Semantics and pragmatics ...in the age of deep learning

Matthijs Westera

Outline

Part 1

The ‘standard picture’ of semantics + pragmatics

Part 2

A radically different approach:
distributional semantics and deep learning

Part 3

What the standard picture gets wrong &
how the two approaches may be integrated

Part 1. The ‘standard picture’ of semantics and pragmatics

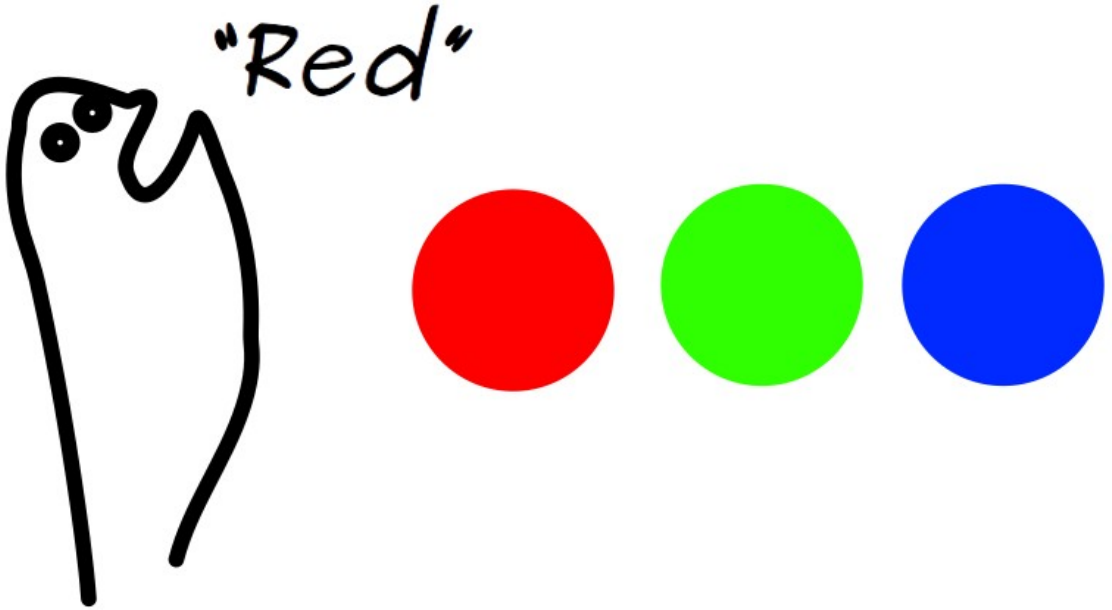
Q: What is semantics? What is pragmatics?

Expression meaning vs. speaker meaning (Grice ‘68)

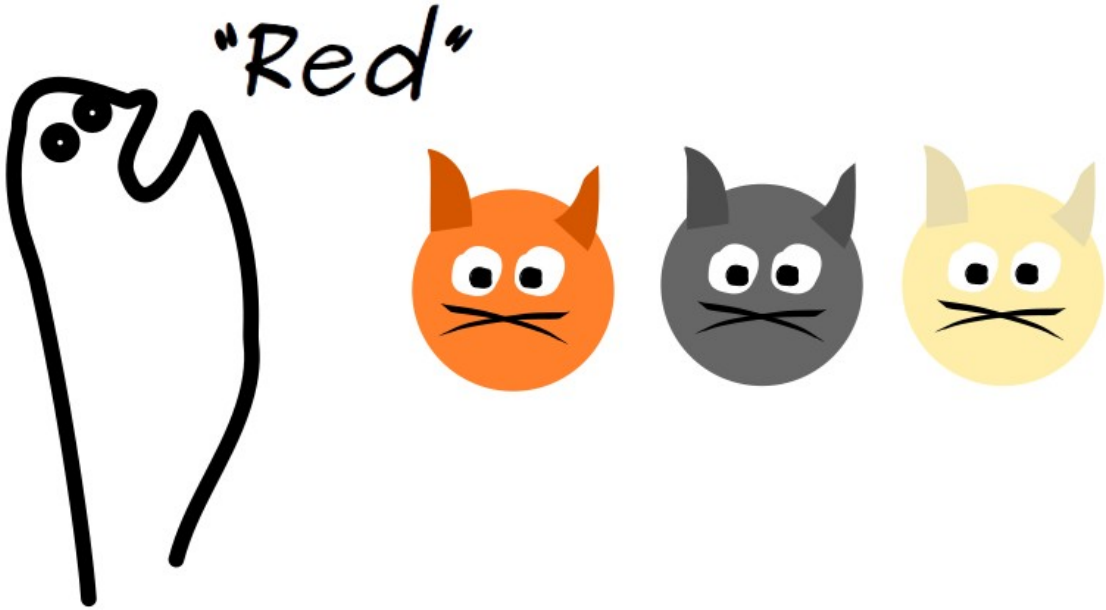
Expression meaning vs. speaker meaning (Grice ‘68)

"Red"

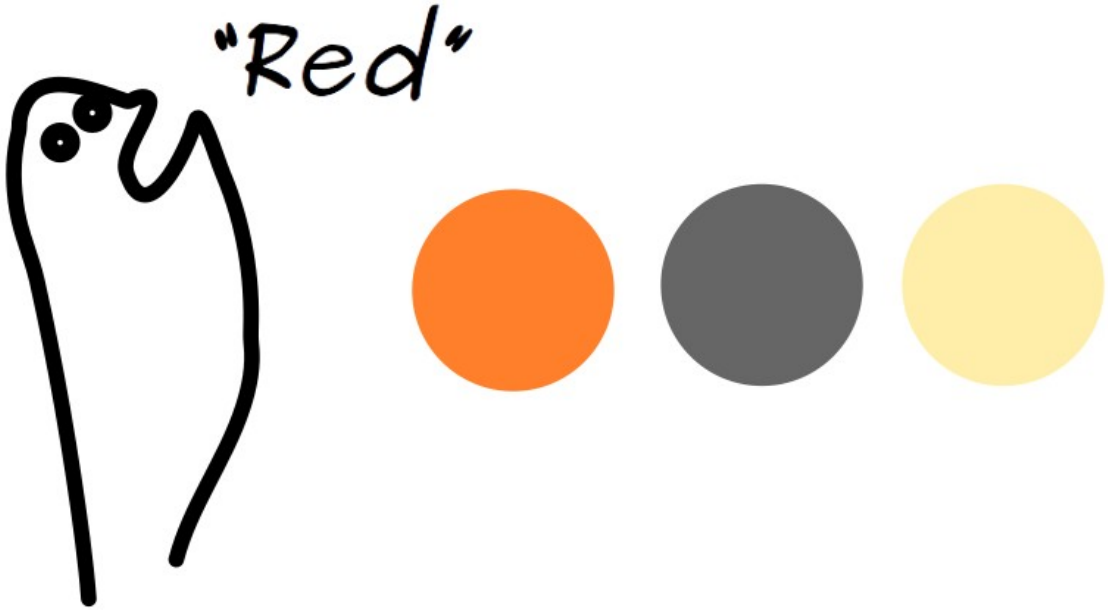
Expression meaning vs. speaker meaning (Grice ‘68)



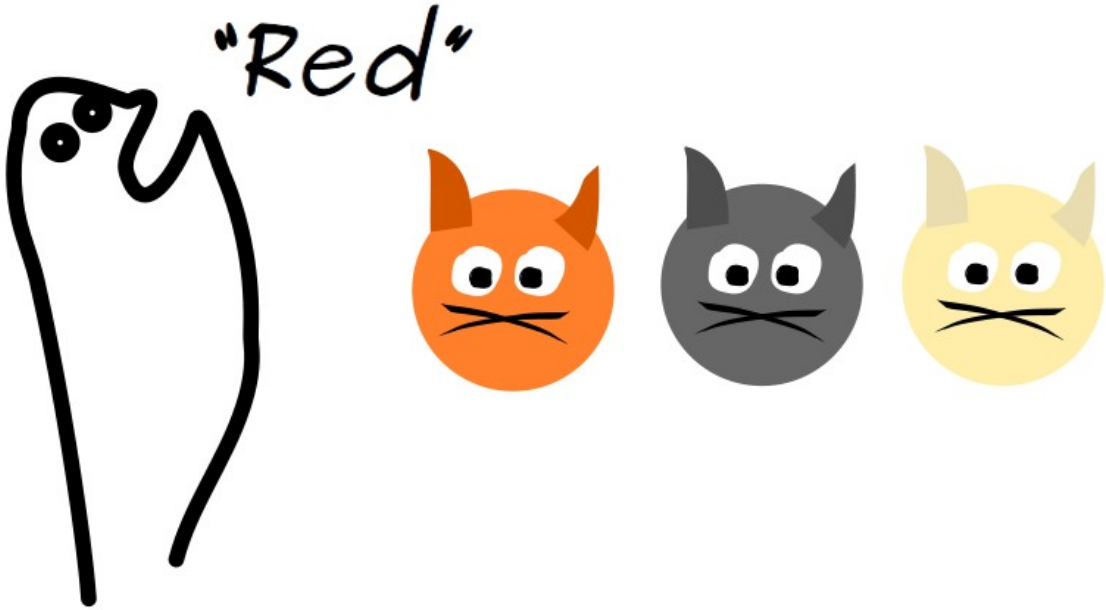
Expression meaning vs. speaker meaning (Grice ‘68)



