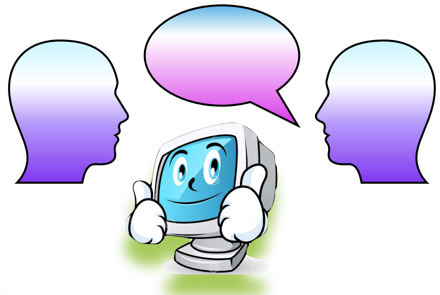


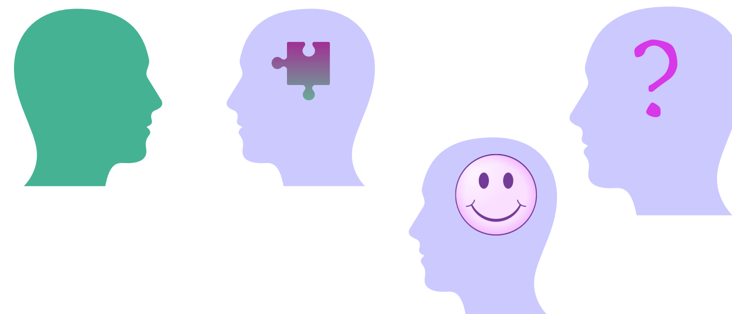
$$log(\frac{p(f_v|T)}{p(f_v|D)})$$



precise psychological and
linguistic theoretical constructs

Quantitative techniques for language use

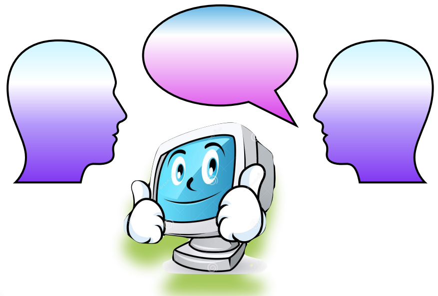


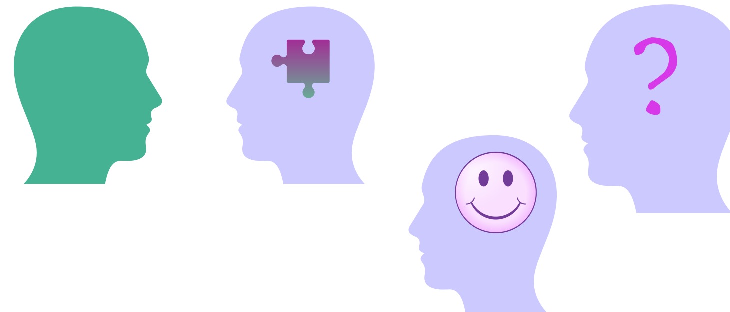


$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

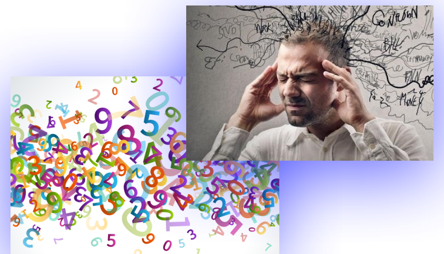
It's just two minutes away from Hyde Park.

Quantitative techniques for language use





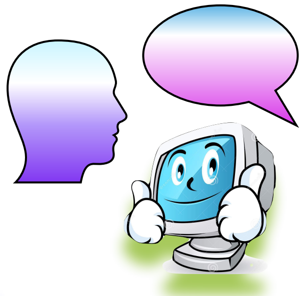
$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

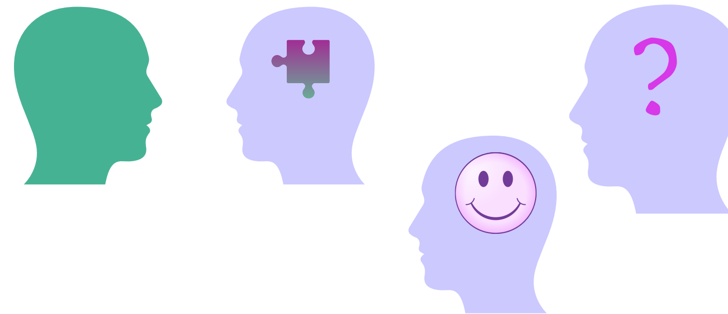


It's just two minutes away from Hyde Park.

