

Today, we'll talk a bit about

Language,

Using computers to work with language, and

Course logistics.

aisapps.vassar.edu/cgi-bin/courses.cgi

COURSE ID	TITLE	UNITS	SP	MAX	ENR	AVL	WL	GMOD	YL	PR	FR	LA	QA	PREREQ	FORMAT	DIV	DEPT	XLIST	DAYS	TIME	LOCATION	INS
CMPU-100-01	Programming with Data	1.0		20	15	5	0	NR					QA		CLS	NS	CMPU		MW	1030AM-1145AM	SC 006	Gor
CMPU-100-02	Programming with Data	1.0		20	15	5	0	NR					QA		CLS	NS	CMPU		R	0100PM-0300PM	SC 006	
CMPU-100-03	Programming with Data	1.0		20	15	5	0	NR					QA		CLS	NS	CMPU		MW	0900AM-1015AM	SP 206	Sta
CMPU-101-01	Problem-Solving & Abstraction	1.0		20	19	1	0	NR					QA		CLS	NS	CMPU		R	0900AM-1100AM	SC 006	
CMPU-101-02	Problem-Solving & Abstraction	1.0		20	12	8	0	NR					QA		CLS	NS	CMPU		MW	1200PM-0115PM	SP 309	Eri
CMPU-101-03	Problem-Solving & Abstraction	1.0		20	13	7	0	NR					QA		CLS	NS	CMPU		R	1100AM-0100PM	SC 006	
CMPU-102-01	Data Structures and Algorithms	1.0		24	13	11	0	NR					QA	Y	CLS	NS	CMPU		MW	0900AM-1015AM	SC 006	Smi
CMPU-102-02	Data Structures and Algorithms	1.0		24	5	19	0	NR					QA	Y	CLS	NS	CMPU		F	0900AM-1100AM	SC 006	
CMPU-145-01	Foundations/Computer Science	1.0		24	3	21	0	NR					QA	Y	CLS	NS	CMPU		MW	1200PM-0115PM	SC 006	Ell
CMPU-145-02	Foundations/Computer Science	1.0		24	8	16	0	NR					QA	Y	CLS	NS	CMPU		F	1100AM-0100PM	SC 006	
CMPU-203-01	Software Design	1.0		20	20	0	0						QA	Y	CLS	NS	CMPU		MW	0130PM-0245PM	SC 006	Lem
CMPU-224-01	Computer Organization	1.0		24	26	-1	0						Y		CLS	NS	CMPU		F	0100PM-0300PM	SC 006	
CMPU-240-01	Theory of Computation	1.0		24	23	1	0						Y		CLS	NS	CMPU		MW	0130PM-0245PM	SP 309	Hun
CMPU-240-02	Theory of Computation	1.0		24	8	16	0						Y		CLS	NS	CMPU		R	1030AM-1145AM	NE 105	
CMPU-241-01	Analysis of Algorithms	1.0		24	24	0	0						Y		CLS	NS	CMPU		MW	0900AM-1015AM	SP 309	Lem
CMPU-250-01	Data and its Discontents	1.0		24	16	8	0						QA	Y	CLS	NS	CMPU		F	0900AM-1100AM	SP 309	
CMPU-290-01	Community-Engaged Learning	0.5	SP	999	0	999	0	SU							OTH	NS	CMPU		MW	0130PM-0245PM	SP 309	Hun
CMPU-290-02	Community-Engaged Learning	1.0	SP	999	0	999	0	SU							OTH	NS	CMPU		R	1030AM-1145AM	NE 105	
CMPU-298-01	Independent Work	0.5	SP	999	0	999	0	SU							OTH	NS	CMPU					
CMPU-298-02	Independent Work	1.0	SP	999	0	999	0	SU							OTH	NS	CMPU					
CMPU-319-01	Modeling Minds,Brains,Behavior	1.0		20	22	-2	0	NR					Y		CLS	NS	COGS	XCOGS	TR	1200PM-1250PM	OH 266	de
CMPU-334-01	Operating Systems	1.0		24	24	2	0						Y		CLS	NS	CMPU		F	0310PM-0510PM	BLS 101	
CMPU-353-01	Bioinformatics	1.0	SP	16	16	0	0								CLS	NS	CMPU	XBIOL	TR	0310PM-0425PM	SP 309	Wat
CMPU-366-01	Computational Linguistics	1.0					0						QA		CLS	NS	CMPU					
CMPU-375-01	Computer Networks	1.0					0						Y		CLS	NS	CMPU		TR	1200PM-1250PM	OH 266	de
CMPU-399-01	Senior Independent Work	0.5	SP	999	0	999	0	SU							OTH	NS	CMPU		F	0310PM-0510PM	BLS 101	
CMPU-399-02	Senior Independent Work	1.0	SP	999	0	999	0	SU							OTH	NS	CMPU		R	1030AM-1145AM	SP 105	

Ask Banner Home Page

catalogue.vassar.edu/content.php?catoid=60&navoid=12464

- [CMPU 353 - Bioinformatics](#)
- [CMPU 365 - Artificial Intelligence](#)

[CMPU 366 - Natural Language Processing](#)

CMPU 366 - Natural Language Processing ★ 🔗 🖨️

Semester Offered: Fall
1 unit(s)

This course examines the intersection of human language and computation, focusing on how computers understand, analyze, and generate natural language. Students explore both traditional techniques (e.g., n-grams and rule-based systems) and modern methods (e.g., neural networks and deep learning) for language modeling and manipulation. Through hands-on projects, they apply these approaches to tasks such as information extraction, text classification, machine translation, and language generation. The course also addresses the ethical implications and societal impacts of language technologies, encouraging critical reflection on their development and use. Jonathan Gordon.

Prerequisite(s): [CMPU 250](#) or [MATH 244](#), or permission of the instructor.

Two 75-minute periods.

Course Format: CLS

- [CMPU 375 - Computer Networks](#)
- [CMPU 377 - Parallel Programming](#)
- [CMPU 378 - Graphics](#)
- [CMPU 395 - Advanced Special Topics](#)

Also me.

Computational Linguistics

Computational Linguistics

Studying language

Using computers

Computational Linguistics

Studying language

Natural Language Processing

Human language

Natural Language Processing

Human language

Natural Language Processing

Doing stuff with language data

We'll do a bit of each – but the emphasis will really be on NLP.

There are around 7,000 living human languages,
distributed across 128 language families.

These can be spoken or signed – and may also be
written down.



Disappearing Languages

The way we speak has an enormous effect on our identities. The ever-changing languages, dialects, or even accents that we use define our relationship with the world, allowing us to express concepts and ideas that might be impossible in other languages.

Iktsuarpok, for example, is an Inuit word for the feeling of anticipation when you're expecting someone that leads you to constantly check to see if they've arrived, while *hiraeth* is a Welsh word that reflects a type of longing for the homeland or the romanticized past. The Japanese word *komorebi* refers to the sunlight that filters through the leaves of trees, while the Yaghan word *mamihlapinatapai* describes the wordless, meaningful look shared by two people who both want to initiate something, but are reluctant to do so.

2469

languages are at risk of extinction around the world

Natural language processing isn't just about English!

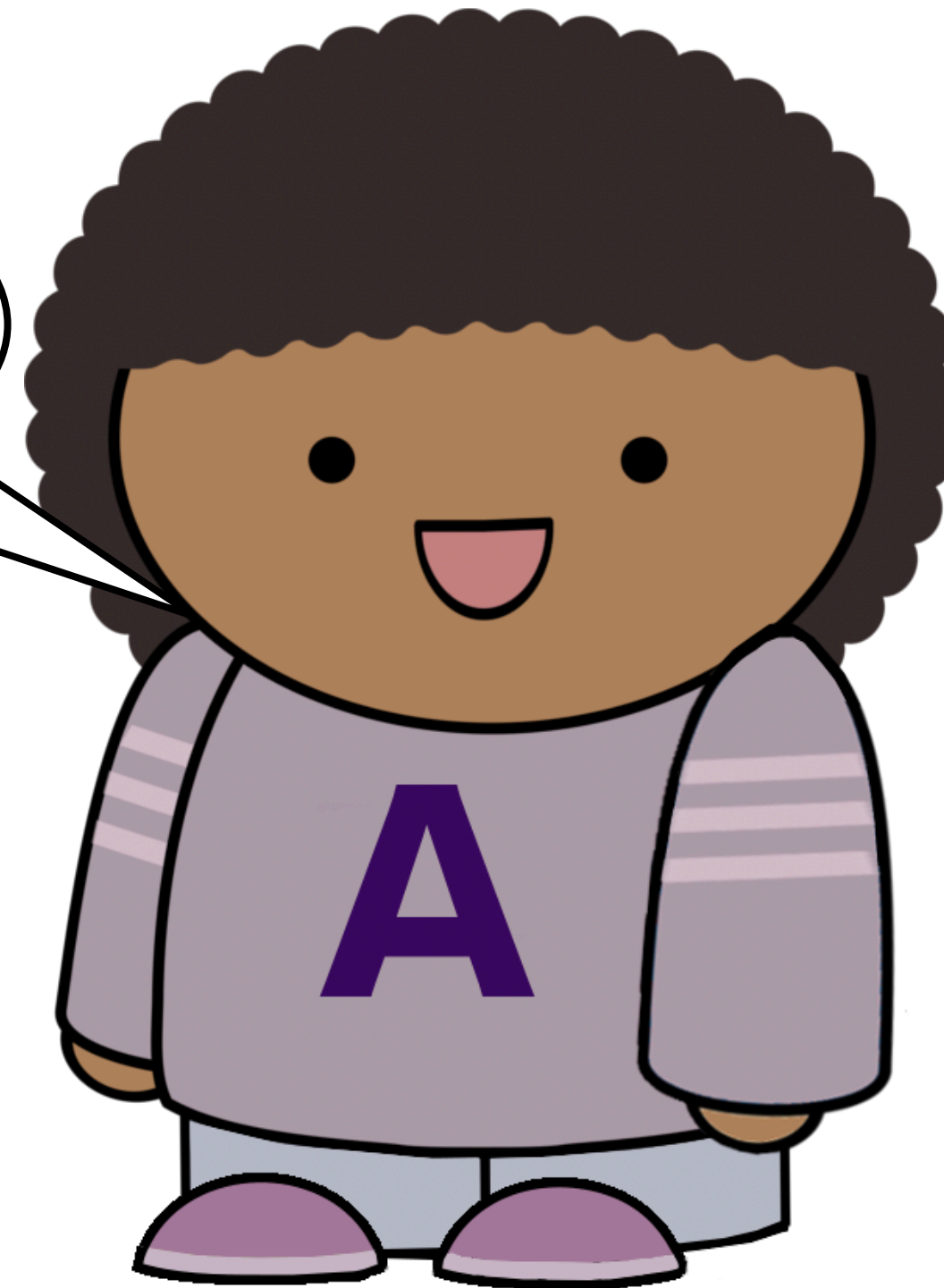
But we're going to talk about English a lot – why is that?

Layers of linguistic representation



Wanna go get
ice cream?