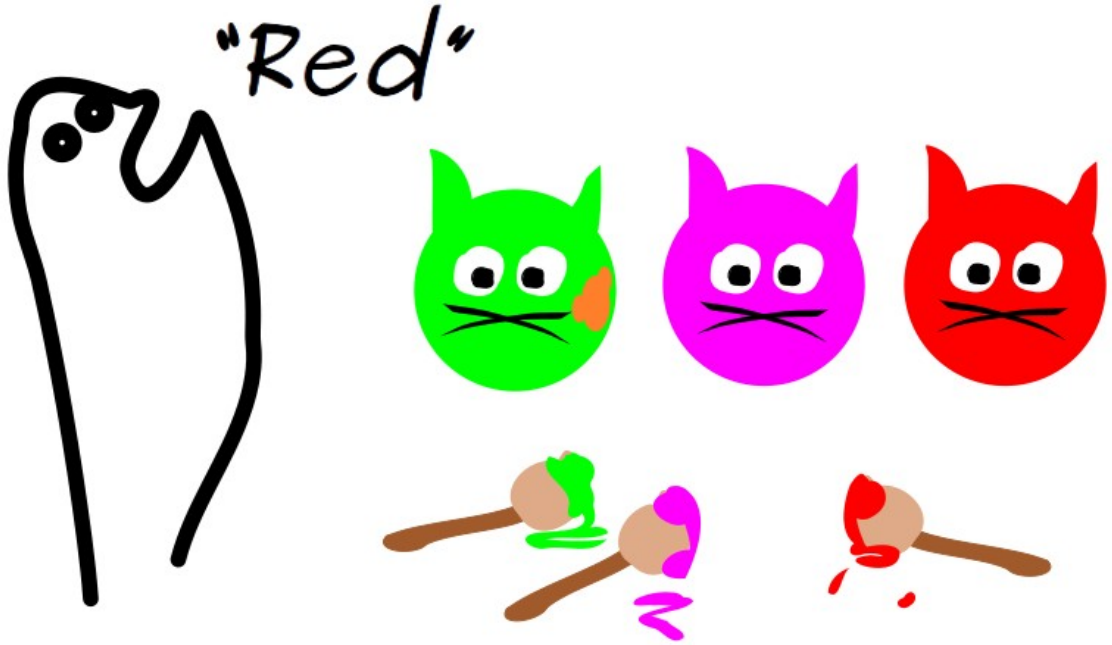
Expression meaning vs. speaker meaning (Grice ‘68)



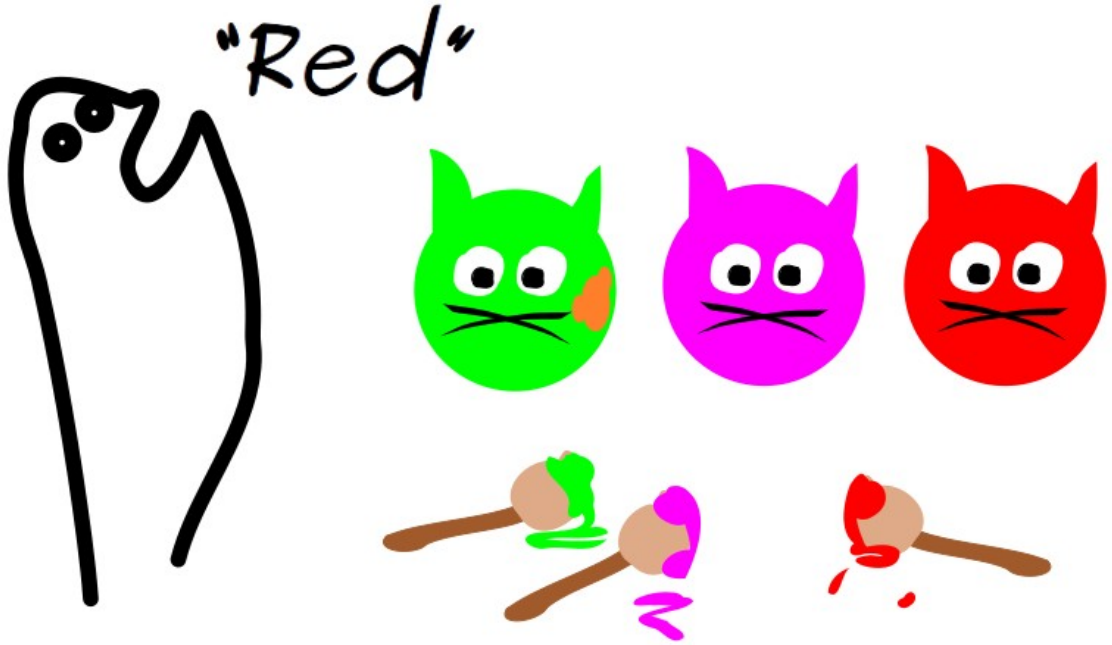
Expression meaning vs. speaker meaning (Grice ‘68)



Expression meaning vs. speaker meaning (Grice ‘68)

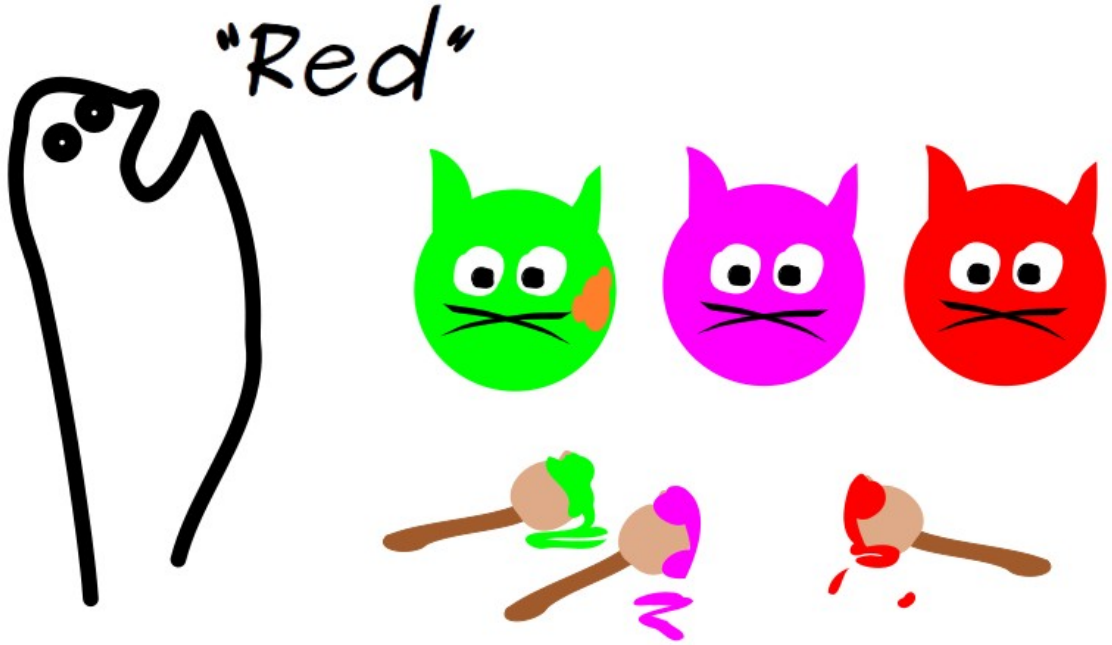


Expression meaning vs. speaker meaning (Grice ‘68)



Speaker meaning:
The speaker’s communicative intention (varies a lot).