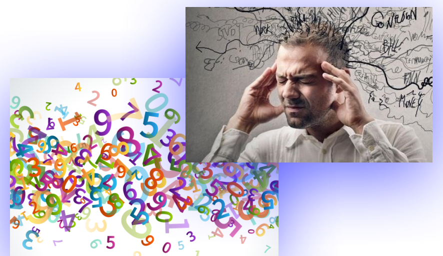
Psychological: "specific details" in a description

Quantitative techniques for language use





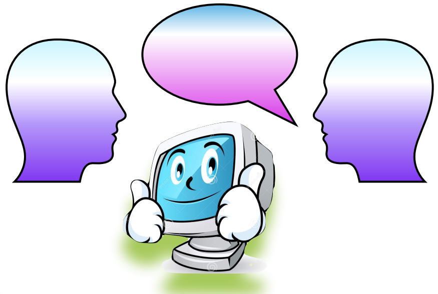
$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

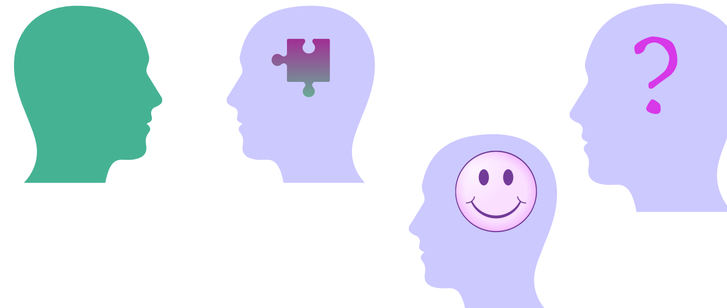


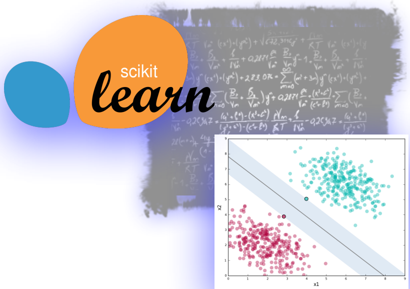
It's just two minutes away from Hyde Park.

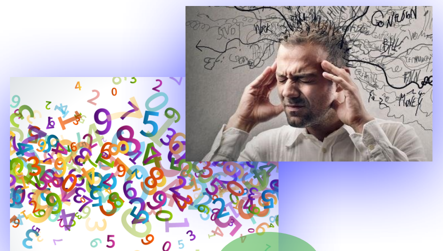
Linguistic: prepositional phrases

Quantitative techniques for language use





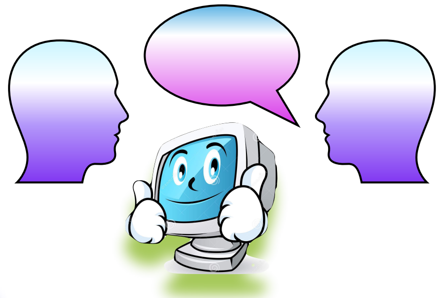


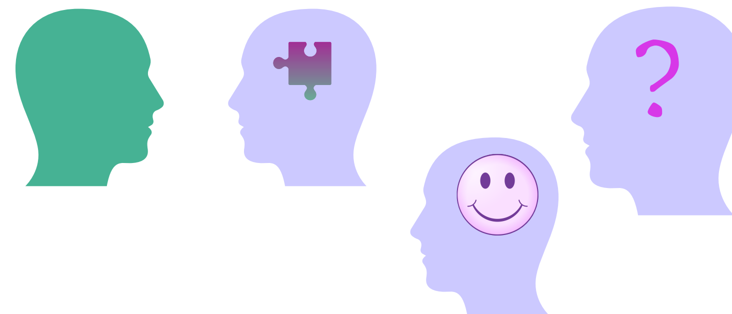


The counts of these features are used in mathematical equations

$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

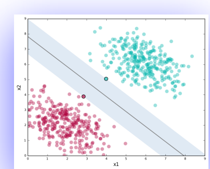
Quantitative techniques for language use



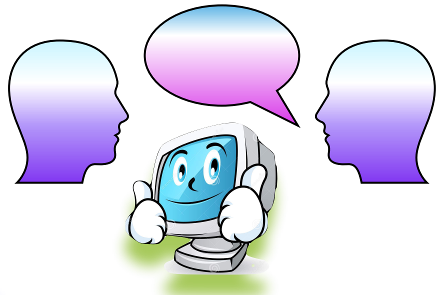


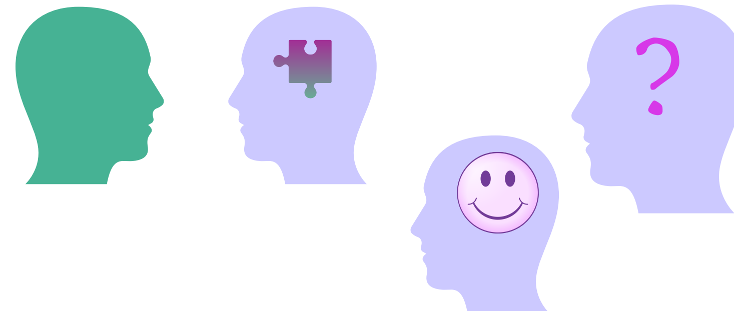
$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