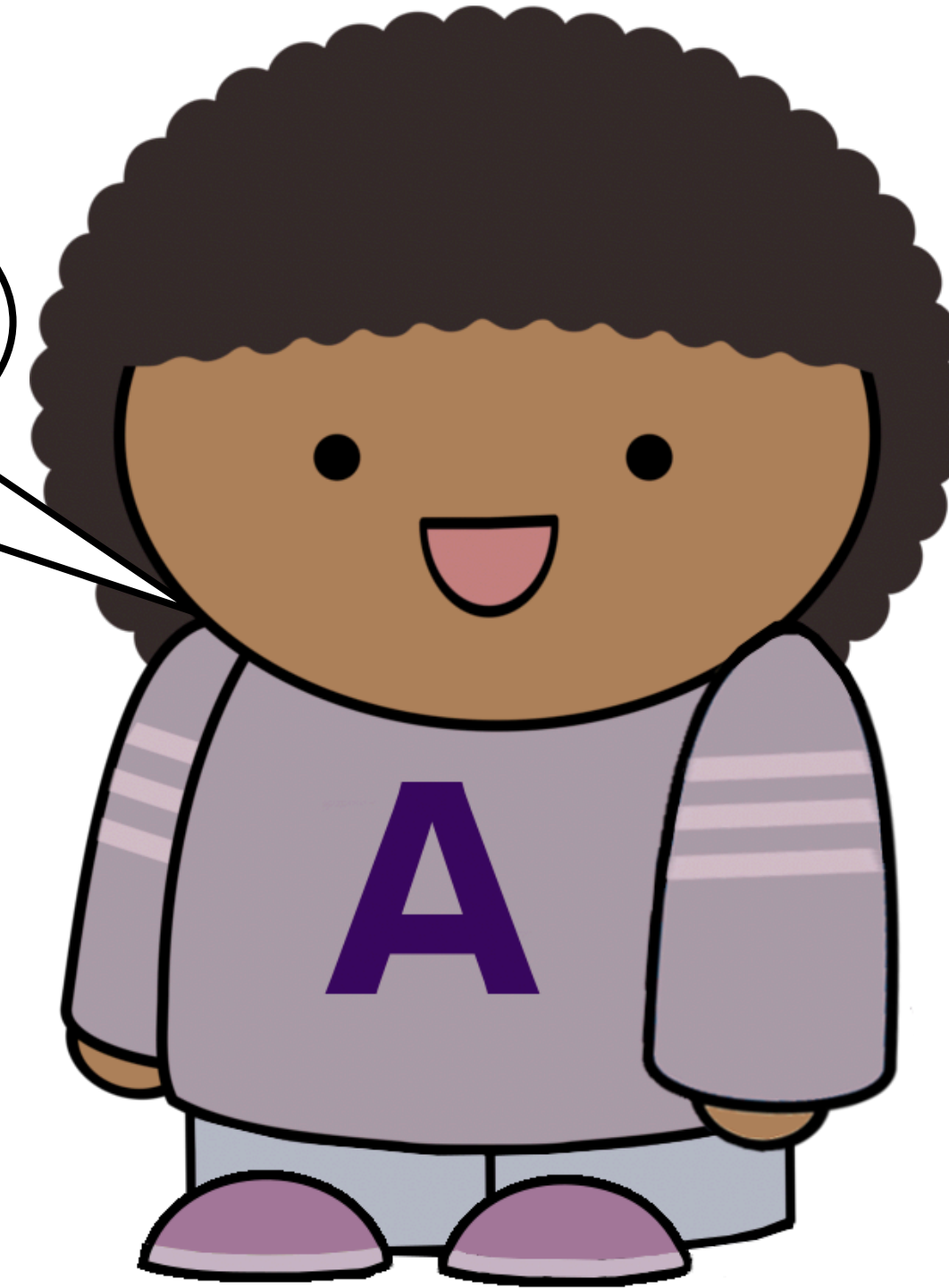
It's 7 am.





*Wanna go get
ice cream?*

It's 7 am.



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

B asks A a yes–no question,
but A does not respond with yes or no.



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

A *literally* says what time it is.



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

A *implies* that it's too early for ice cream.



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

Pragmatics: The meaning of sequences of sentences.



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

How do we know what B's question means?



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

What does the sequence of words *wanna go get ice cream* mean?



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

Genuinely too complicated to explain here!
(Go to graduate school in linguistics.)



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

What does the sequence of words *it's 7 am* mean?



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

[[*it's 7 am*]] = NOW(7am)



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

$$[[it's\ 7\ am]]_{c,w} = \begin{cases} \text{True} & \text{if } w_t = 7am \\ \text{False} & \text{otherwise} \end{cases}$$



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

Basically: *it's 7 am* is a *function* that takes a *world* and returns true for some worlds and false for others.



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

Semantics: the meaning of a sentence is its truth conditions (i.e., the conditions under which it is true!)



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

How do we determine the order of the words?



Pragmatics

Semantics

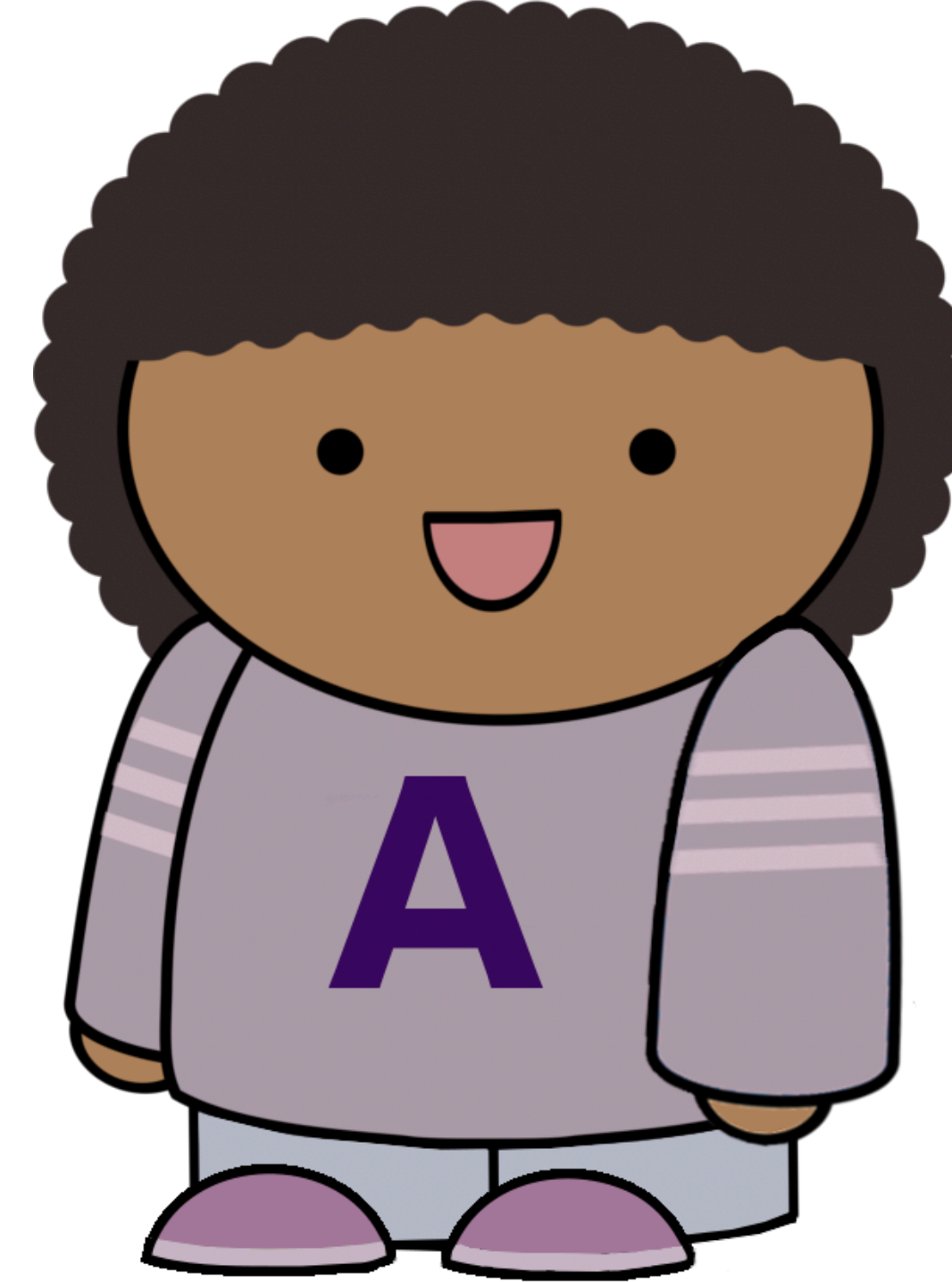
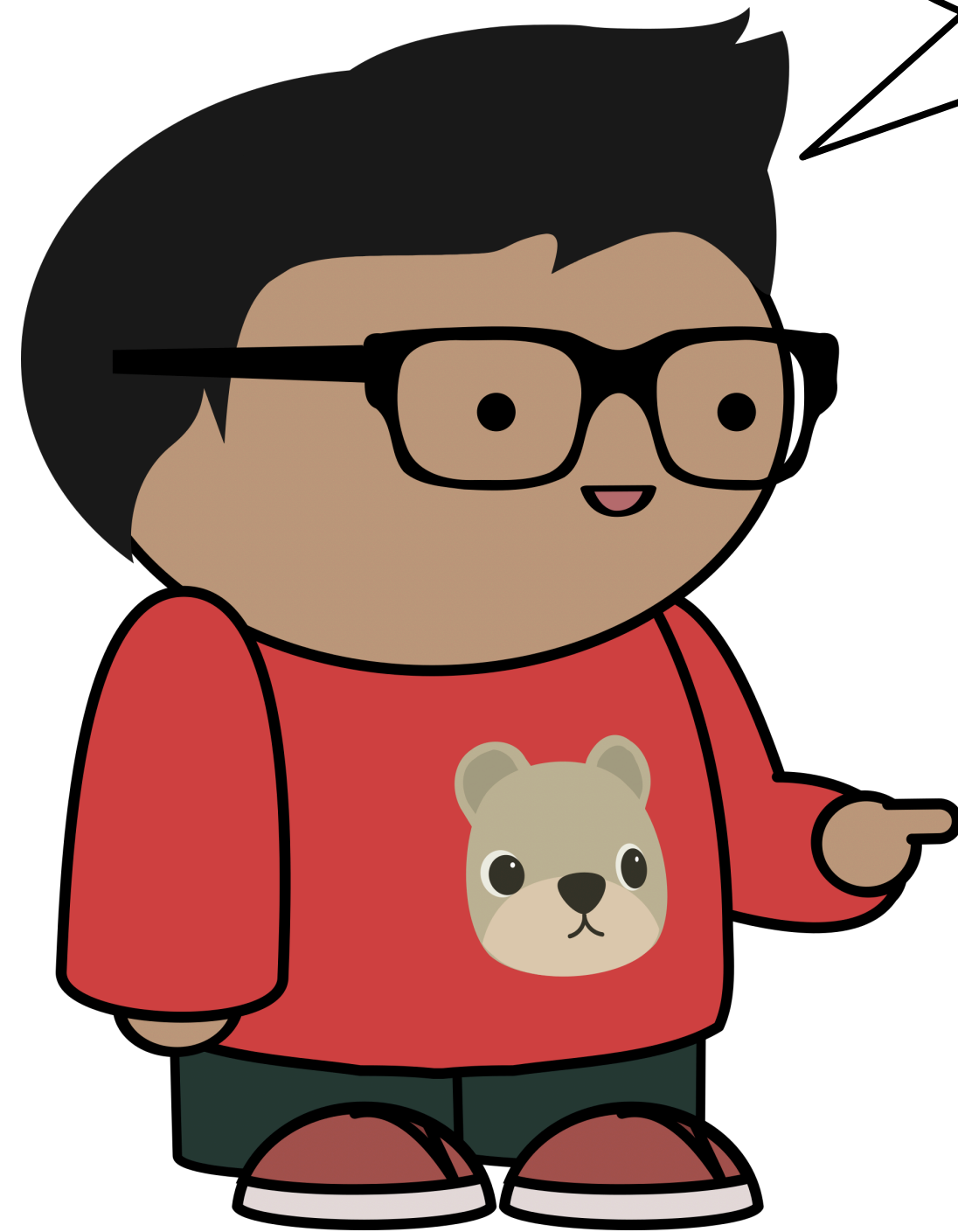
Syntax

Morphology

Phonology

Phonetics

How do we determine the order of the words?



*Must be
German...*

Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

How do we determine the order of the words?



Syntax: the structure of a sentence is determined by a set of language-specific syntactic rules.



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

Morphology: the rules that determine how words are formed.



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

Morphology: the rules that determine how words are formed.