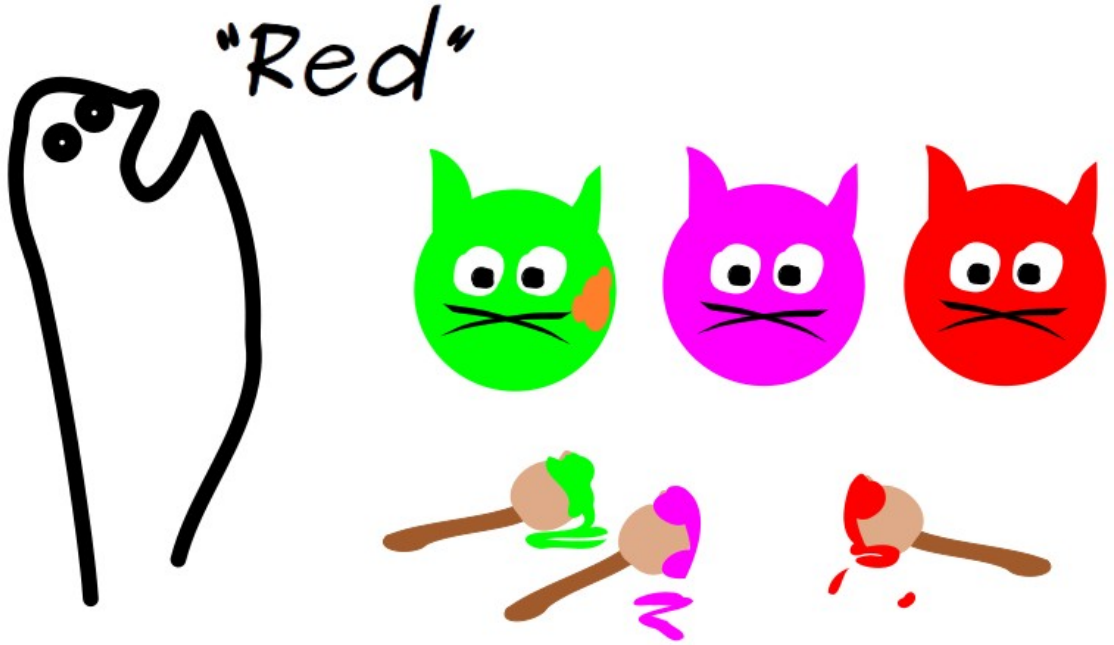
Expression meaning vs. speaker meaning (Grice ‘68)



Speaker meaning:
The speaker’s communicative intention (varies a lot).

Expression meaning:
Some kind of fixed backbone of all uses.

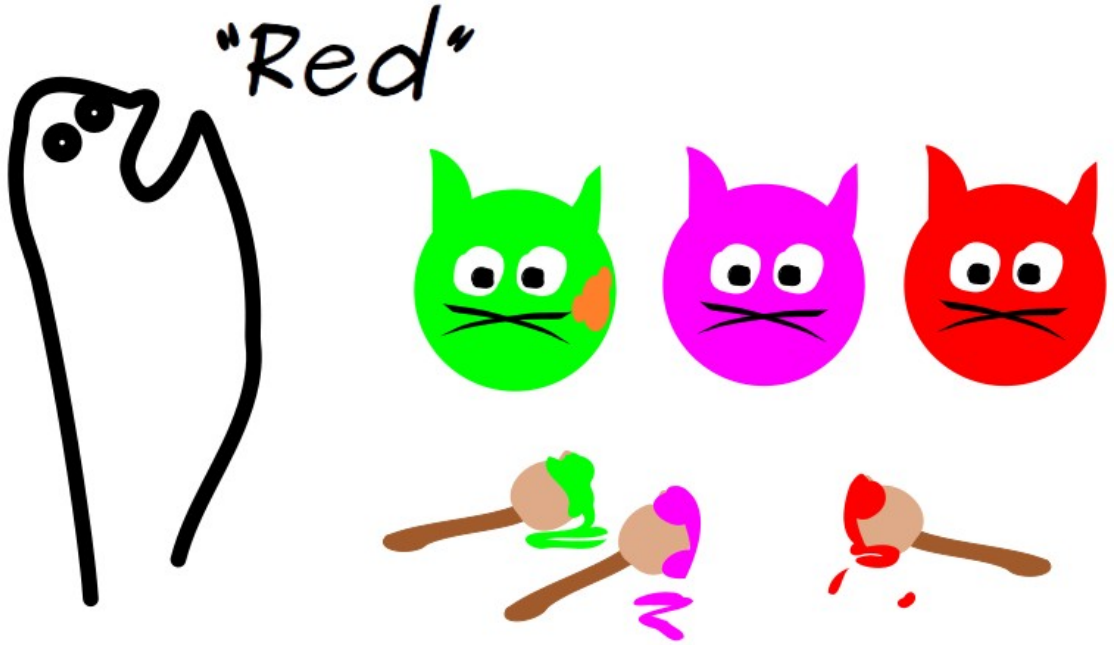
Expression meaning vs. speaker meaning (Grice ‘68)



Pragmatics: **Speaker meaning**:
The speaker’s communicative intention (varies a lot).

Expression meaning:
Some kind of fixed backbone of all uses.

Expression meaning vs. speaker meaning (Grice ‘68)



Pragmatics: **Speaker meaning**:
The speaker’s communicative intention (varies a lot).

Semantics: **Expression meaning**:
Some kind of fixed backbone of all uses.

Q: What was the point of Grice’s seminal paper “Logic and conversation”?

Grice’s cooperative principle

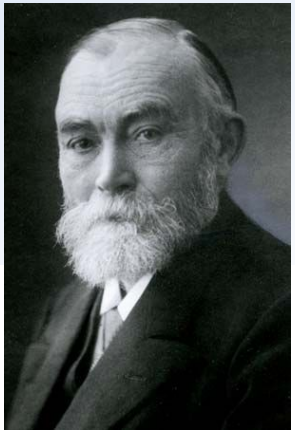
- Quality: be truthful
- Relation: be relevant
- Quantity: provide the right amount of information
- Manner: be clear, concise

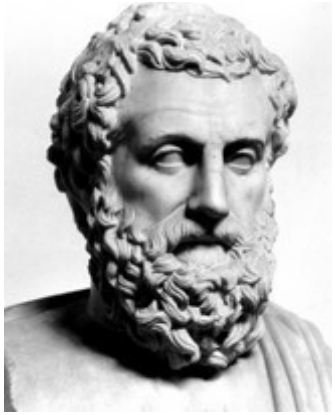
Grice’s cooperative principle

- Quality: be truthful
- Relation: be relevant
- Quantity: provide the right amount of information
- Manner: be clear, concise

But these maxims were *not* the main point...

Historical intermezzo





λv (λw λP λQ ∀x [P(w)(λu.x) → Q(w)(λu.x)])(v) (λu λ

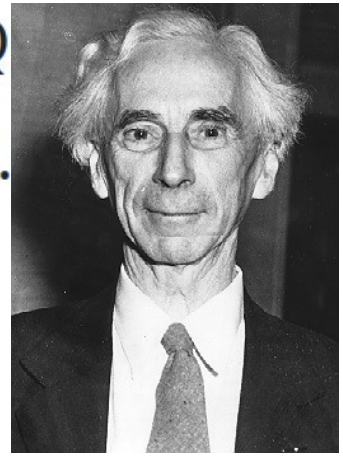
λv (λP λQ ∀x [P(v)(λu.x) → Q(v)(λu.x)]) (λu λy. STU

λv (λQ ∀x [((λu λy.STUDENT(u)(y(u))))(v)(λu.x) → Q

λv (λQ ∀x [(λy.STUDENT(v)(y(v)))(λu.x) → Q(v)(λu.x

λv (λQ ∀x [STUDENT(v)(λu.x (v)) → Q

λv (λQ ∀x [STUDENT(v)(x) → Q(v)(λu.



∃x A(x,y) ∧ B(x)

∃x∃y A(x,y) → B(x)

∃x∃y[A(x,y) → B(x)]

¬∃x∃y A(x,y) → B(x)

∀x¬∃y A(x,y)

¬B(x) → (¬∀y[¬A(x,y) v B(x)] → C(y))

∃x[A(x,y) v B(y)]

Historical intermezzo

For a large class of cases – though not for all – in which we employ the word "meaning" it can be defined thus: ***the meaning of a word is its use in the language****.*

(Wittgenstein, PU 43)

Historical intermezzo

*We become **obsessed** with "truth" when discussing statements, just as we become obsessed with "freedom" when discussing conduct.*

(Austin, 1950a: p.130)

...our ordinary words are much subtler in their uses, and mark many more distinctions, than philosophers [≈ logicians] have realized.

(Austin, 1962, p.3)

Historical intermezzo

There is in my opinion no important theoretical difference between natural languages and the artificial languages of logicians; [...].