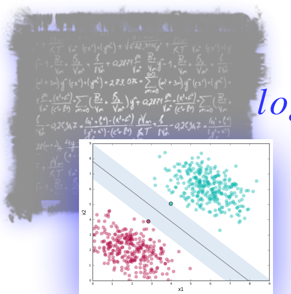
The counts of these features are used in mathematical equations that underlie machine learning techniques



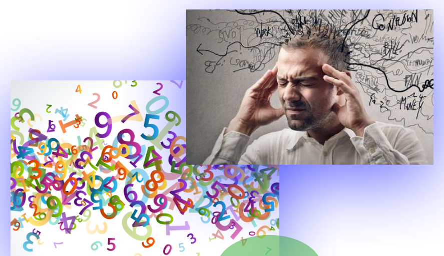
Quantitative techniques for language use





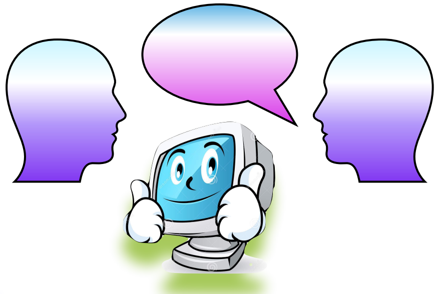


$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

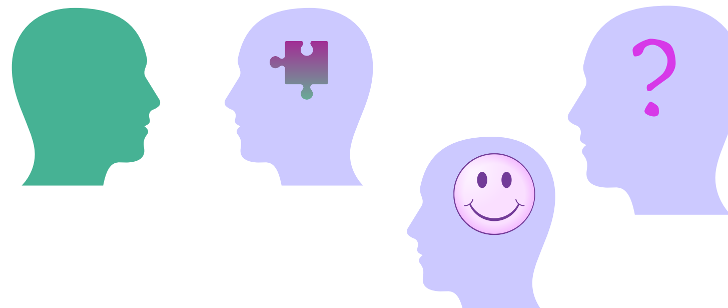


The counts of these features are used in mathematical equations that underlie machine learning techniques available in common computational tools.

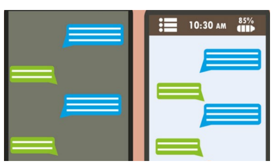
Quantitative techniques for language use



$$log(\frac{p(f_v|T)}{p(f_v|D)})$$



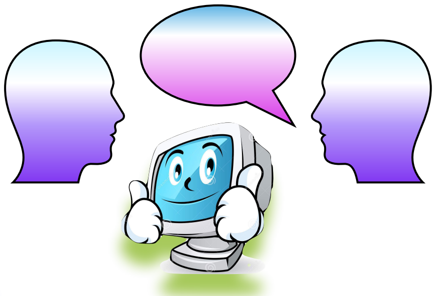
Some discoveries from CoLaLab about subtle information in language text alone

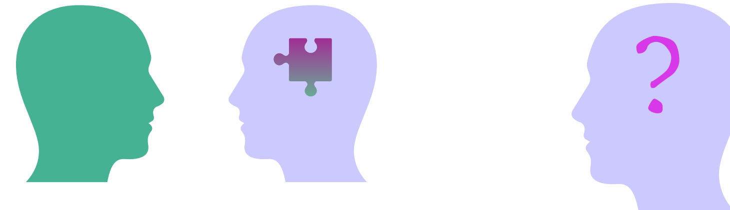


electronic
(more conversational)

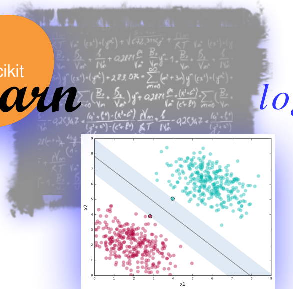
written text

Quantitative techniques for language use





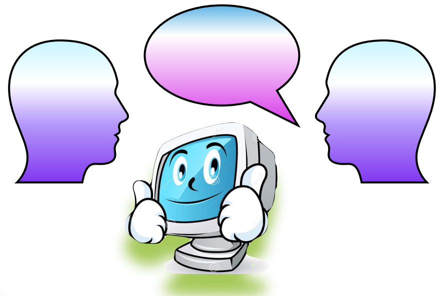
scikit learn

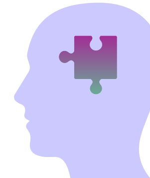


$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

Pearl & Steyvers 2013, Pearl & Enverga 2015:
Detecting emotions, attitudes, and intentions
in short messages

Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

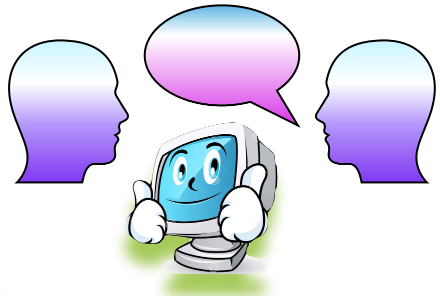
Pearl & Steyvers 2013, Pearl & Enverga 2015:
Detecting emotions, attitudes, and intentions
in short messages