Pragmatics

Semantics

Syntax

Morphology

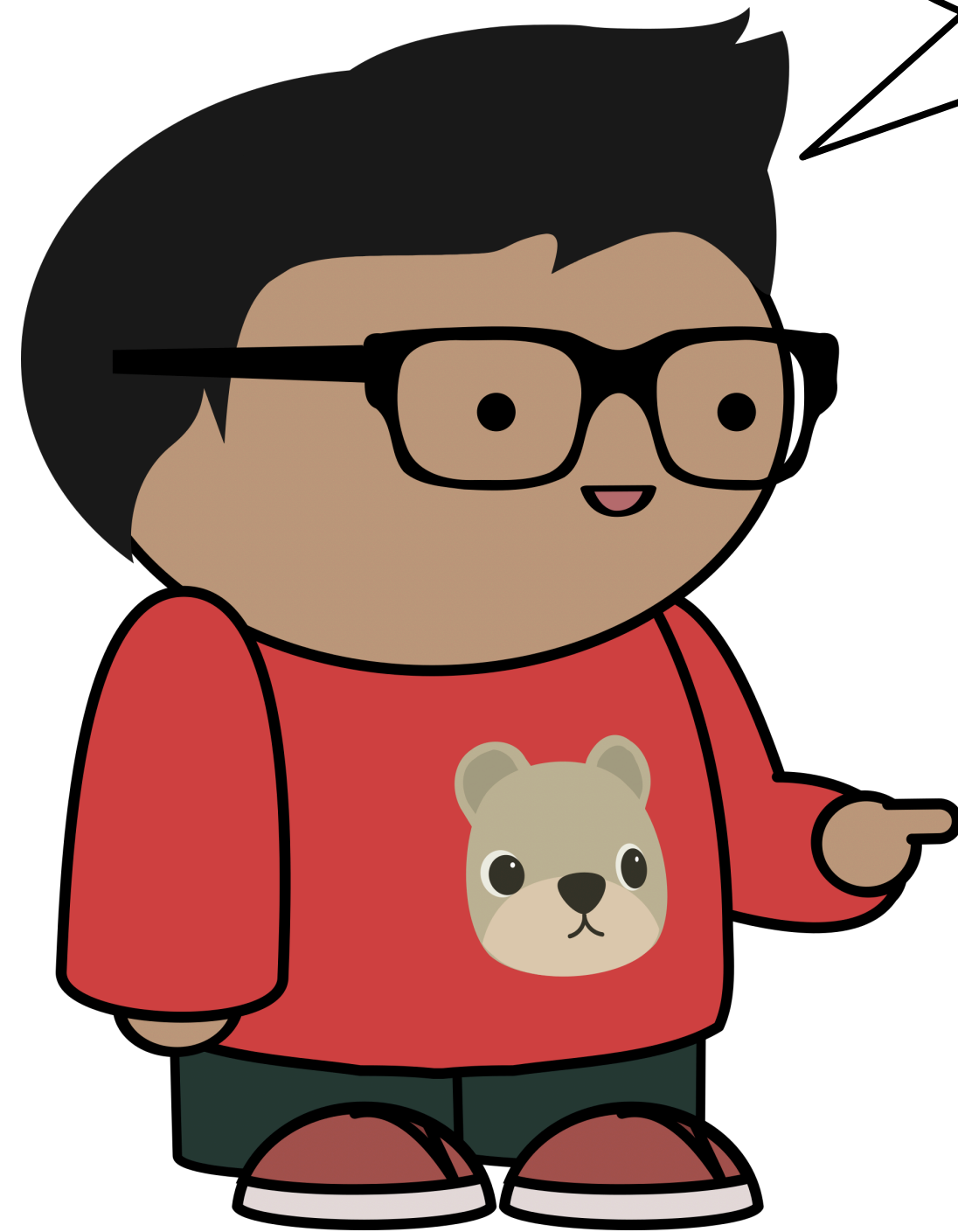
Phonology

Phonetics

Morphology: the rules that determine how words are formed.

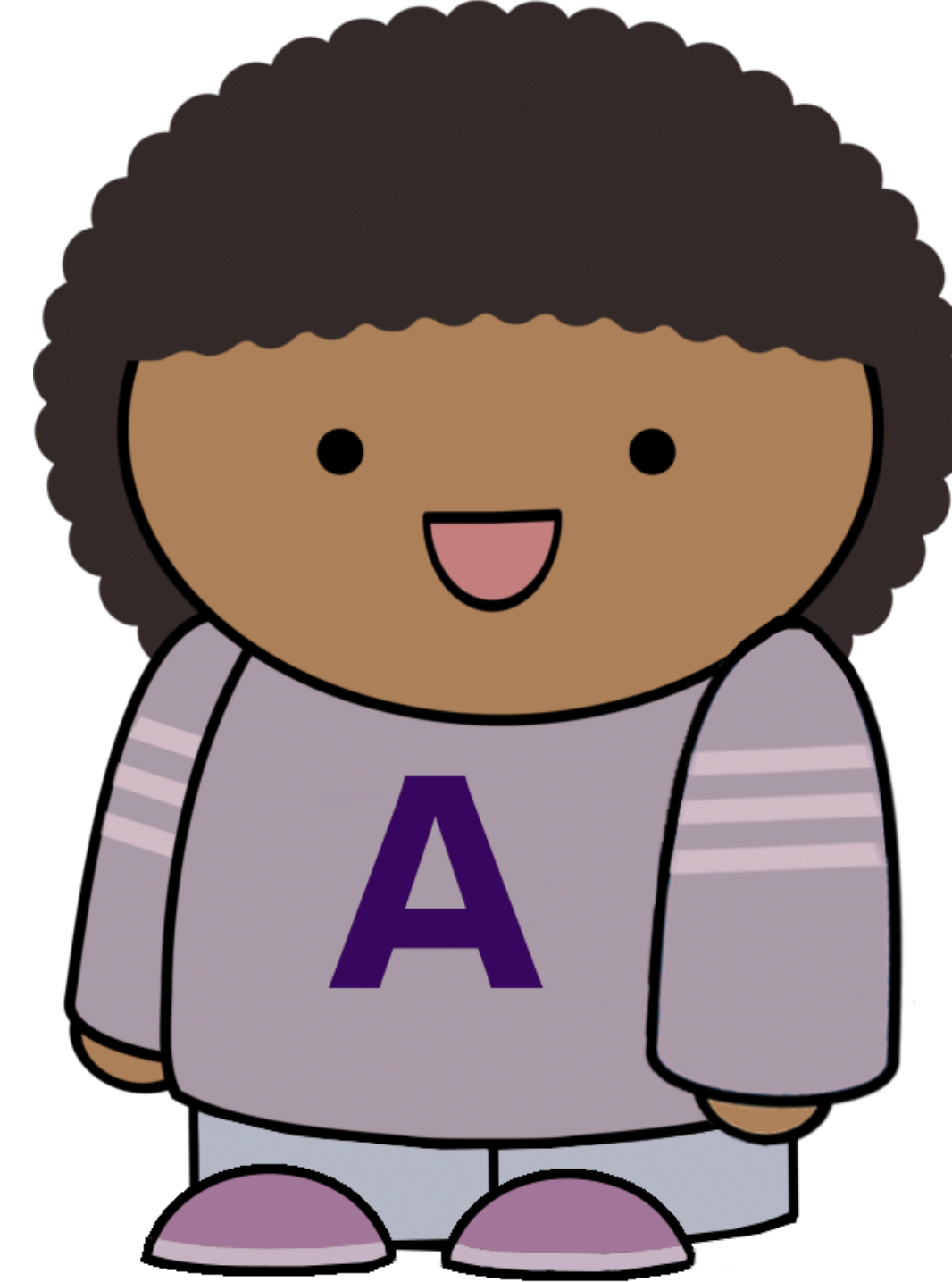


Morphology: the rules that determine how words are formed.



'wanə ɡoʊ ɡɛt
aɪs krim?

*Must be a weird
American...*



Pragmatics

Semantics

Syntax

Morphology

Phonology

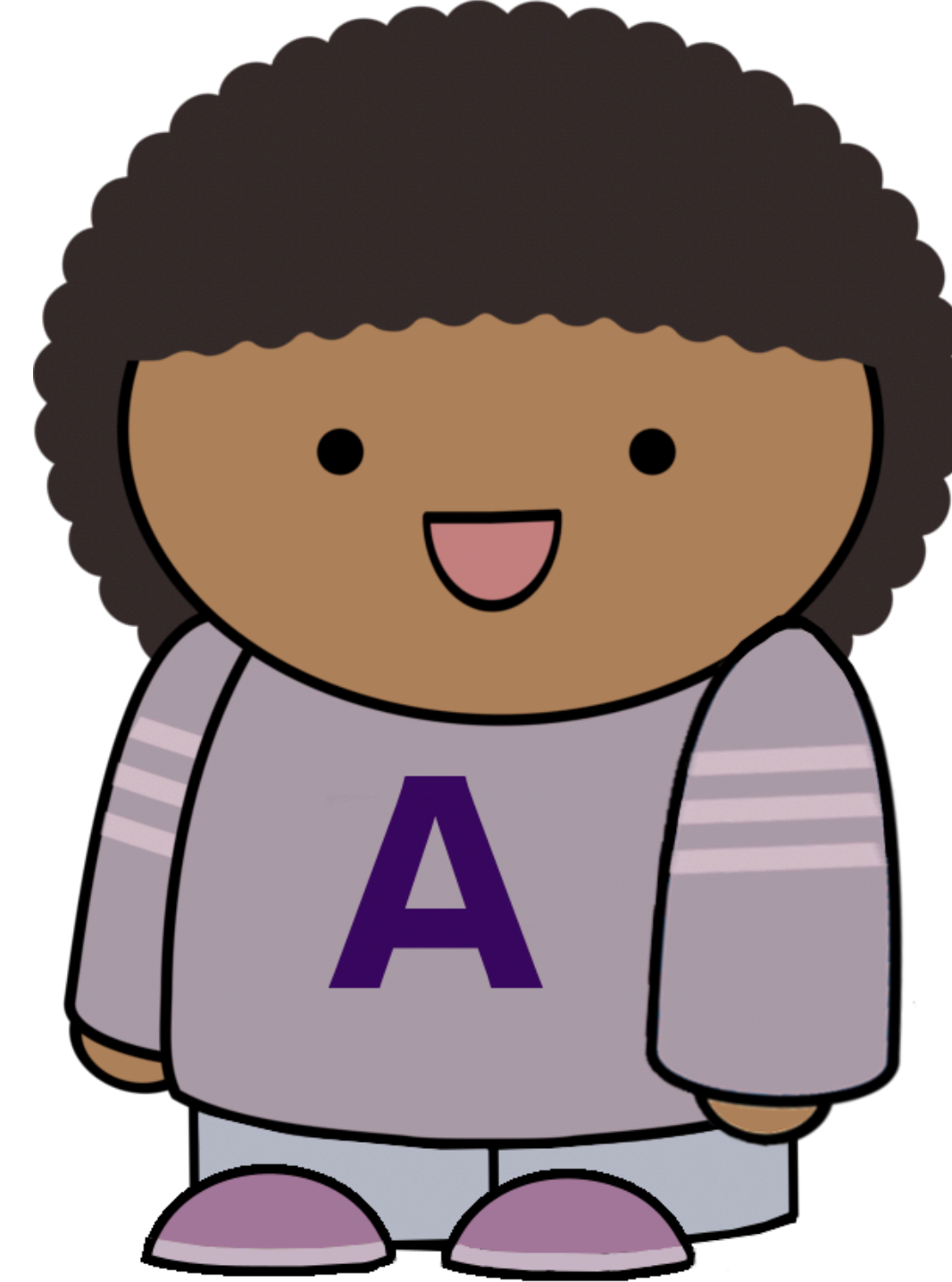
Phonetics

Phonology: the rules that determine how the sounds/signs of a language are organized



'wɒnə ɡəʊ ɡɛt
aɪs kri:m?