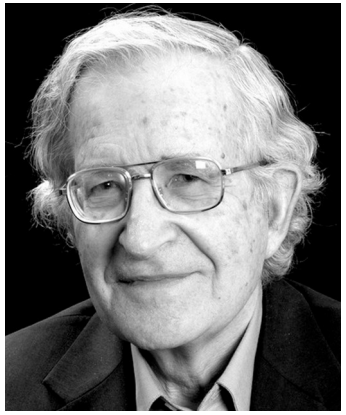
On this point I differ from a number of philosophers, but agree, I believe, with Chomsky and his associates.

(Montague, 1970, *Universal Grammar)*

Historical intermezzo

So there is that part of syntax and there certainly is pragmatics in some general sense of what you do with words and so on. But whether there is semantics in the more technical sense is an open question. I don't think there's any reason to believe that there is.

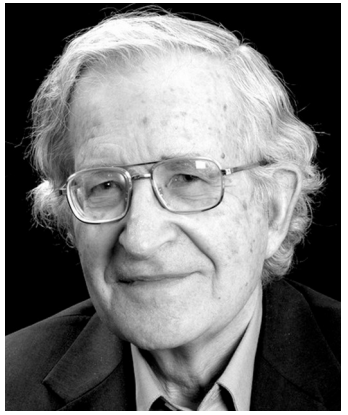
(Chomsky, 2000, p.73)



Historical intermezzo

So there is that part of syntax and there certainly is pragmatics in some general sense of what you do with words and so on. But whether there is semantics in the more technical sense is an open question. I don't think there's any reason to believe that there is.

(Chomsky, 2000, p.73)



Historical intermezzo

It is a commonplace of philosophical logic that there are, or appear to be, ***divergences in meaning*** *between, on te one hand, the* ***formal devices*** *–* ¬, ∧, ∨, → *– and, on the other, what are taken to be their analogues or counterparts in* ***natural language*** *– such expressions as* not, and, or, if...

(Grice 1967)

I have no intention of entering the fray on behalf of either contestant. I wish, rather, to maintain that the common assumption of the contestants that the divergences do in fact exist is (broadly speaking) a common mistake, and that ***the mistake arises from inadequate attention to the nature and importance of the conditions governing conversation****.*

Historical intermezzo

It is a commonplace of philosophical logic that there are, or appear to be, ***divergences in meaning*** *between, on te one hand, the* ***formal devices*** *– ¬, ^, ˅, → – and, on the other, what are taken to be their analogues or counterparts in* ***natural language*** *– such expressions as* not, and, or, if...

(Grice 1967)

I have no intention of entering the fray on behalf of either contestant. I wish, rather, to maintain that the common assumption of the contestants that the divergences do in fact exist is (broadly speaking) a common mistake, and that ***the mistake arises from inadequate attention to the nature and importance of the conditions governing conversation****.*

The history in a nutshell

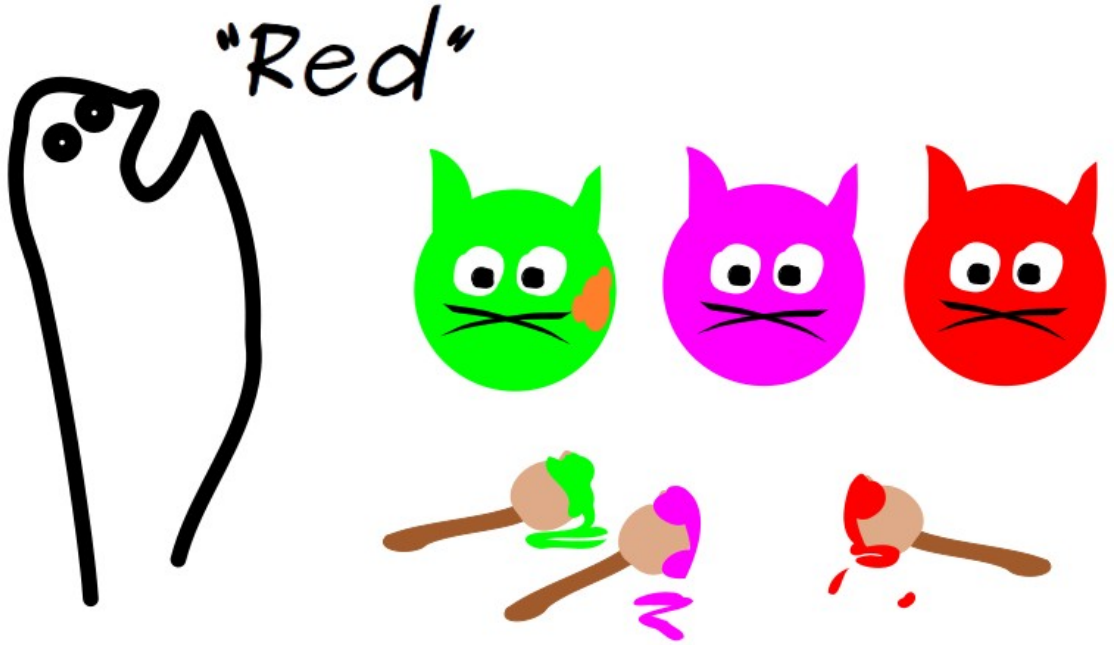
- Let’s analyze natural language meaning with logic!
- No, language seems too illogical…
- Well, actually, the language itself might be very logical, it is what speakers *do* with it that tends to mess things up a bit.

The history in a nutshell

- Let’s analyze natural language meaning with logic!
- No, language seems too illogical…
- Well, actually, the language itself might be very logical, it is what speakers *do* with it that tends to mess things up a bit.

The ‘standard picture’:
formal semantics + Gricean pragmatics.

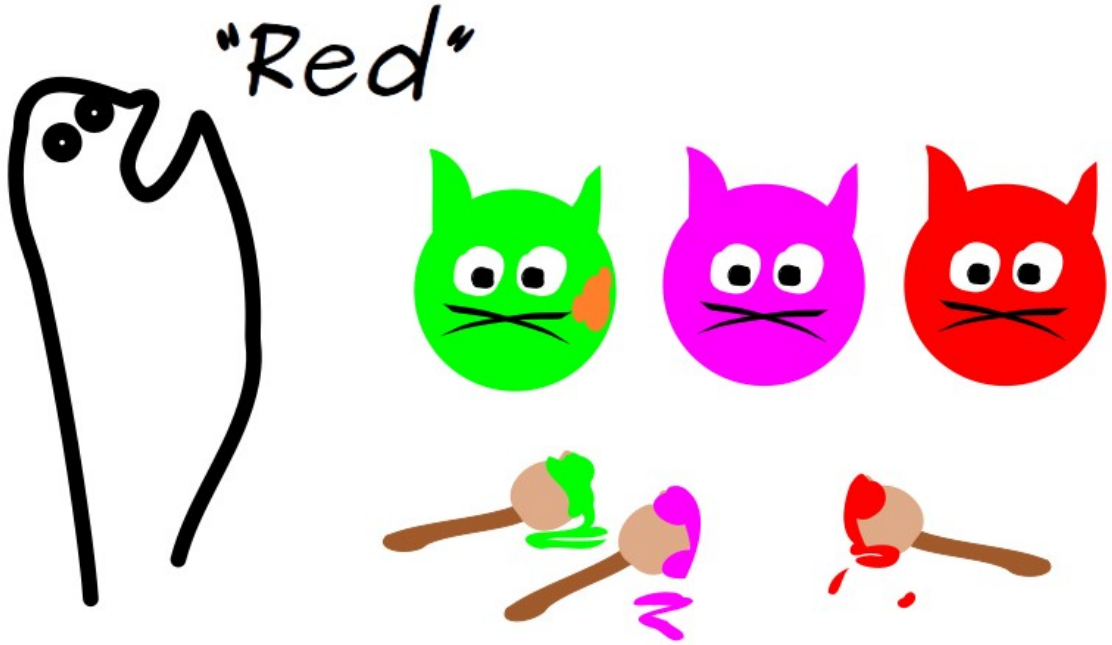
Expression meaning vs. speaker meaning (Grice ‘68)



Pragmatics: **Speaker meaning**:
The speaker’s communicative intention (varies a lot).

Semantics: **Expression meaning**:
Some kind of fixed backbone of all uses.