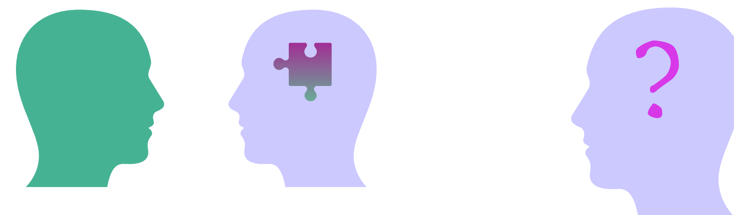
What features we counted:
n-grams (strings of n units)
that abstracted across
linguistic constructs

the+best
the+brightest
the+most+fantastic
the+most+fun

*the+*POSITIVE-ADJECTIVE-IN-THE-SUPERLATIVE

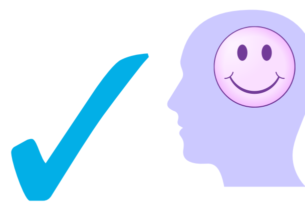
Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

Pearl & Steyvers 2013, Pearl & Enverga 2015:
Detecting emotions, attitudes, and intentions
in short messages

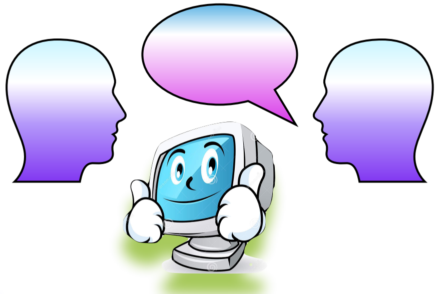


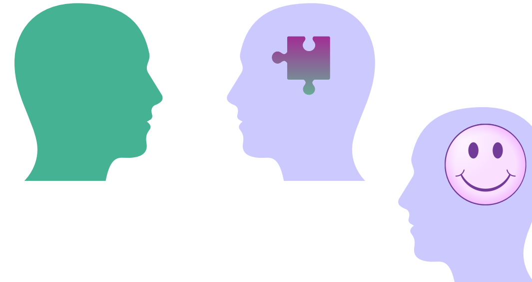
We then used standard computational tools to achieve better accuracy than previous approaches.

the+best
the+brightest
the+most+fantastic
the+most+fun

the+POSITIVE-ADJECTIVE-IN-THE-SUPERLATIVE

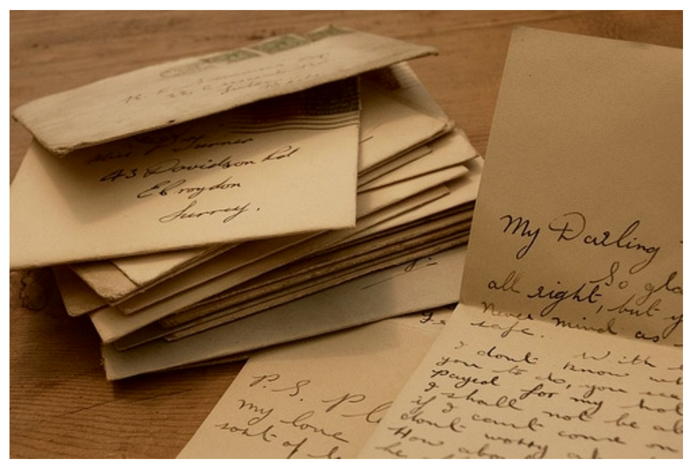
Quantitative techniques for language use



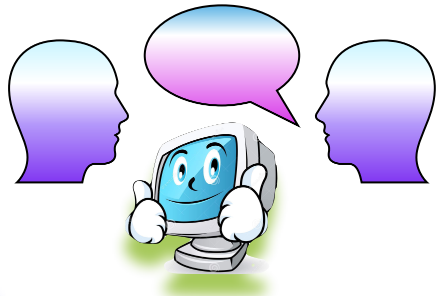


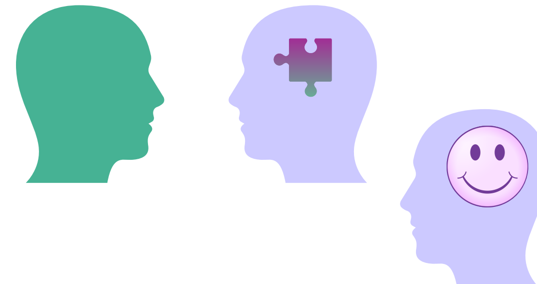
$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

Pearl, Lu, & Haghighi 2016: Authorship in epistolary novels — can one person (the author) really write in the style of multiple other people?