*Must be a weird
Brit...*



Pragmatics

Semantics

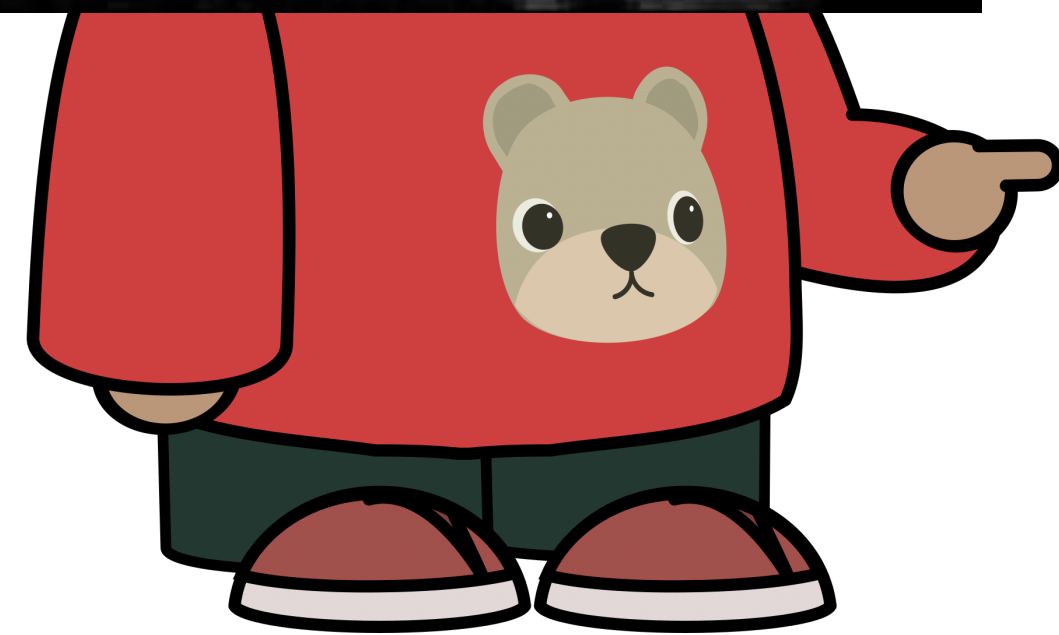
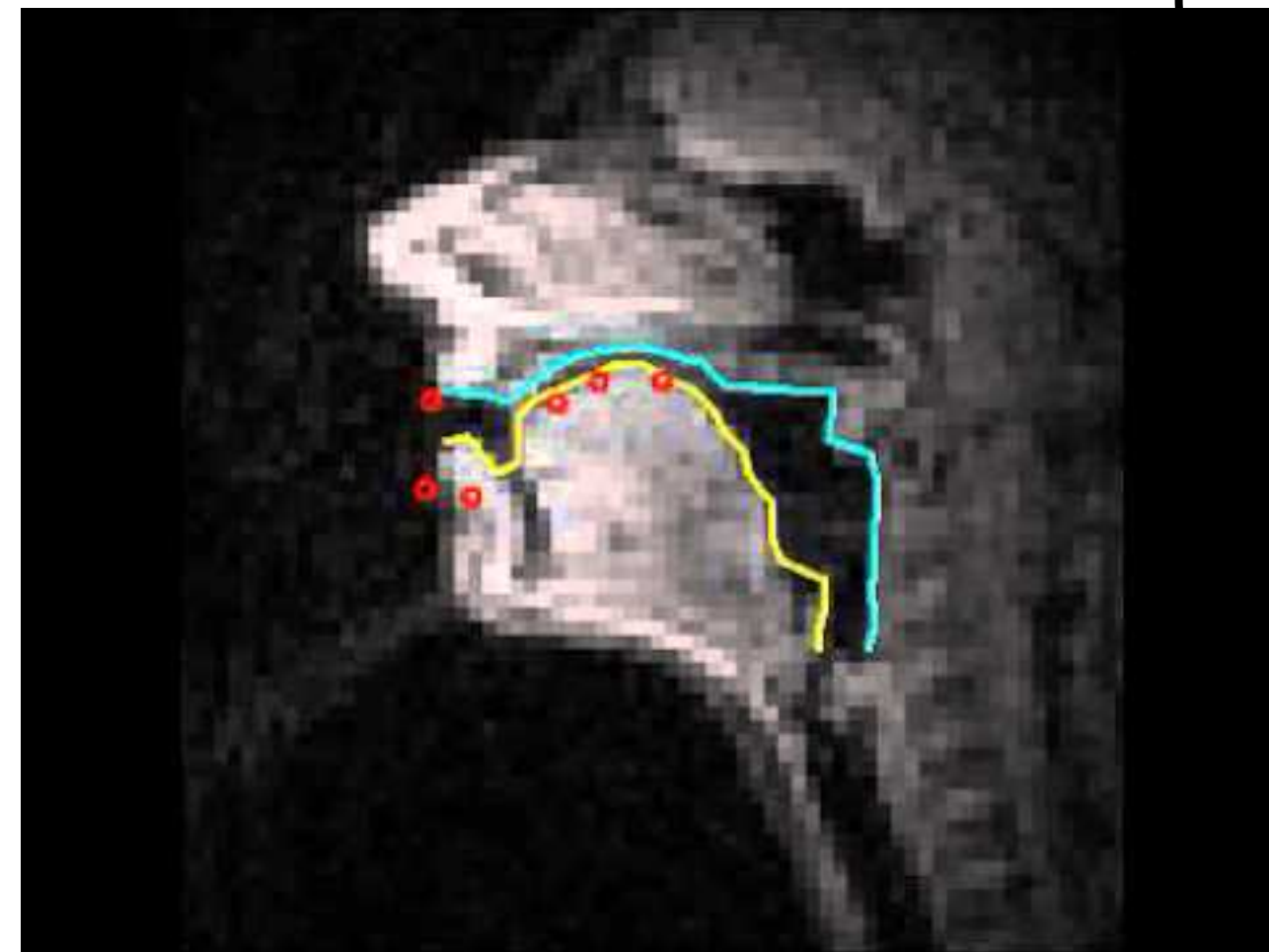
Syntax

Morphology

Phonology

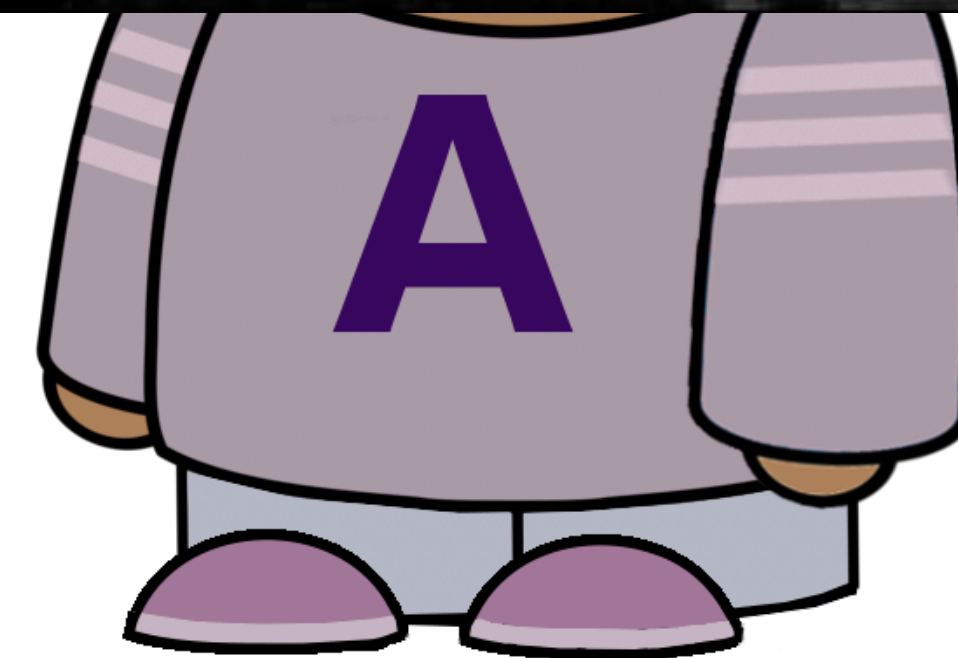
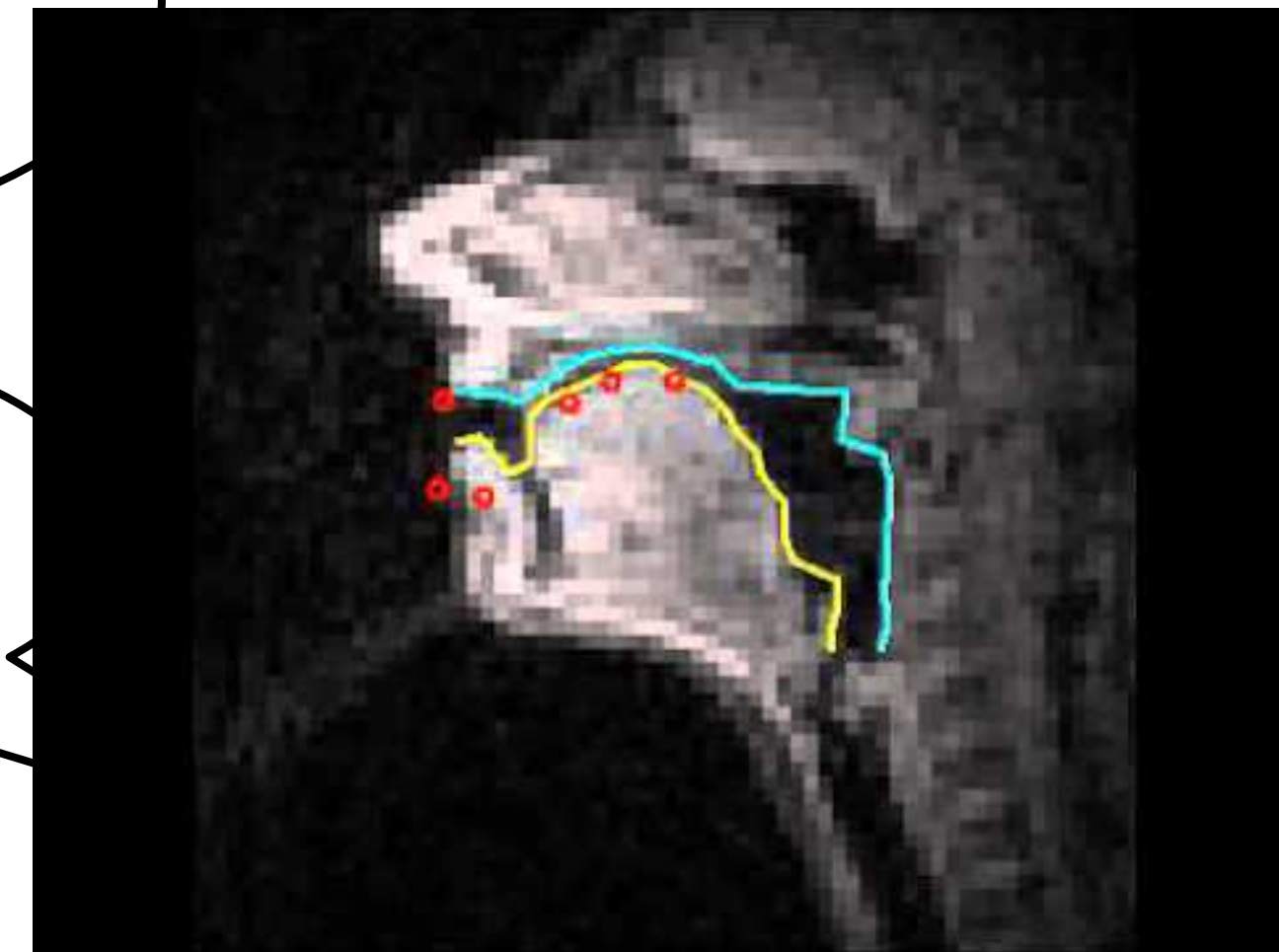
Phonetics

Phonology: the rules that determine how the sounds/signs of a language are organized



'wanə ɡoʊ ɡet
aɪs krim?

ɪts 'sevn ə ɛm.



Pragmatics

Semantics

Syntax

Morphology

Phonology

Phonetics

Phonetics: how do language users produce the building blocks of language?