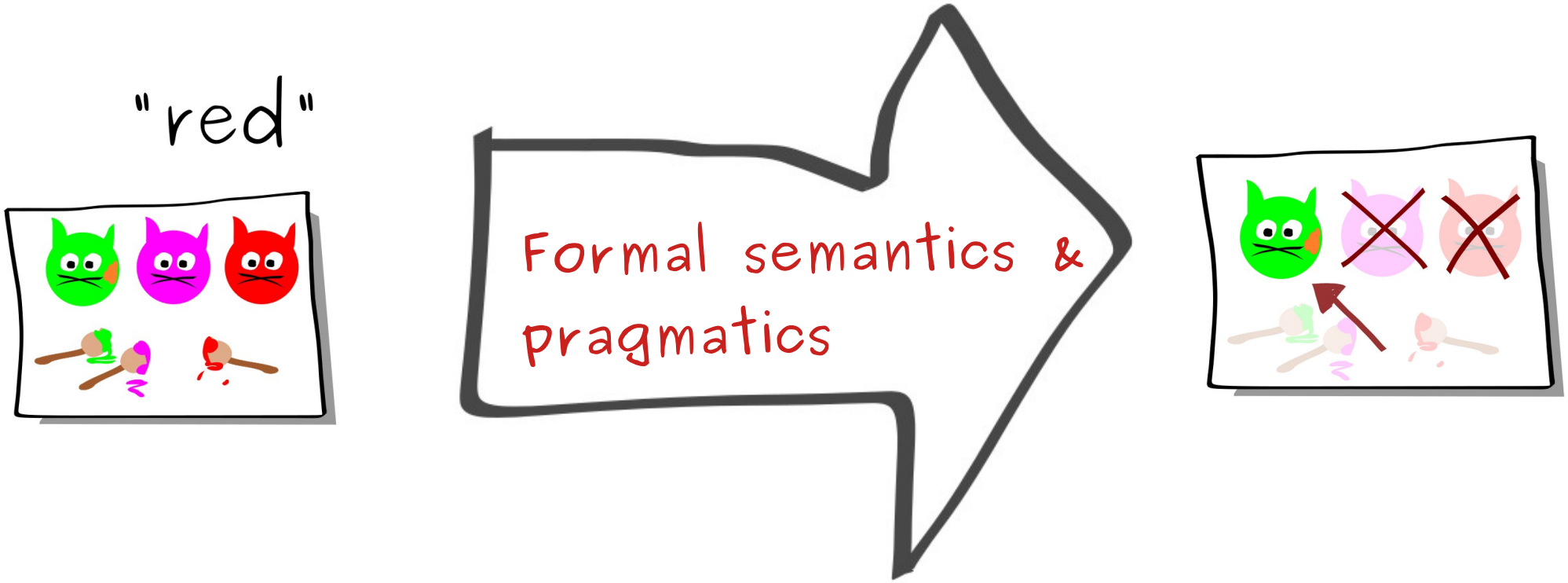
Expression meaning vs. speaker meaning (Grice ‘68)



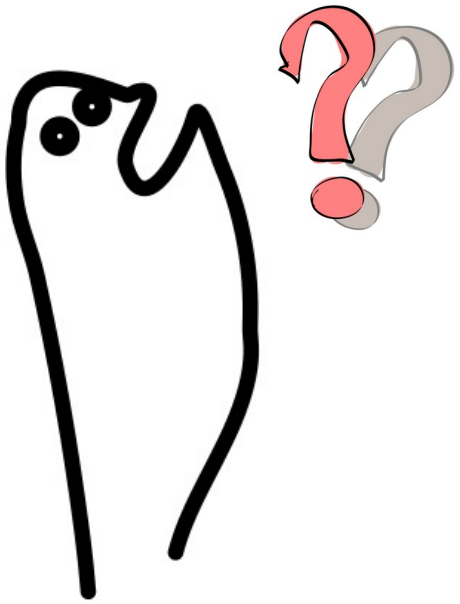
Pragmatics: **Speaker meaning**:
The speaker’s communicative intention (varies a lot).

FORMAL Semantics: **Expression meaning**:
Some kind of fixed backbone of all uses.

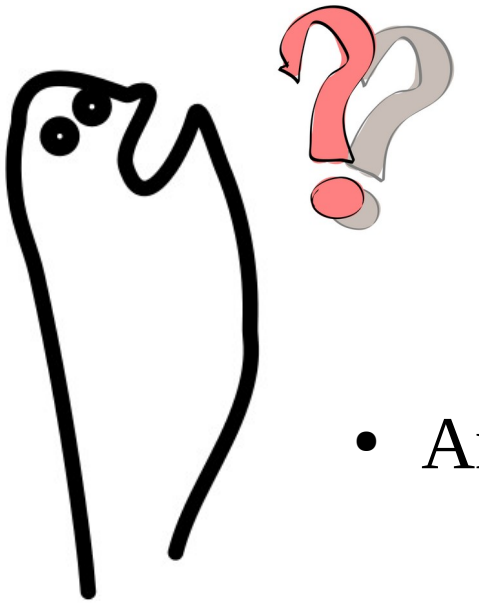
The ‘standard picture’



Main questions at this point

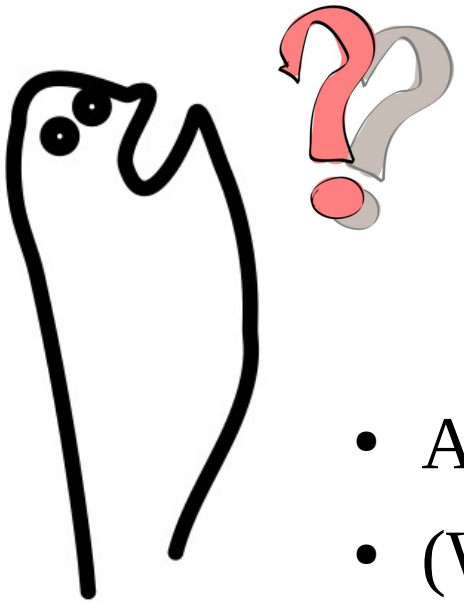


Main questions at this point



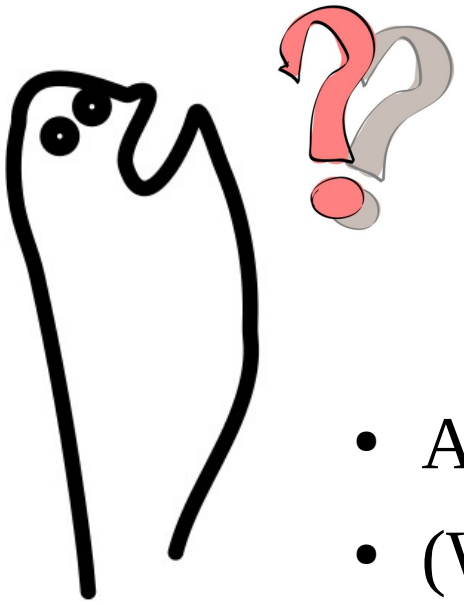
- And they lived happily ever after?

Main questions at this point



- And they lived happily ever after?
- (Why might Chomsky not believe in semantics?)

Main questions at this point

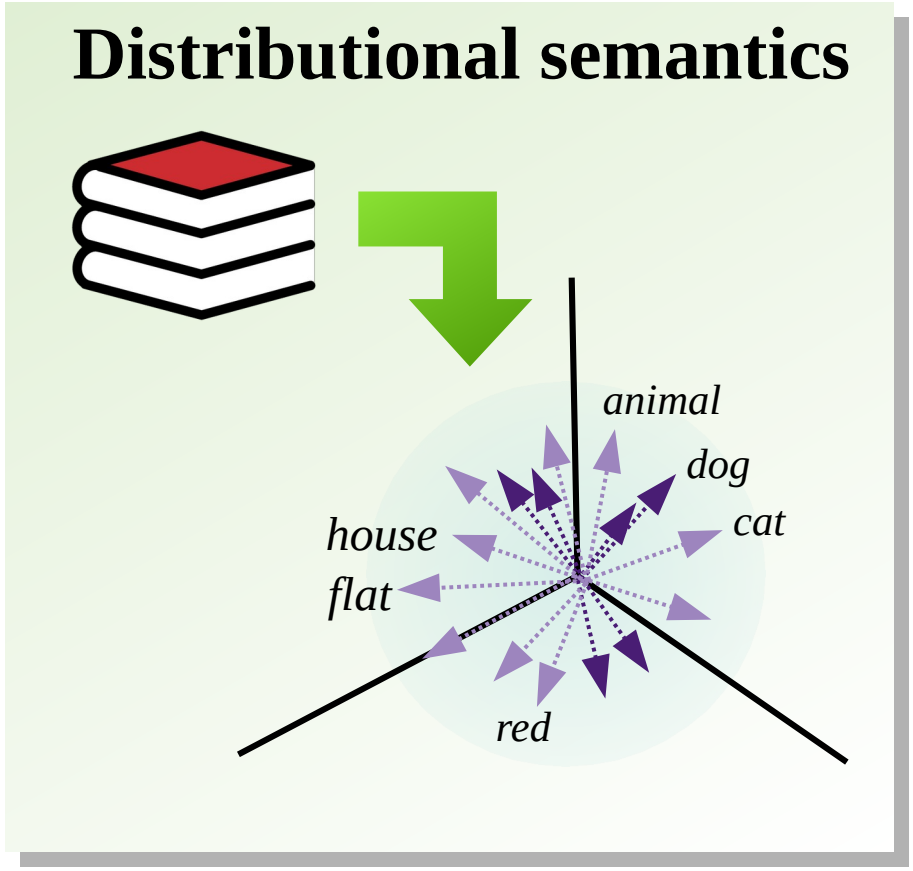


- And they lived happily ever after?
- (Why might Chomsky not believe in semantics?)
- What else was in the title of this talk again?