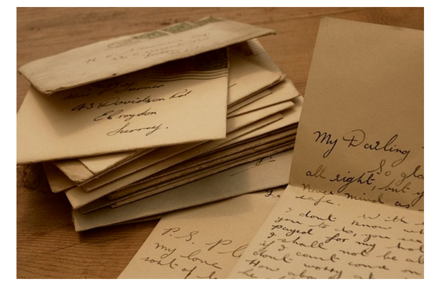


Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

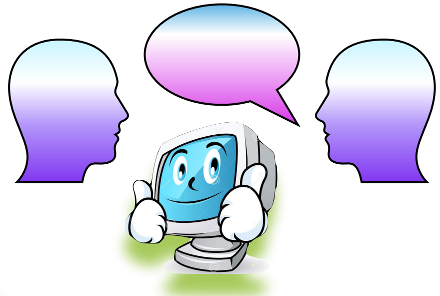


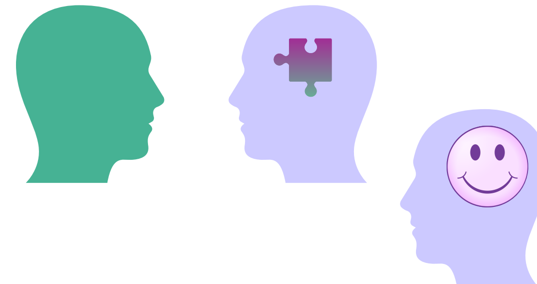
Pearl, Lu, & Haghighi 2016: Authorship in epistolary novels — can one person (the author) really write in the style of multiple other people?

✓ ✗

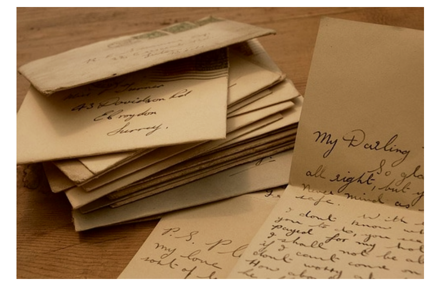
Answer: Yes and no.
The features the author manipulated (which did create several fairly distinct characters) weren't the ones that signified his own style. His own style features were still present.

Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

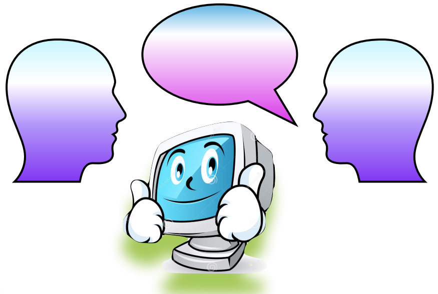


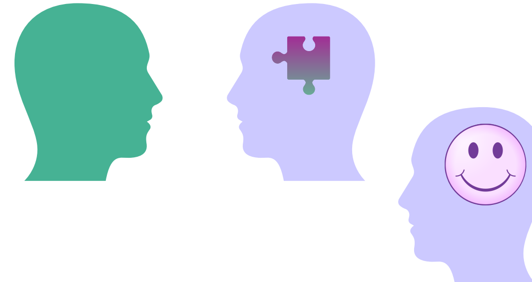
Pearl, Lu, & Haghighi 2016: Authorship in epistolary novels — can one person (the author) really write in the style of multiple other people? ✓✗

"I HAVE both your letters at once. It is very unhappy, my dear, … -And why? Shall I venture to tell you?- Because they are nearer…"

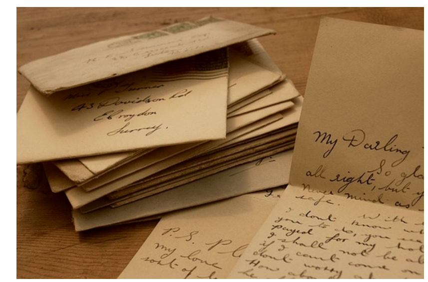
What features we counted:
character-level (like punctuation)

Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$



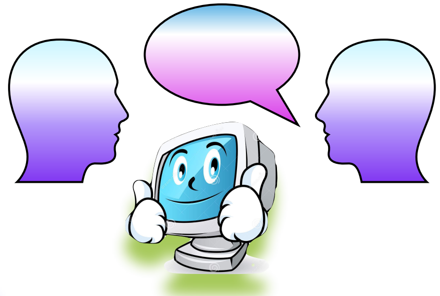
Pearl, Lu, & Haghighi 2016: Authorship in epistolary novels — can one person (the author) really write in the style of multiple other people? ✓✗

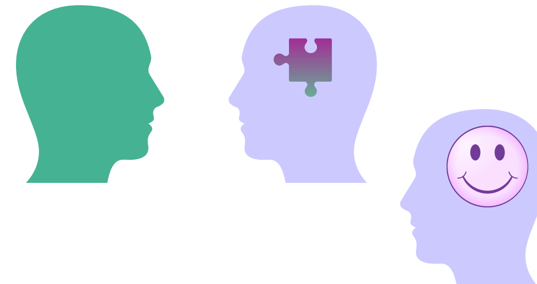
What features we counted:
word-level (like total words)