Layers of linguistic abstraction

Pragmatics *It's 7 am* \Rightarrow *It's a weird time for ice cream so I don't know how to respond.*

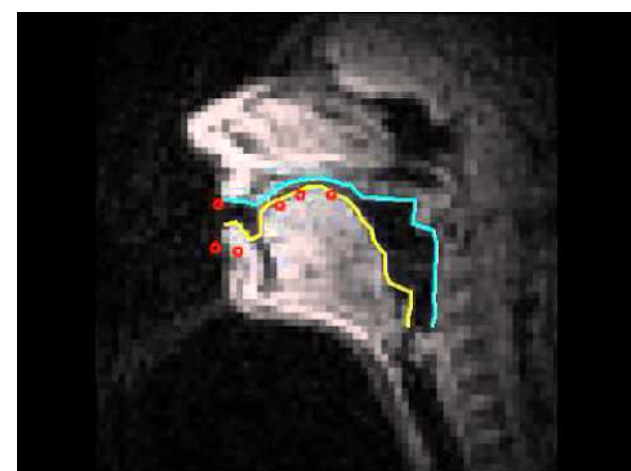
Semantics $[[\textit{It's 7 am}]] \rightarrow \text{True if now(w) == 7 am else False}$

Syntax $\{7, \textit{it's}, \textit{am}\} \rightarrow \textit{It's 7 am}.$

Morphology $\{7 \rightarrow \textit{7}, [\textit{it is}] \rightarrow \textit{it's}, \textit{am} \rightarrow \textit{am}\}$

Phonology *its* 'sɛvən ə ɛm.

Phonetics



Layers of linguistic abstraction

Pragmatics *It's 7 am* \Rightarrow *It's a weird time for ice cream so I don't know how to respond.*

Semantics $[[\textit{It's 7 am}]] \rightarrow \text{True if now(w) == 7 am else False}$

Syntax $\{7, \textit{it's}, \textit{am}\} \rightarrow \textit{It's 7 am}.$

Morphology $\{7 \rightarrow 7, [\textit{it is}] \rightarrow \textit{it's}, \textit{am} \rightarrow \textit{am}\}$

Orthography I t ' s 7 a m

Humans learn language instinctively:

Language has a *critical acquisition period*.

Language acquisition begins before birth and follows predictable developmental stages.

Humans can't decide *not* to learn language!

Language acquisition does not seem to correlate with intelligence.

All human cultures have language, and all human languages are equally expressive.

Child language acquisition



Example 1



Example 2

Child language acquisition



Example 1

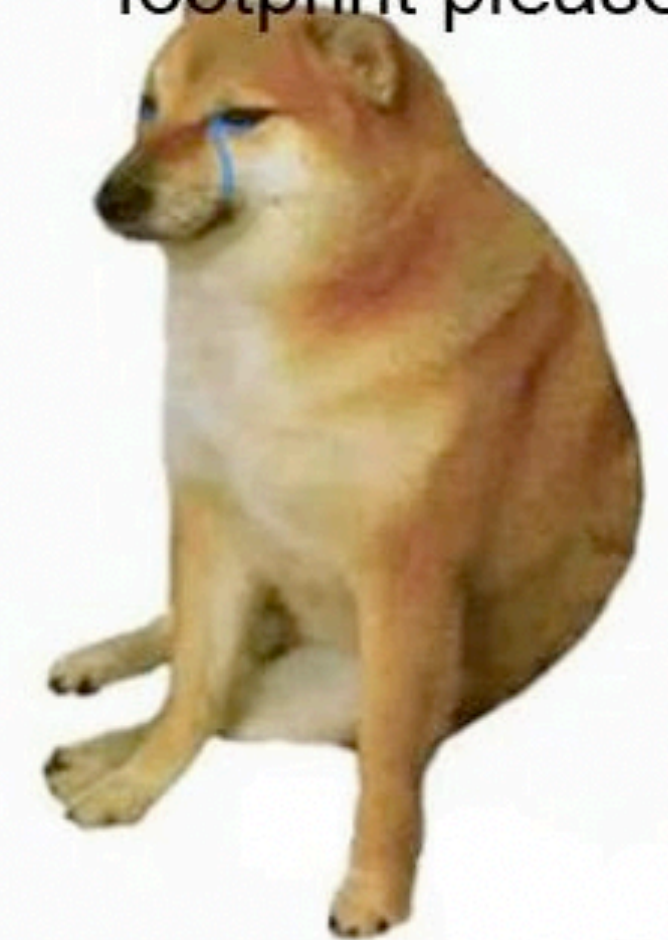


Example 2



human infants

with fast mapping, i
can learn the meaning
of a word in 3
exposures in my
human brain powered
by food



large language models

1 trillion
parameters and
a carbon
footprint please

What does it mean for a computer to understand language?

VOL. LIX. No. 236.]

[October, 1950

MIND

A QUARTERLY REVIEW

OF

PSYCHOLOGY AND PHILOSOPHY



I.—COMPUTING MACHINERY AND
INTELLIGENCE

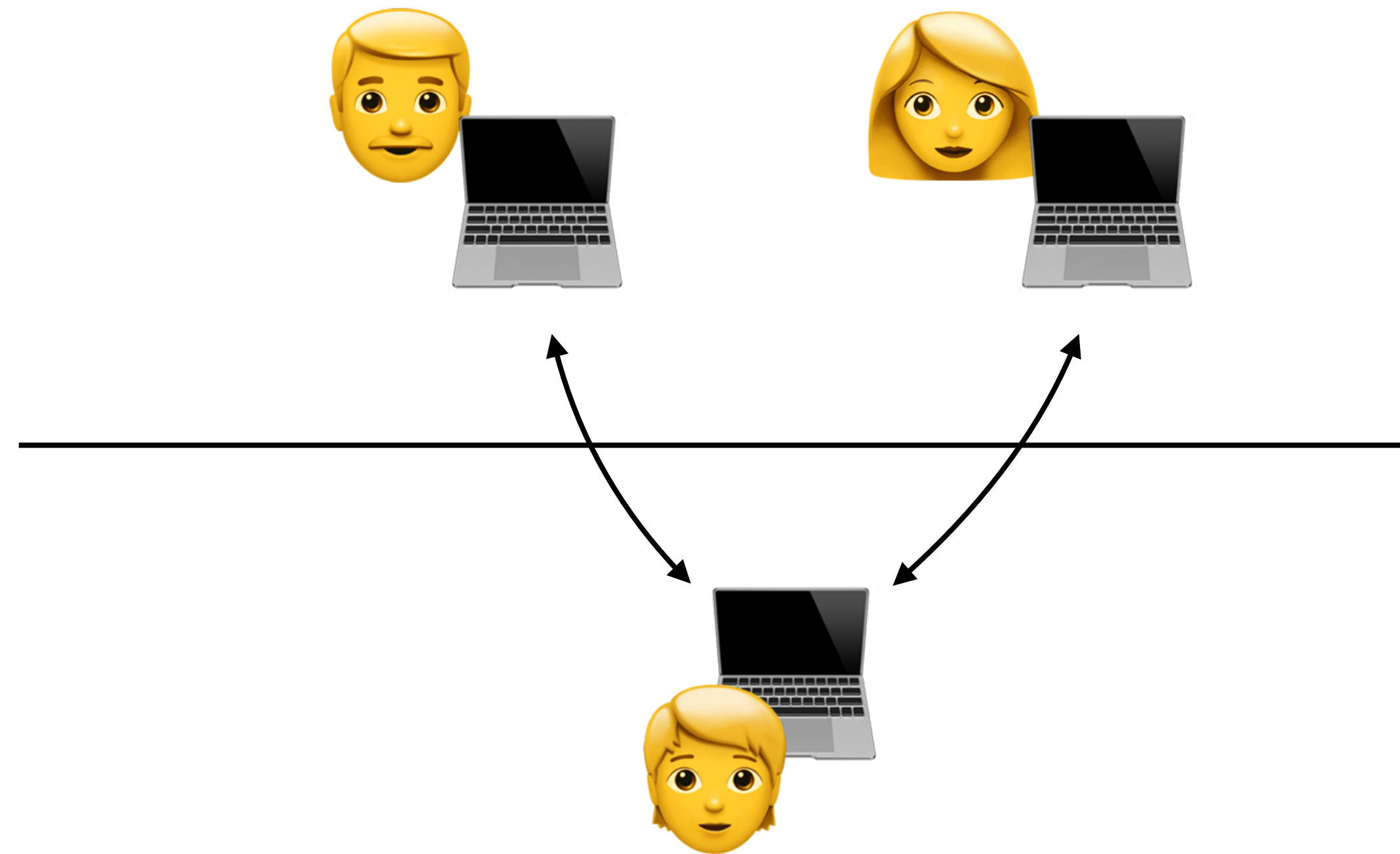
By A. M. TURING

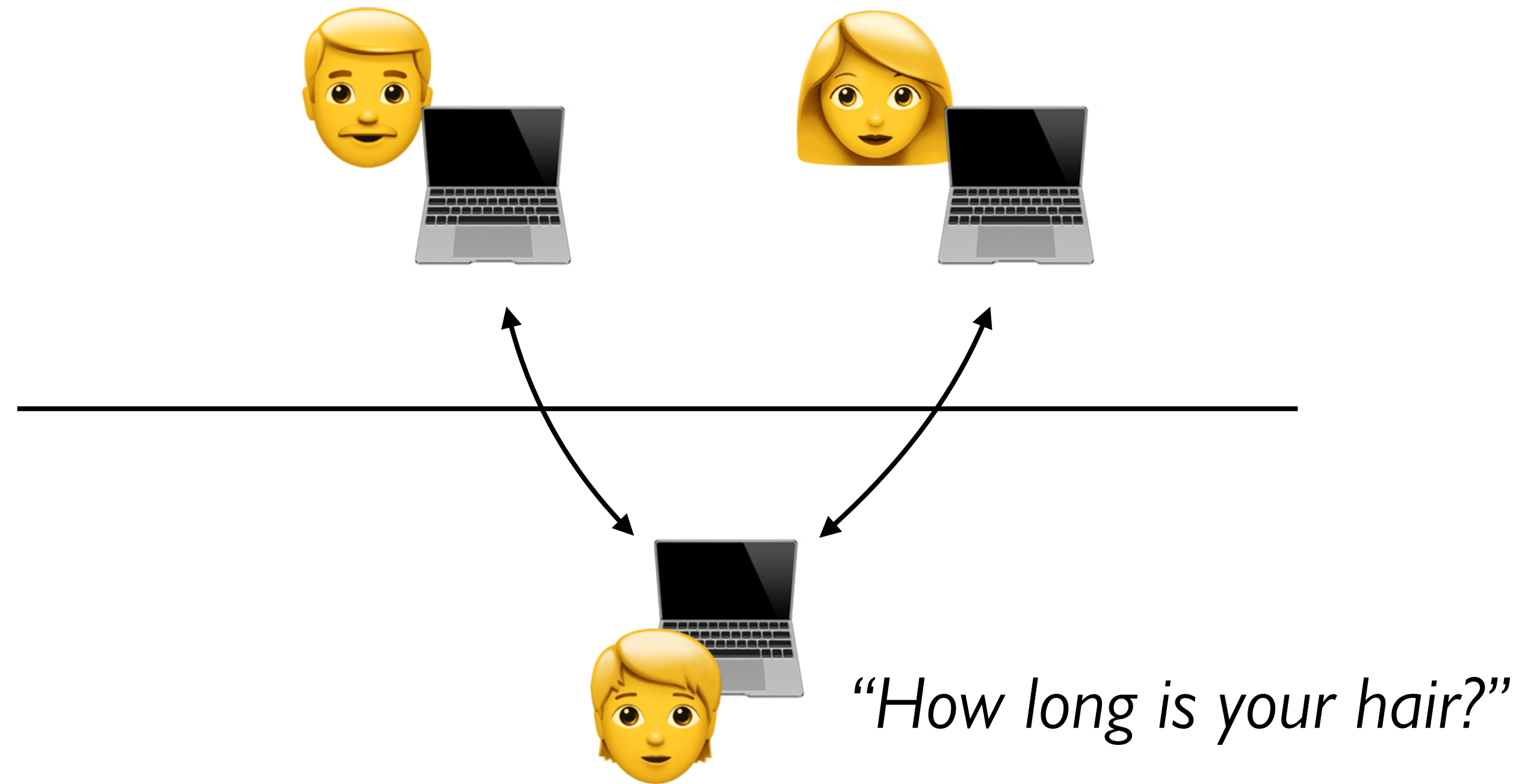
Can machines think?