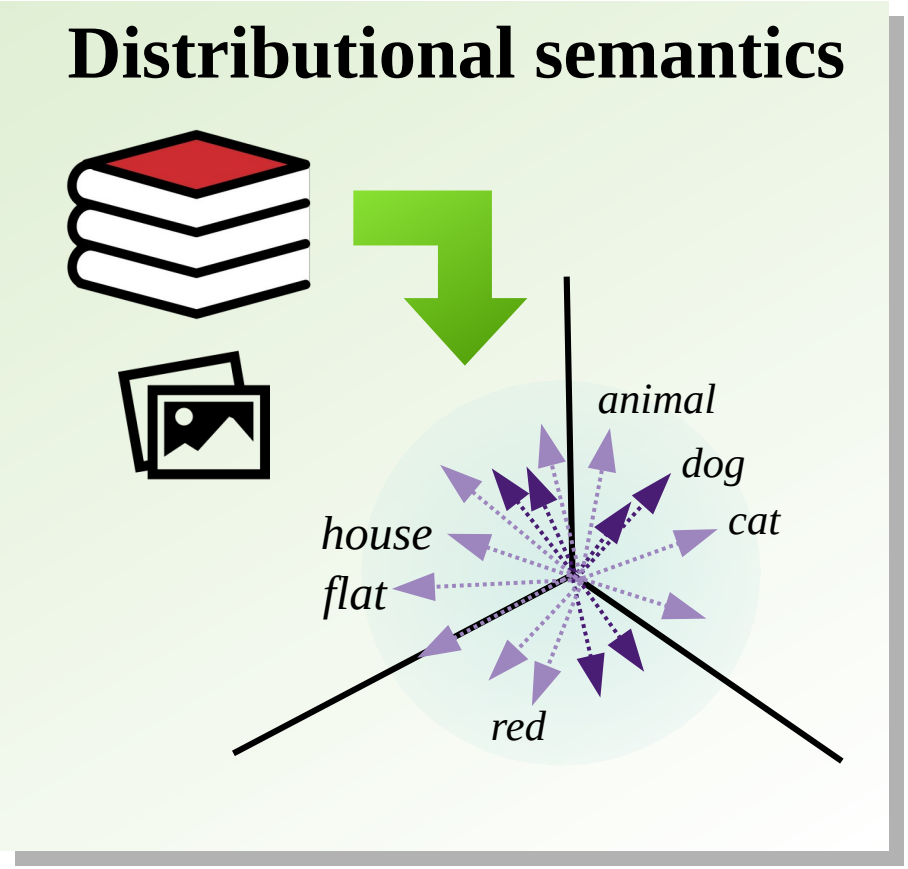
Part 2. A radically different approach: distributional semantics and deep learning

Q: What is “distributional semantics”?

A vector-space model of meaning



A vector-space model of meaning



Early roots

For a large class of cases – though not for all – in which we employ the word "meaning" it can be defined thus: ***the meaning of a word is its use in the language****.*

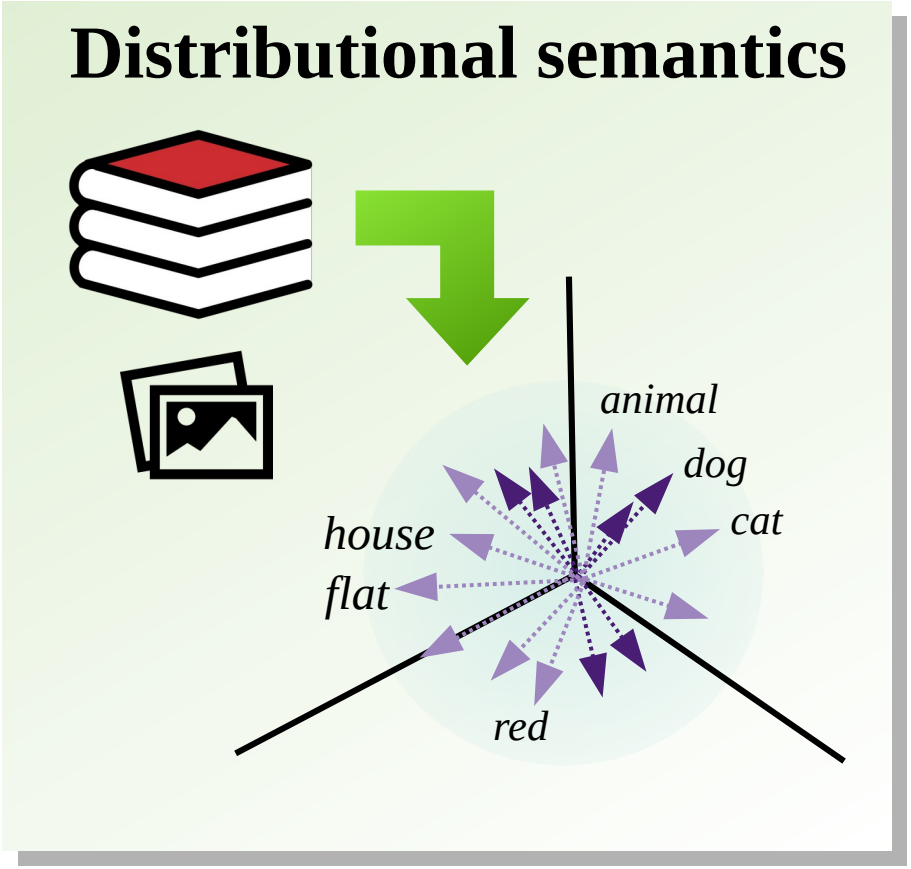
(Wittgenstein, PU 43)

Early adoption in linguistics

As Wittgenstein says, 'the meaning of words lies in their use.' The day-to-day practice of playing language games recognizes customs and rules. It follows that a text in such established usage may contain sentences such as 'Don't be such an ass!', 'You silly ass!', 'What an ass he is!' In these examples, the word ass is in familiar and habitual company, commonly collocated with you silly—, he is a silly—, don't be such an—. You shall know a word by the company it keeps!

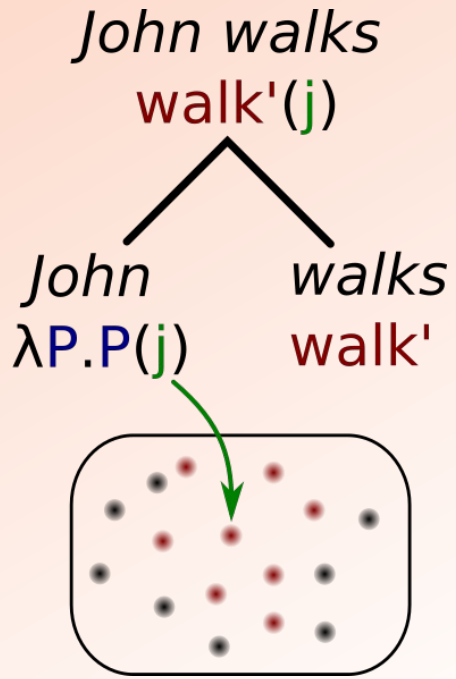
(Firth, 1957)