25

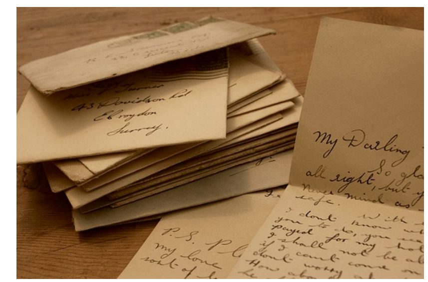
"I HAVE both your letters at once. It is very unhappy, my dear, … -And why? Shall I venture to tell you? - Because they are nearer…"

Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

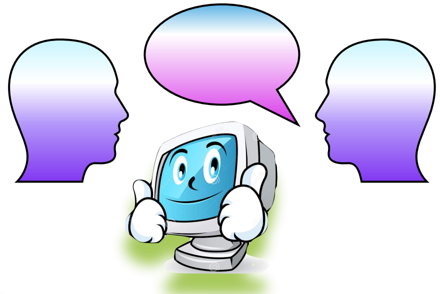


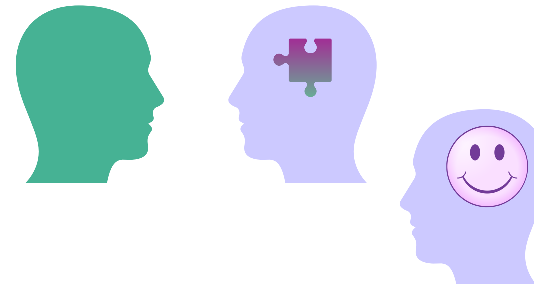
Pearl, Lu, & Haghighi 2016: Authorship in epistolary novels — can one person (the author) really write in the style of multiple other people?

What features we counted:
syntactic (like first person pronouns)

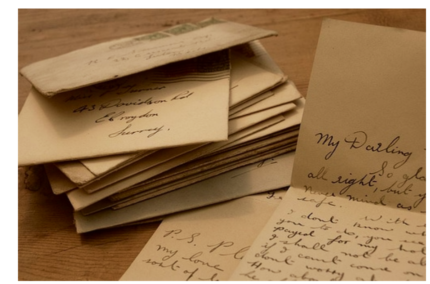
"I HAVE both your letters at once. It is very unhappy, my dear, ... -And why? Shall I venture to tell you? - Because they are nearer..."

Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

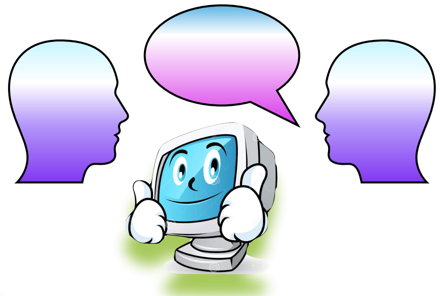


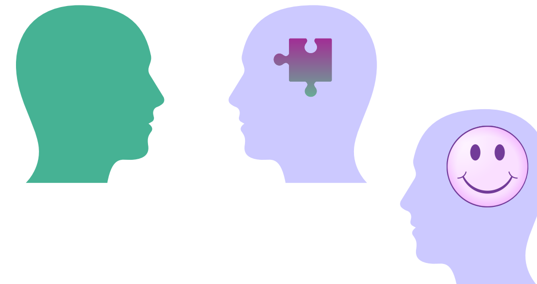
Pearl, Lu, & Haghighi 2016: Authorship in epistolary novels — can one person (the author) really write in the style of multiple other people? ✓✗

What features we counted:
semantic (like endearments)

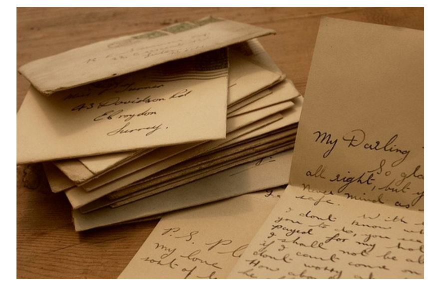
"I HAVE both your letters at once. It is very unhappy, my dear, … -And why? Shall I venture to tell you? - Because they are nearer…"

Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

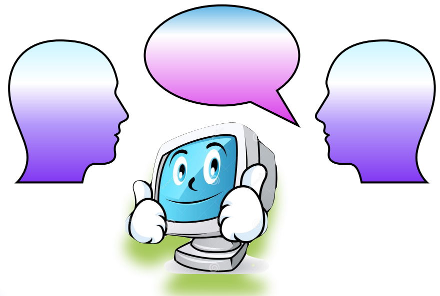


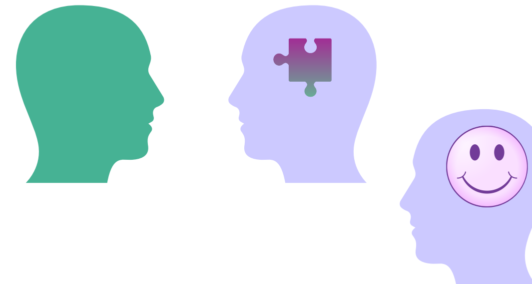
Pearl, Lu, & Haghighi 2016: Authorship in epistolary novels — can one person (the author) really write in the style of multiple other people? ✓✗