~~Can machines think?~~

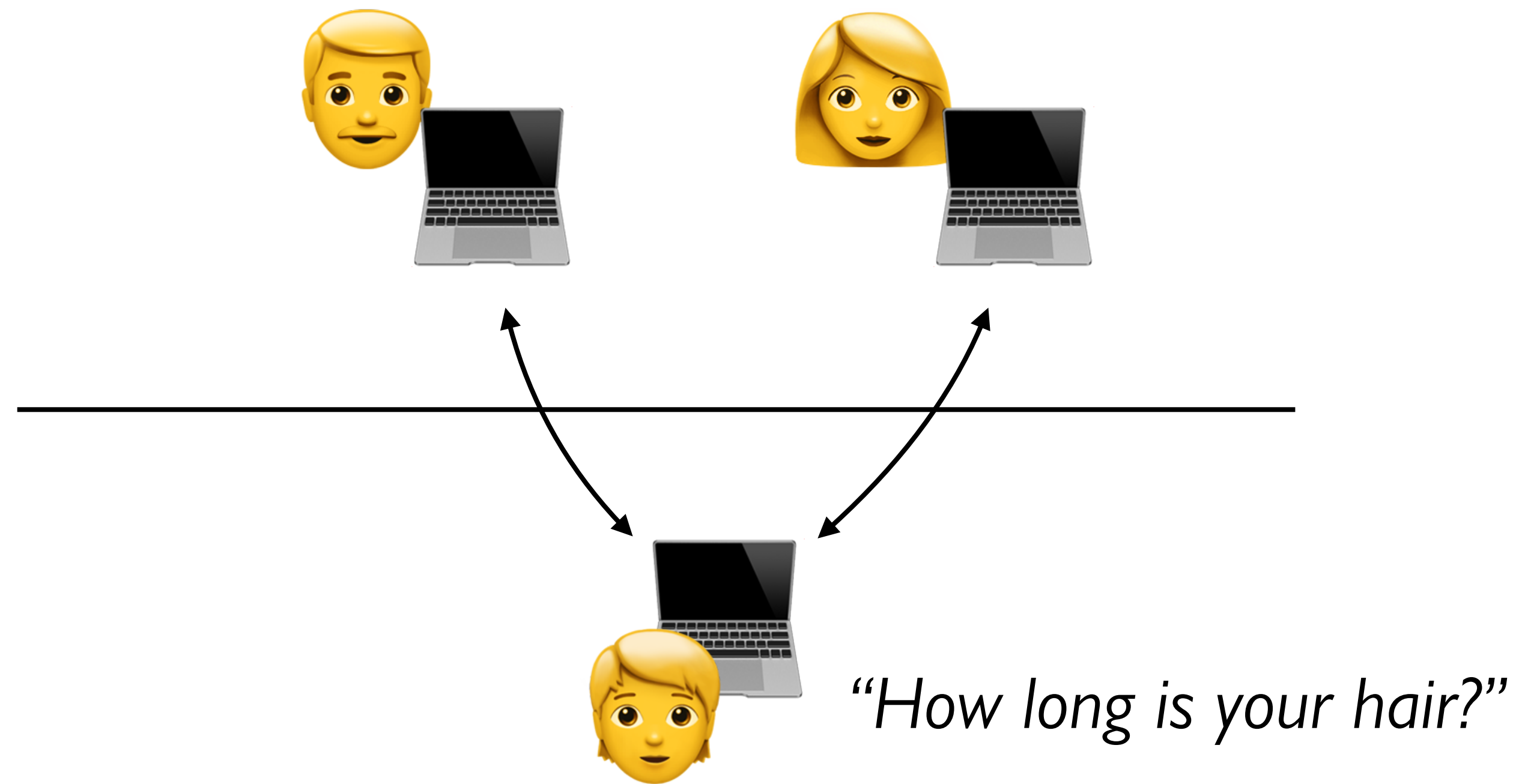
“I shall replace the question by another, which is closely related to it and is expressed in relatively unambiguous words.

“The new form of the problem can be described in terms of a game which we call the ‘imitation game’...”

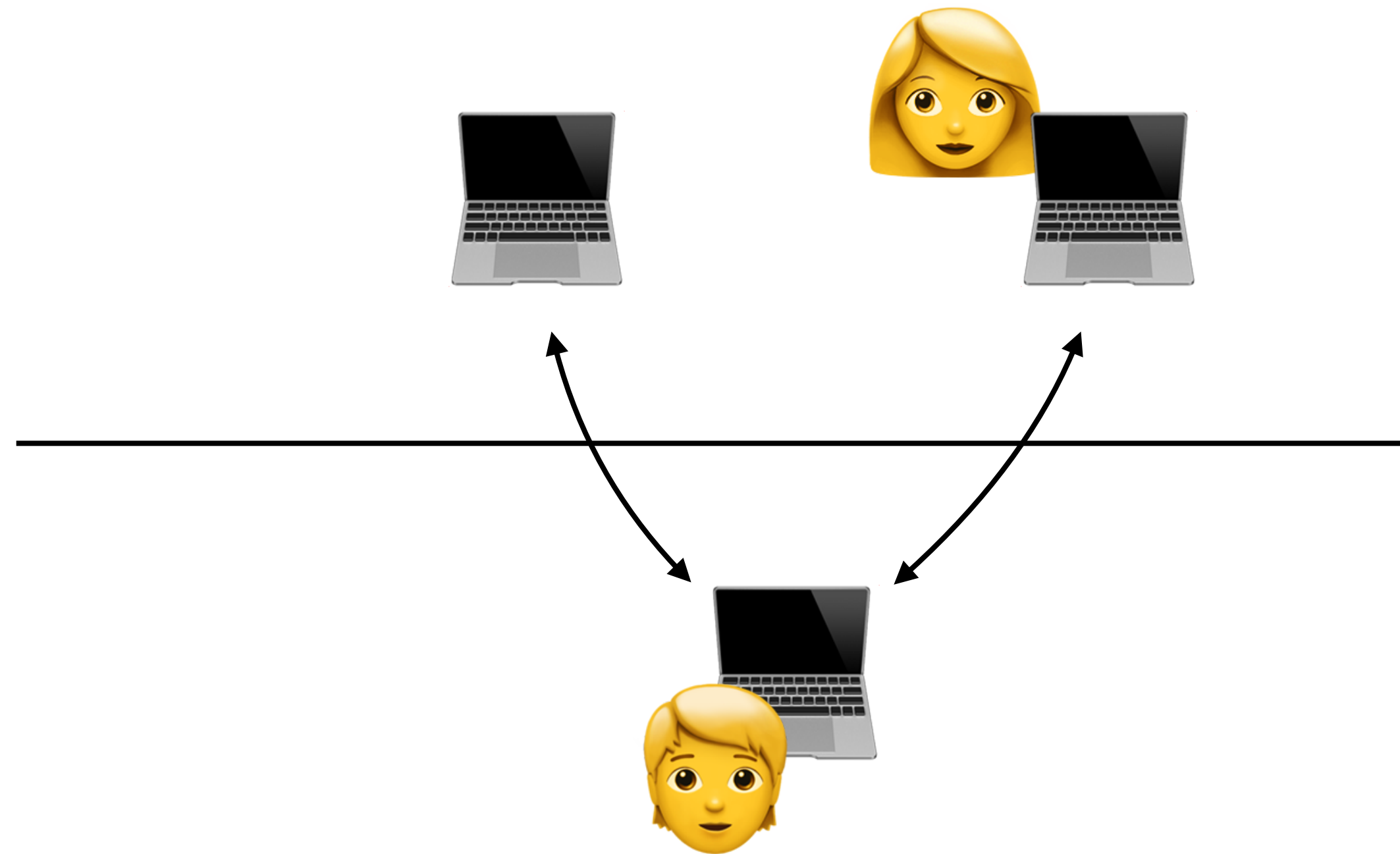




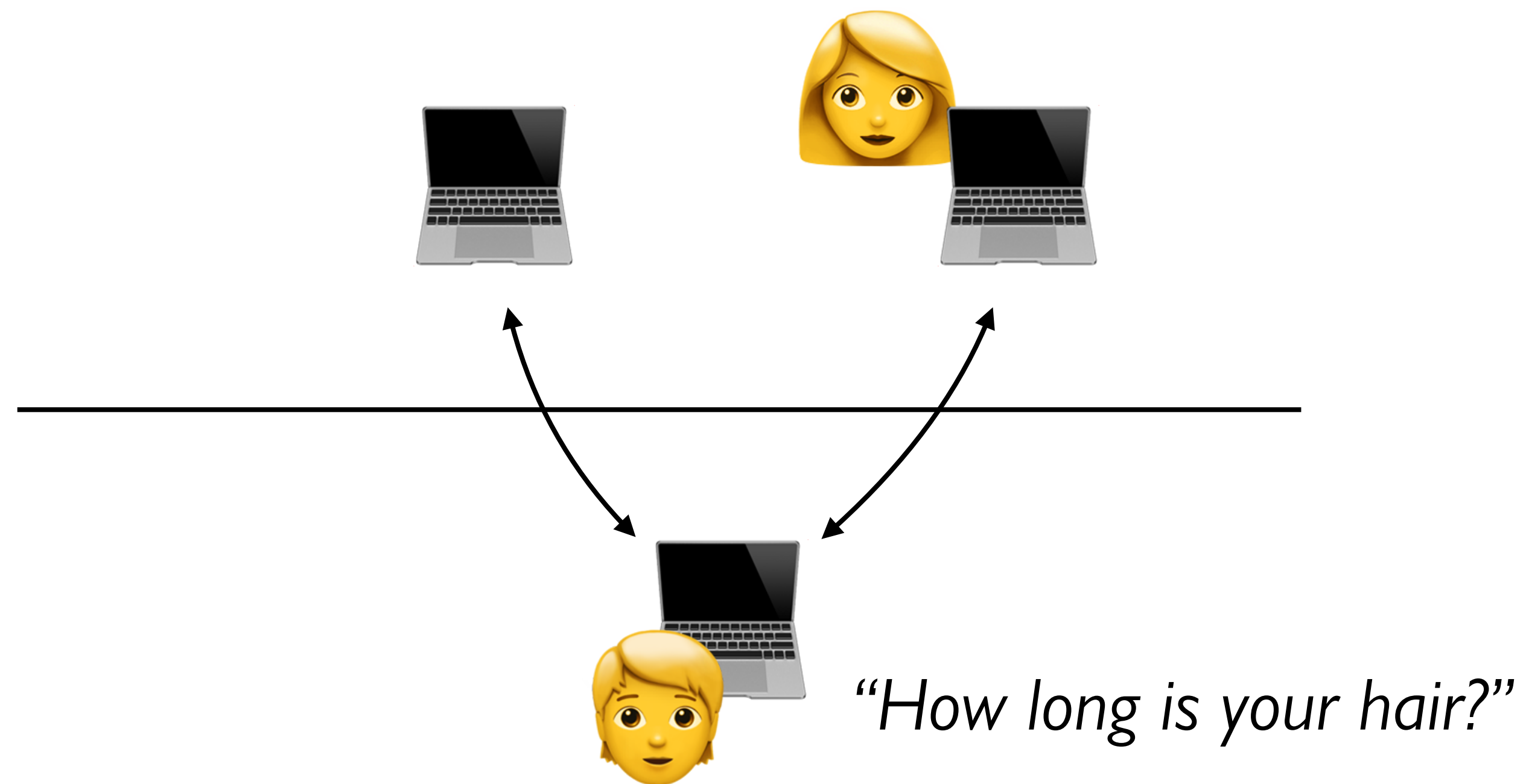
“This is so 1950 it hurts.”