A vector-space model of meaning

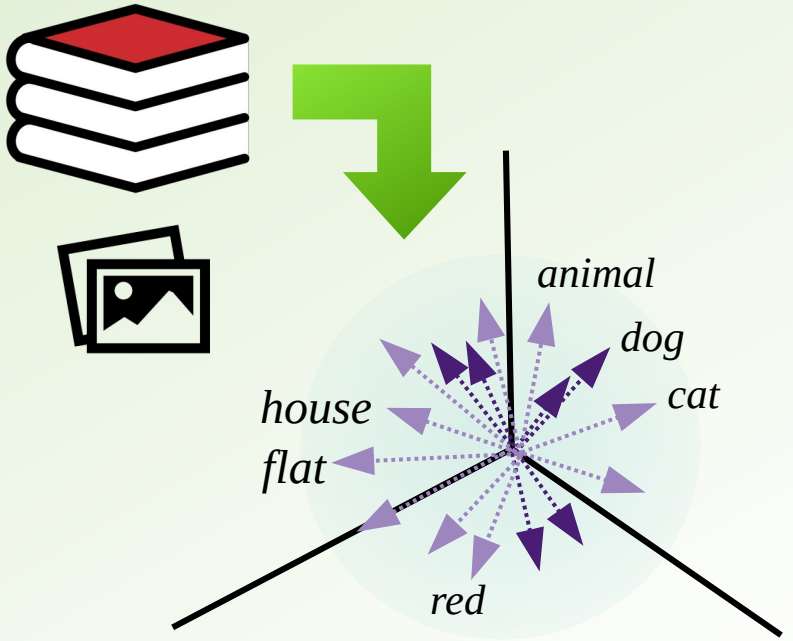


A vector-space model of meaning

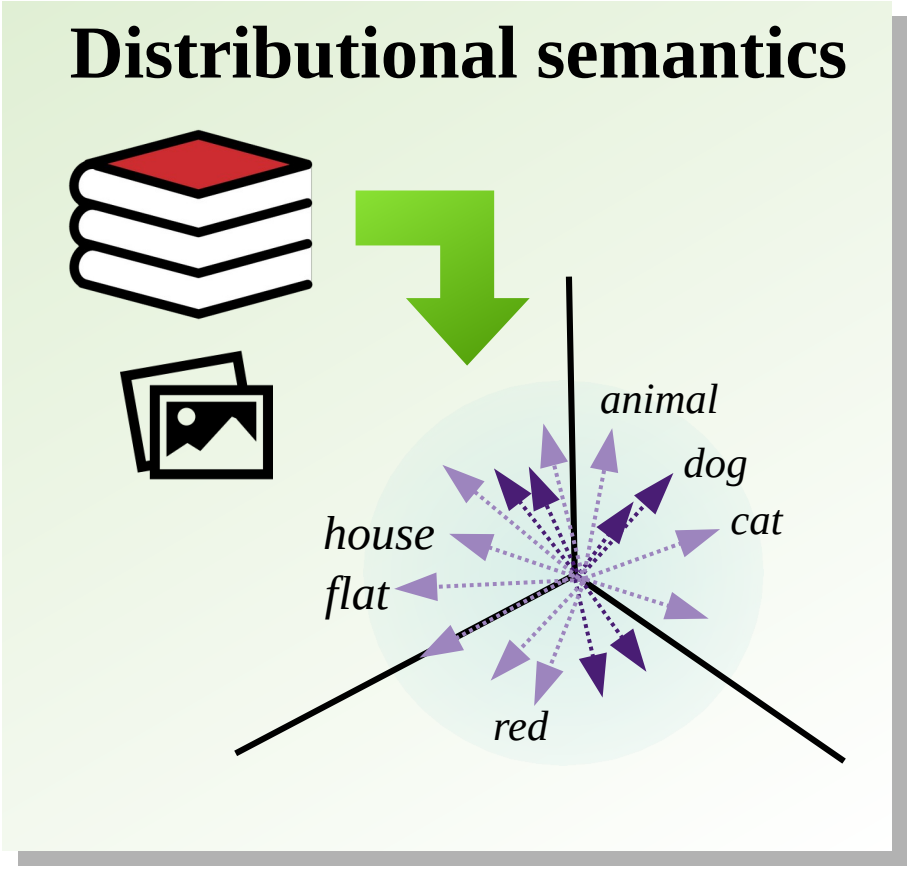
Formal semantics



Distributional semantics

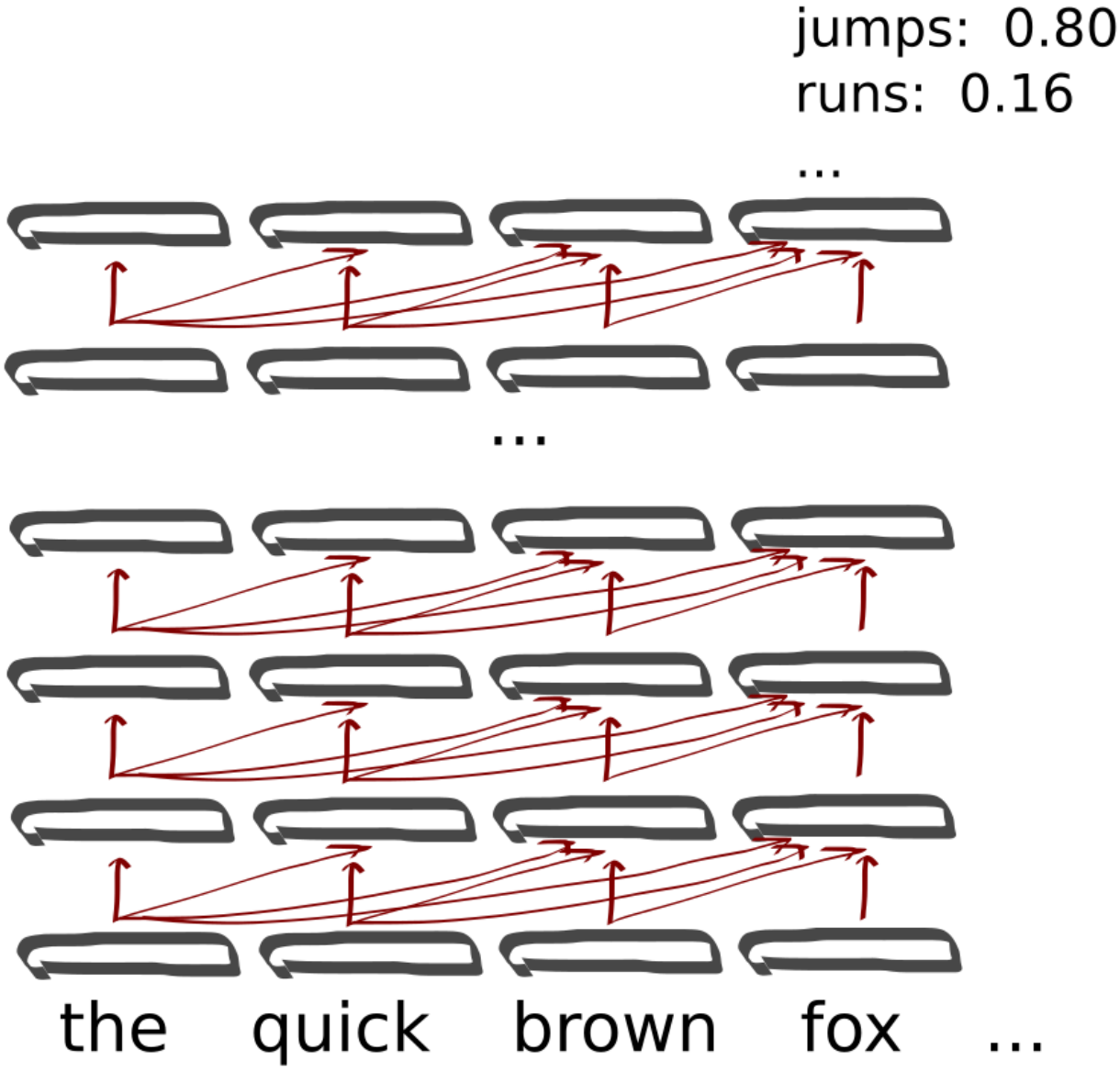