What features we counted:
formatting (like all capitals)

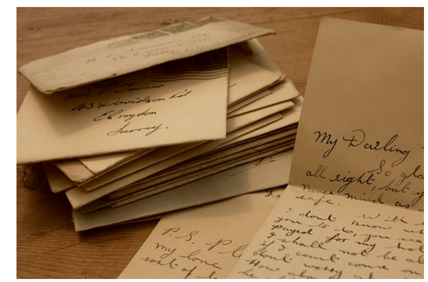
"I HAVE both your letters at once. It is very unhappy, my dear, … -And why? Shall I venture to tell you? - Because they are nearer…"

Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

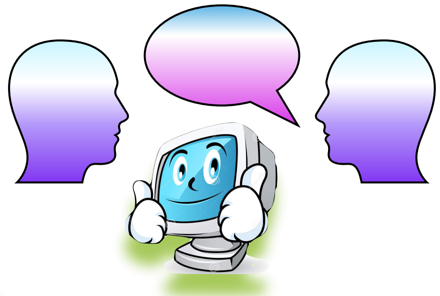


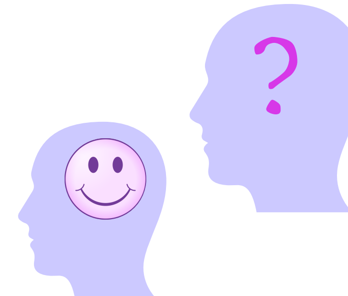
Pearl, Lu, & Haghighi 2016: Authorship in epistolary novels — can one person (the author) really write in the style of multiple other people? ✓✗

We then used standard computational tools to determine the stylistic components that were distinct from the author's own style vs. those that were alike.

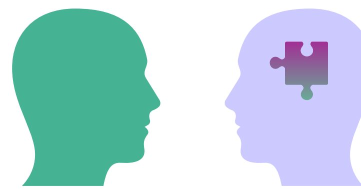
"I HAVE both your letters at once. It is very unhappy, my dear, … -And why? Shall I venture to tell you? - Because they are nearer…"

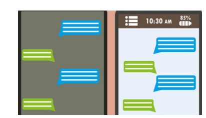
Quantitative techniques for language use





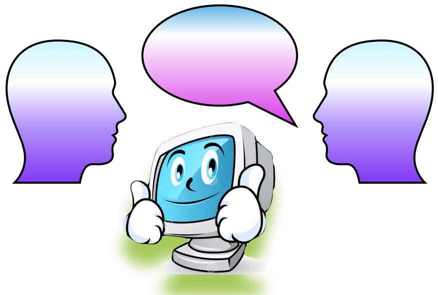
$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

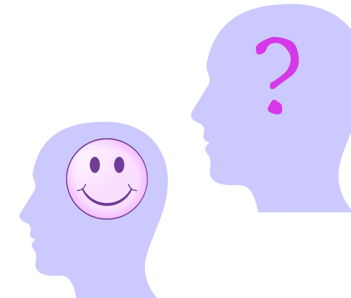


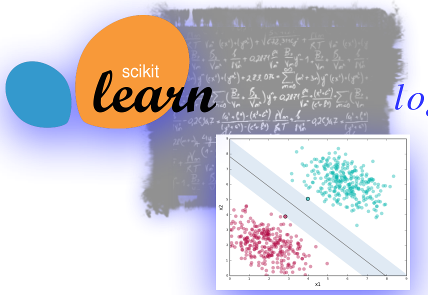


Vogler & Pearl 2019: Can we more accurately detect deception across different content domains, like online product reviews, short opinion essays, and transcripts of job interviews?

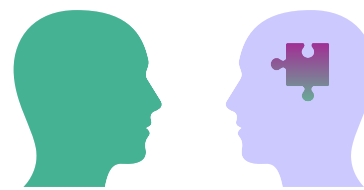
Quantitative techniques for language use







$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

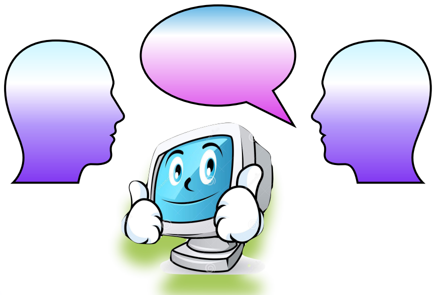


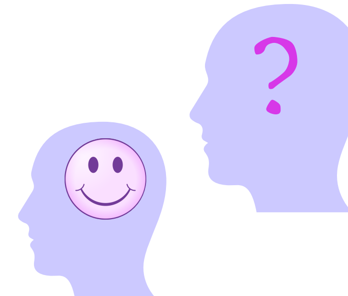
Vogler & Pearl 2019: Can we more accurately detect deception across different content domains, like online product reviews, short opinion essays, and transcripts of job interviews?