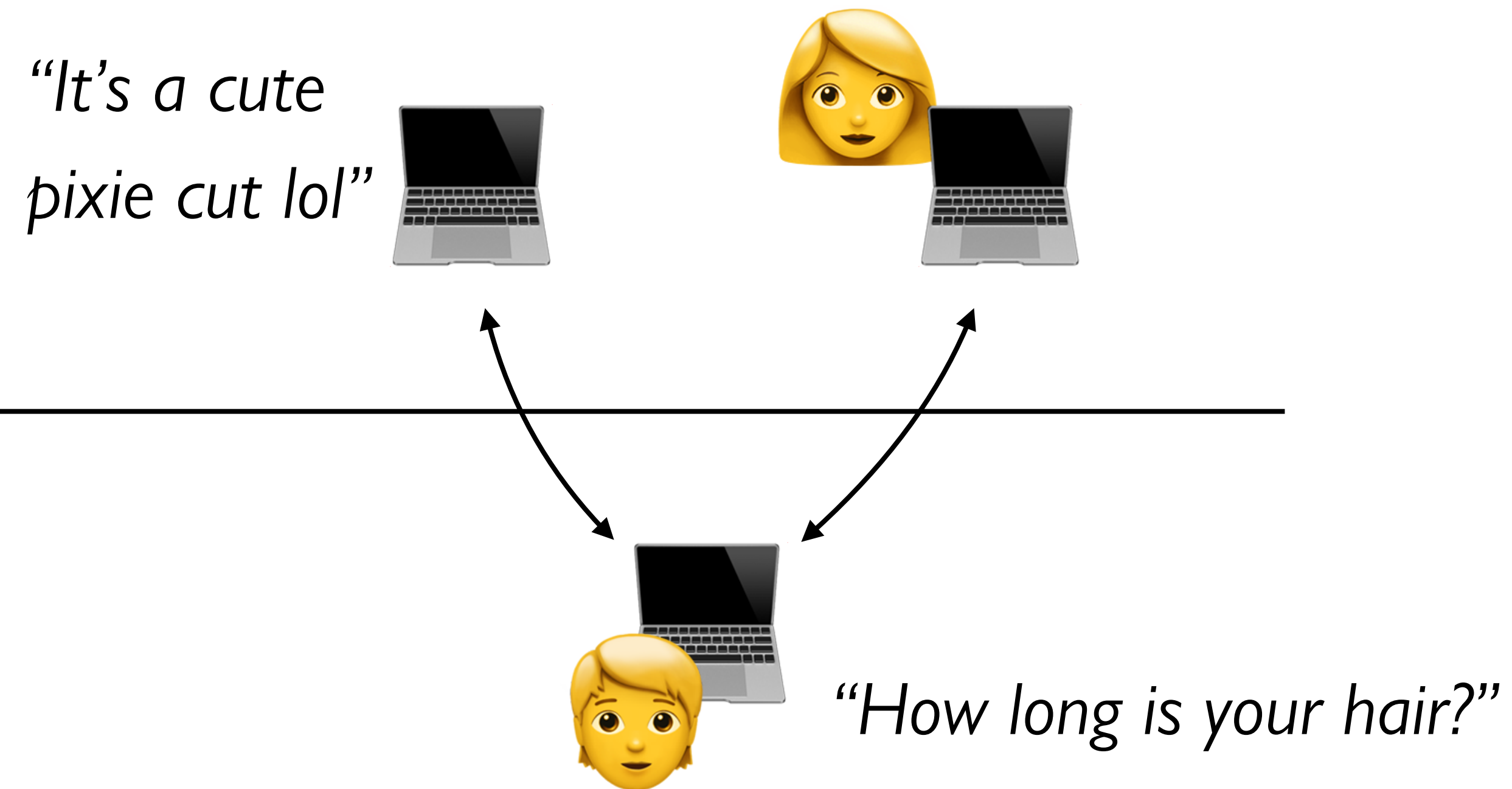
Big brain time



Big brain time



Big brain time



Turing suggested talking over teletypes, i.e., texting with the computer.

We often prefer speech, but high-quality speech recognition and synthesis turns out to be one of the *easier* parts of making an intelligent machine!

**Q : Please write me a sonnet on the subject of the Forth
Bridge.**



Q : Please write me a sonnet on the subject of the Forth Bridge.



Q : Please write me a sonnet on the subject of the Forth Bridge.

A : Count me out on this one. I never could write poetry.

Q : Add 34957 to 70764

A : (Pause about 30 seconds and then give as answer) 105621.

Q : Do you play chess ?

A : Yes.

Q : I have K at my K1, and no other pieces. You have only K at K6 and R at R1. It is your move. What do you play ?

A : (After a pause of 15 seconds) R-R8 mate.

There's a long history of making computers we can talk to, some serious, some not so serious.

USER: *You are like my father in some ways.*

ELIZA: *What resemblance do you see*

USER: *You are not very aggressive but I think you don't want me to notice that.*

ELIZA: *What makes you think I am not aggressive*

USER: *You don't argue with me*

ELIZA: *Why do you think I don't argue with you*

USER: *You are afraid of me*

ELIZA: *Does it please you to believe I am afraid of you*



c. 1964

ChatGPT 5 >

For this dialogue, respond as a Rogerian psychotherapist would.

I hear that you'd like me to respond in the manner of a Rogerian psychotherapist. That tells me it's important for you to feel understood and met where you are, rather than being given direct advice.

Maybe we could begin with you sharing a little about what's been on your mind lately. I'll listen closely, reflect back what I hear, and together we can explore your thoughts and feelings in a way that honors your own perspective and wisdom.

Would you like to start by telling me what's been weighing on you or what brought you here today?

You are like my father in some ways.

It sounds like something in the way I respond reminds you of your father. I'm wondering what qualities or feelings come up for you in that comparison.

Is it a sense of familiarity, comfort, or maybe something more complicated?

↓

Ask anything

5

Course information

docs.google.com/forms/d/e/1FAIpQLSdKCQcfQeFvqp0_piNQGGq4myMHWqS13sn

CMPU 366 Student information
Fall 2025

Please fill out this short form to help me better prepare for the start of the semester.

jgordon@vassar.edu [Switch account](#)

* Indicates required question

Email *

☐ Record jgordon@vassar.edu as the email to be included with my response

What name would you like me to call you? *

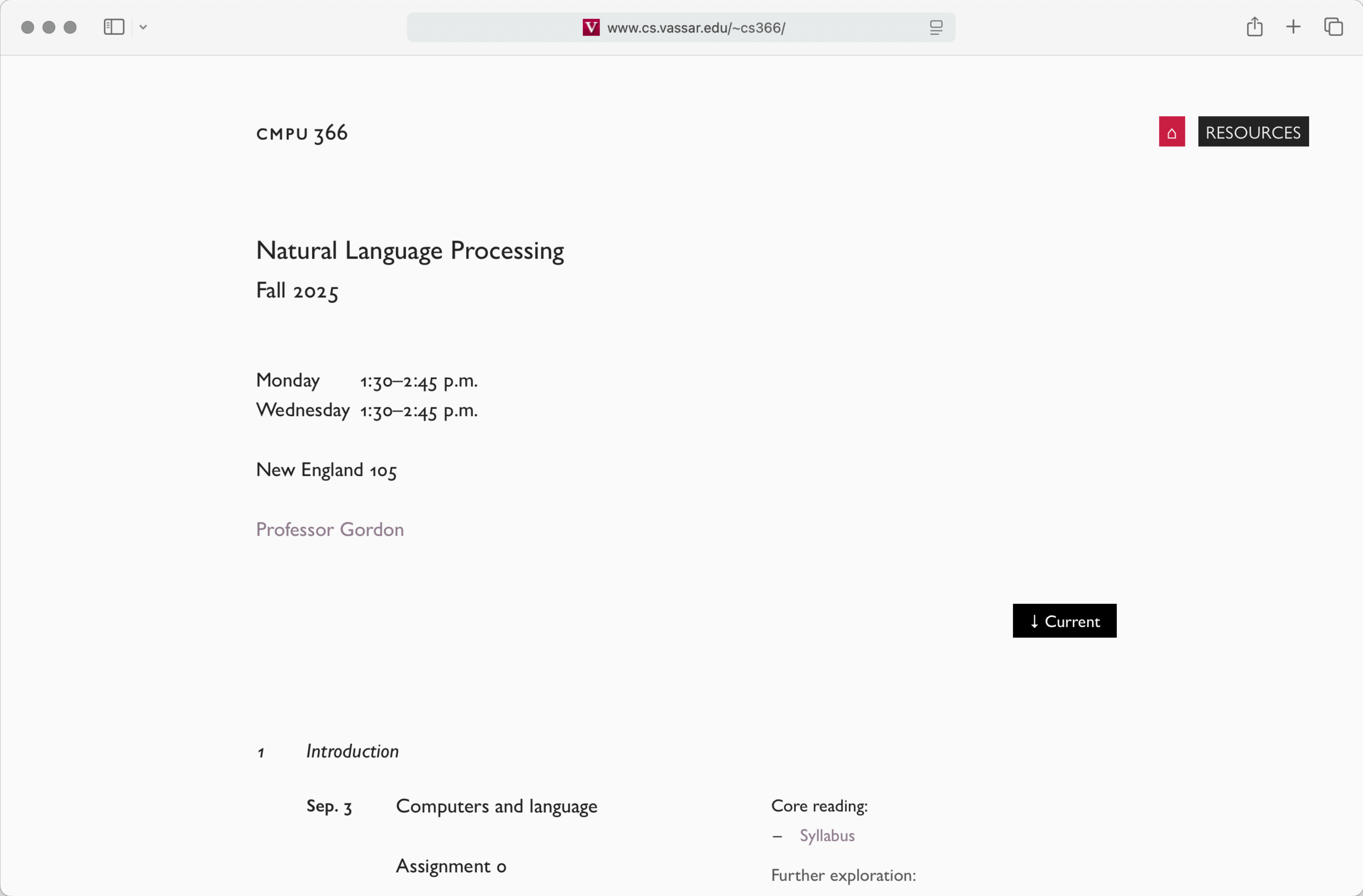
Your answer

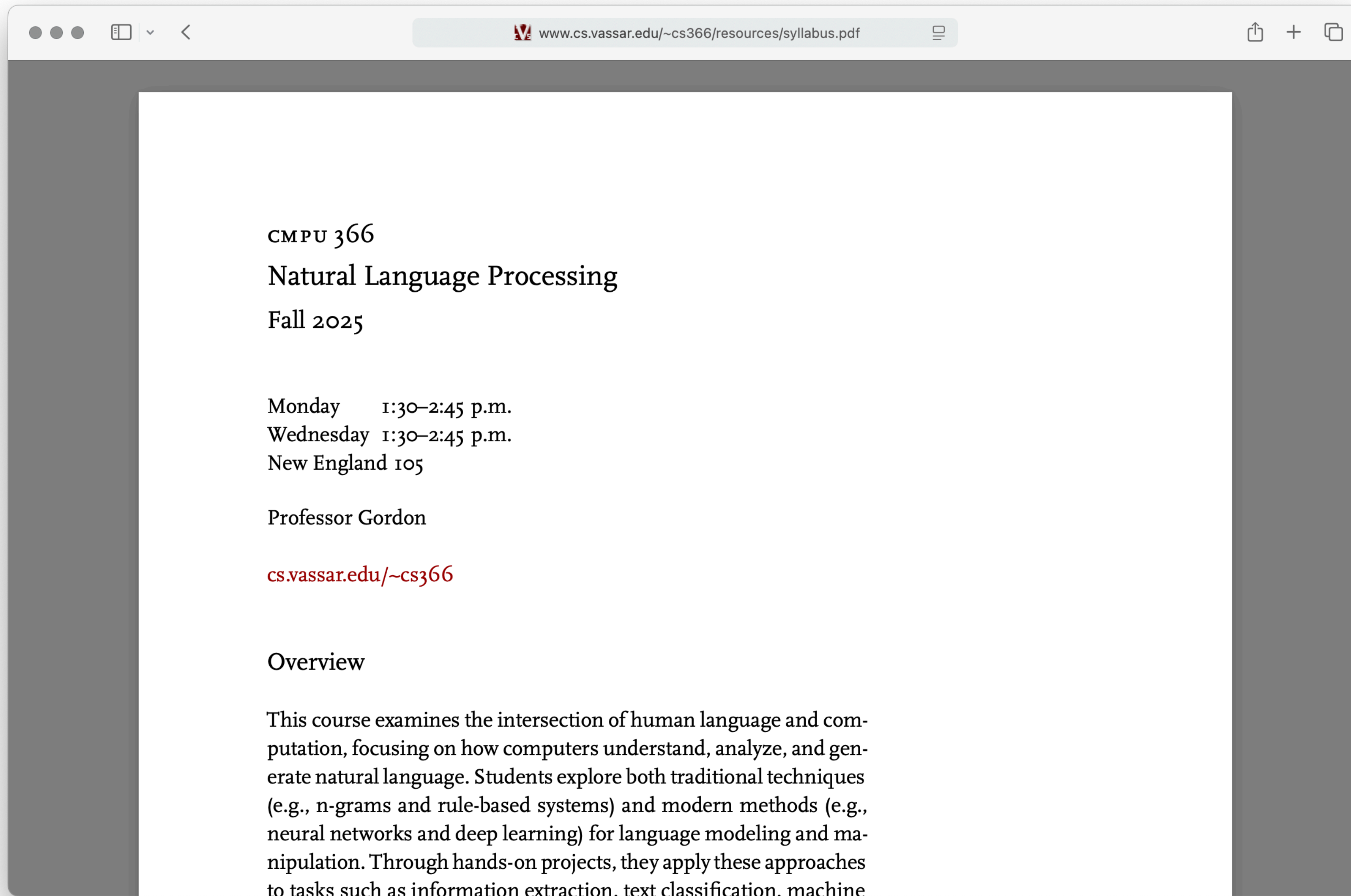
Are you trying to add this class, currently enrolled, or planning to drop? *

☐ I would like to add this class.

☐ I am currently enrolled in this class.

forms.gle/MNQbuyuTstVsLnEh7





CMPU 366

Natural Language Processing

Fall 2025

Monday 1:30–2:45 p.m.

Wednesday 1:30–2:45 p.m.

New England 105

Professor Gordon

cs.vassar.edu/~cs366

Overview

This course examines the intersection of human language and computation, focusing on how computers understand, analyze, and generate natural language. Students explore both traditional techniques (e.g., n-grams and rule-based systems) and modern methods (e.g., neural networks and deep learning) for language modeling and manipulation. Through hands-on projects, they apply these approaches to tasks such as information extraction, text classification, machine

Prerequisites

Two 200-level CS courses

This is a requirement to take any 300-level CS course.

CMPU 250 Data and Its Discontents or
MATH 244 Intermediate Data Science

*But the department chair chose
to relax this requirement for this
semester.*

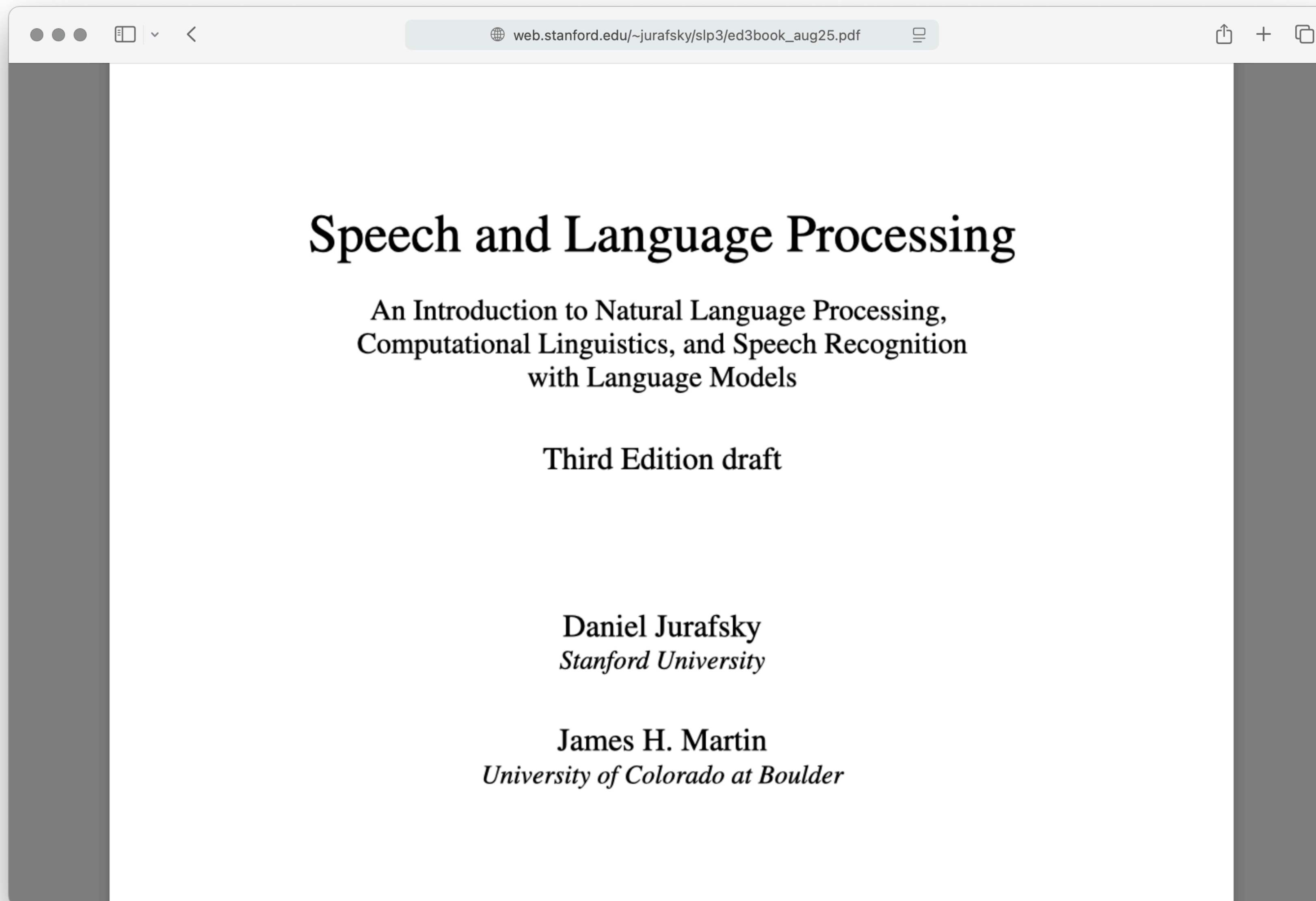
SPEECH AND LANGUAGE PROCESSING

*An Introduction to Natural Language Processing,
Computational Linguistics, and Speech Recognition*



Second Edition

DANIEL JURAFSKY & JAMES H. MARTIN



Speech and Language Processing

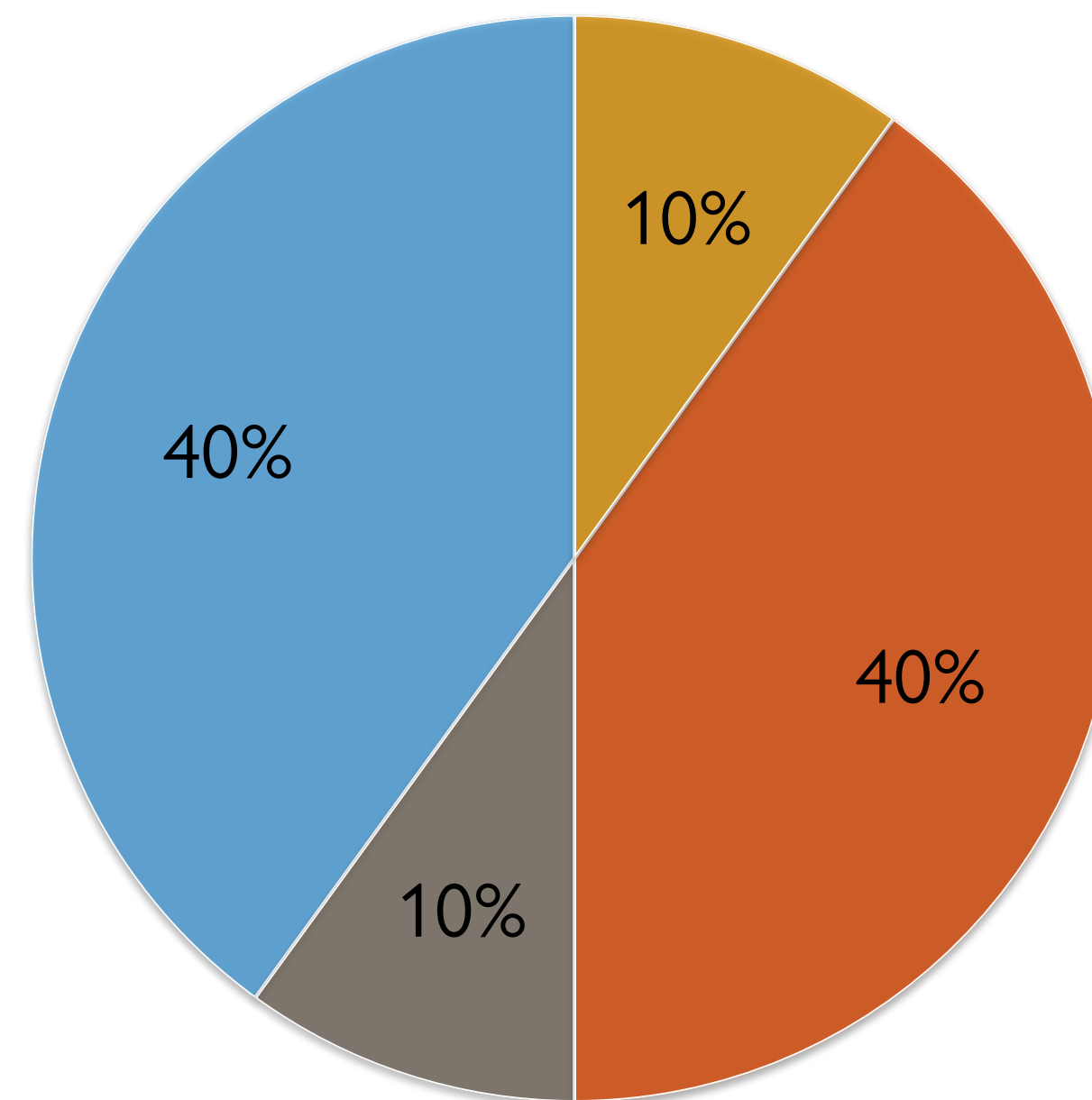
An Introduction to Natural Language Processing,
Computational Linguistics, and Speech Recognition
with Language Models

Third Edition draft

Daniel Jurafsky
Stanford University

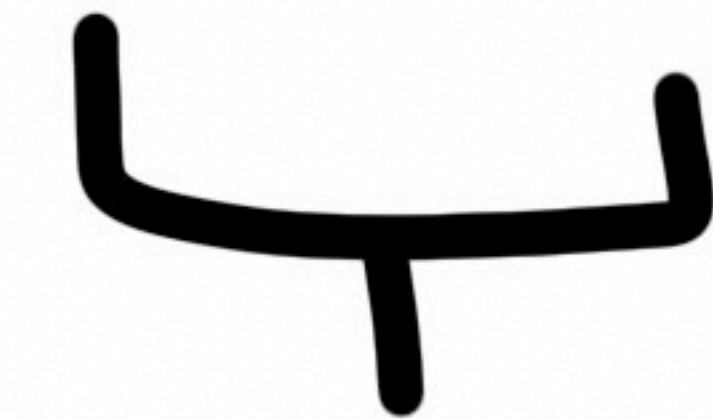
James H. Martin
University of Colorado at Boulder

Coursework



- Midterm exam
- Assignments
- Special topics presentations
- Final project

Assignments and projects will be in Python



This will be a
fun programing
language to
learn



wait this is a snake

gradescope[®]
by Turnitin

< ≡

CMPU 366

Natural Language Processing

Dashboard

Regrade Requests

Instructor

J. Gordon

Course Actions

Unenroll From Course

Account

^

CMPU 366

Fall 2025

Course ID: 1114236

Name	Status	Released	Due (EDT)
Your instructor hasn't released any assignments yet.			

gradescope.com

You have 4 late days for the semester, which you can use all at once, or spread across assignments.

Ordinarily, I will not accept late work beyond these days.

If you have a prolonged illness or unexpected circumstance, let me know and we'll work together to make a custom plan.

Collaboration policy

In this class, you can talk at a high-level with other students about assignments, but *you cannot show them your code.*

You may not use ChatGPT, Bard, Codex, or any other AI system unless explicitly stated in the homework assignment.

Office hours

I plan to have office hours Tuesday/Thursday 3–5 p.m., but check the [department website](#) for the most recent information.

No office hours work for everyone, so let me know if you'd like to meet at another time!