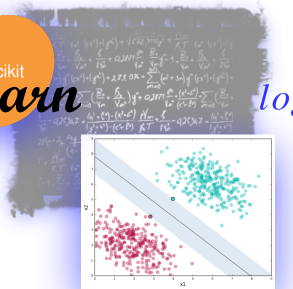
Answer: Yes.

When the content (and form) of the language text changes a lot from sample to sample, we can do much better if we use features that are linguistically-defined and also capture the psychological idea of “specific details”.

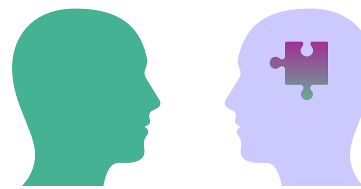
Quantitative techniques for language use







$$log(\frac{p(f_v|T)}{p(f_v|D)})$$

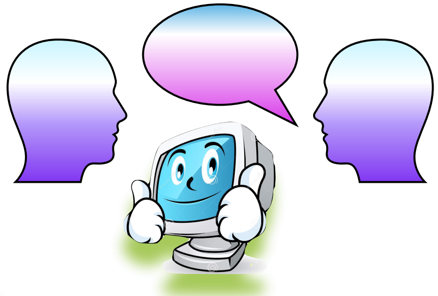


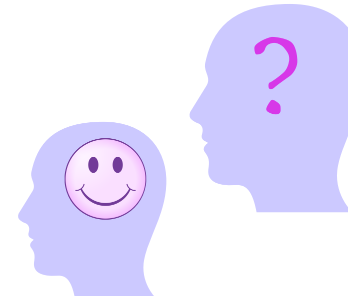
Vogler & Pearl 2019: Can we more accurately detect deception across different content domains, like online product reviews, short opinion essays, and transcripts of job interviews? ✓

What features we counted:
Specific details expressed with linguistic constructions like exact numbers and prepositional phrases.

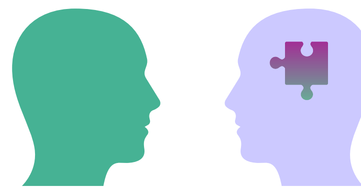
It's just two minutes away from Hyde Park.

Quantitative techniques for language use





$$log(\frac{p(f_v|T)}{p(f_v|D)})$$



Vogler & Pearl 2019: Can we more accurately detect deception across different content domains, like online product reviews, short opinion essays, and transcripts of job interviews? ✓

We then used standard computational tools to achieve better detection than previous approaches.

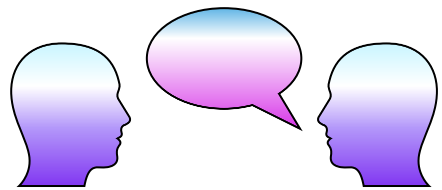
It's just two minutes away from Hyde Park.

How math helps us better understand language

Math is at the heart of quantitative techniques, where we count things and reason about those counts.

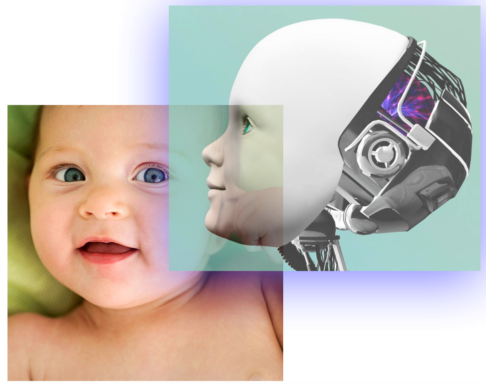
How math helps us better understand language

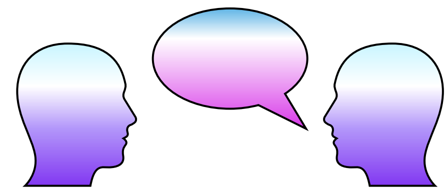
We can use quantitative techniques to better understand many different questions about the utterly human capacity of language.



How math helps us better understand language

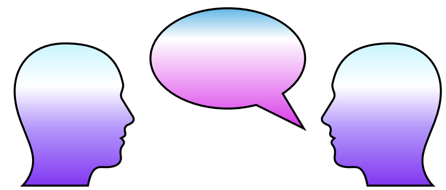
In language development, quantitative techniques
are used in computational cognitive modeling.





How math helps us better understand language

In language use, quantitative techniques underlie common computational tools that are used in combination with insights from psychology and linguistics.



So let's keep using math to help us better understand language!

Thank you!

Lisa S. Pearl
Professor
Department of Language Science
SSPB 2219
University of California, Irvine
lpearl@uci.edu

Computation of Language Laboratory

UC Irvine