A vector-space model of meaning

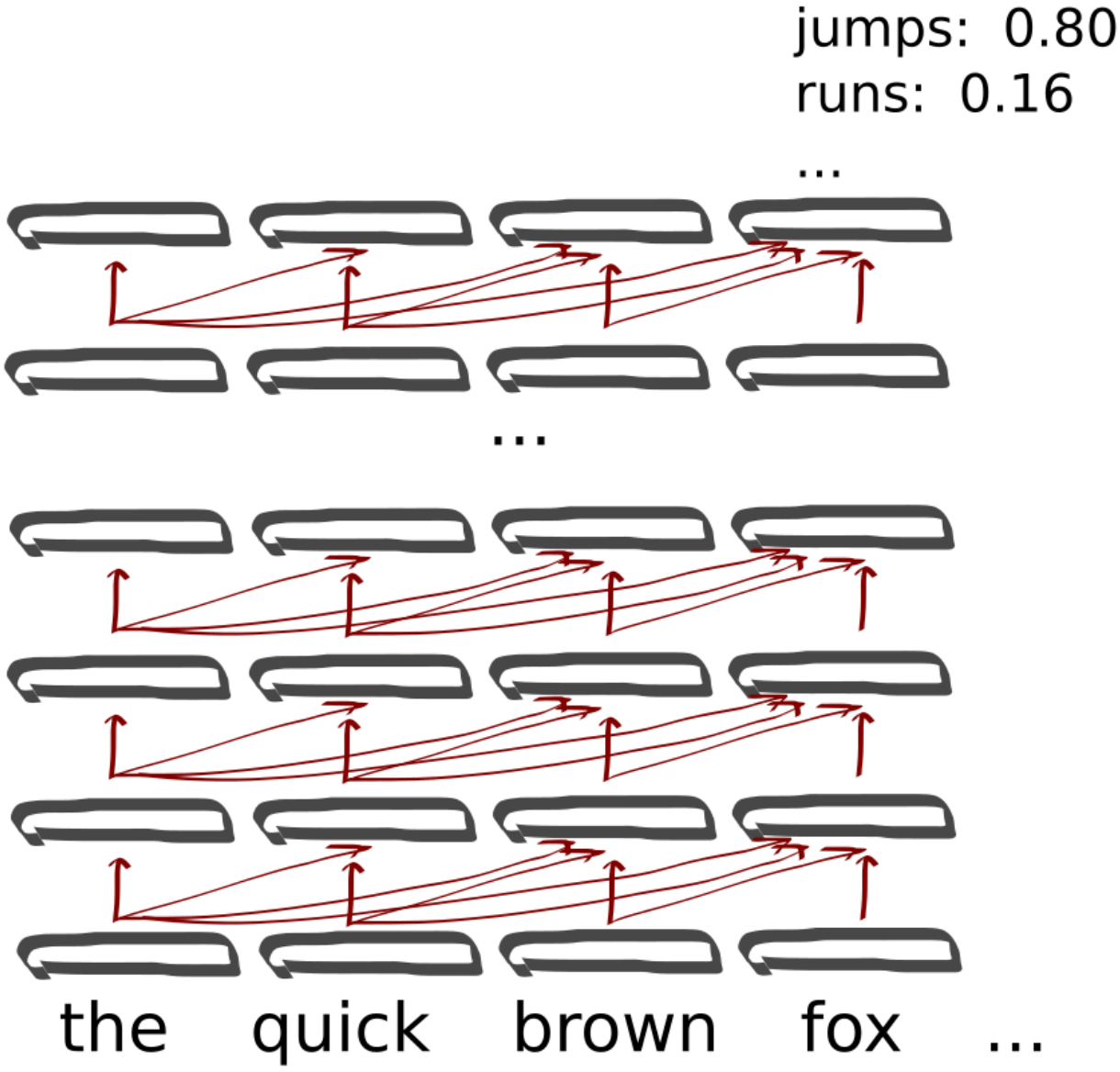


Modern intermezzo: ChatGPT

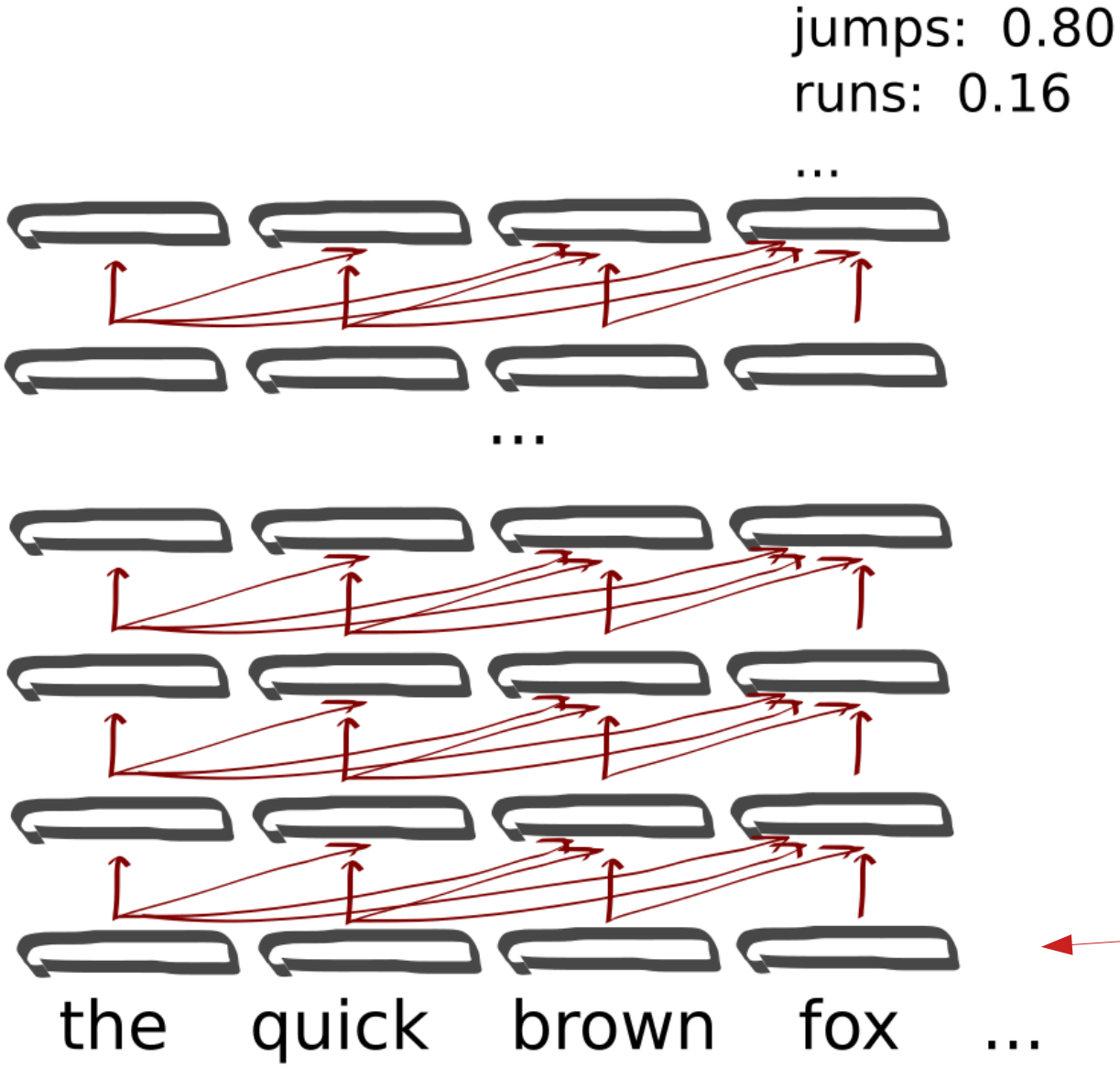
Modern intermezzo: ChatGPT



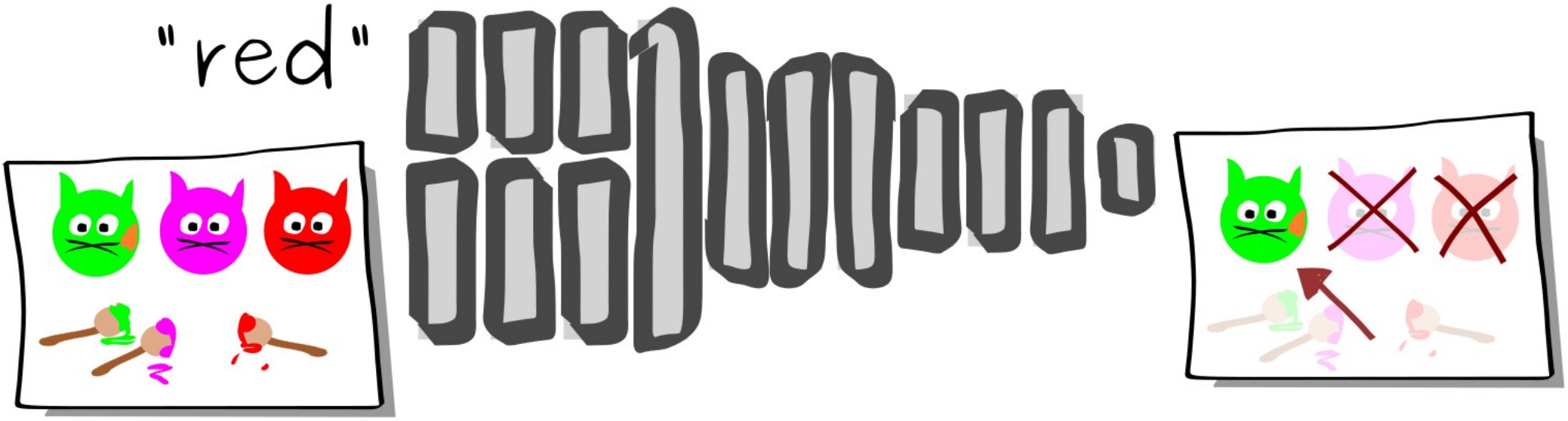
Modern intermezzo: ChatGPT



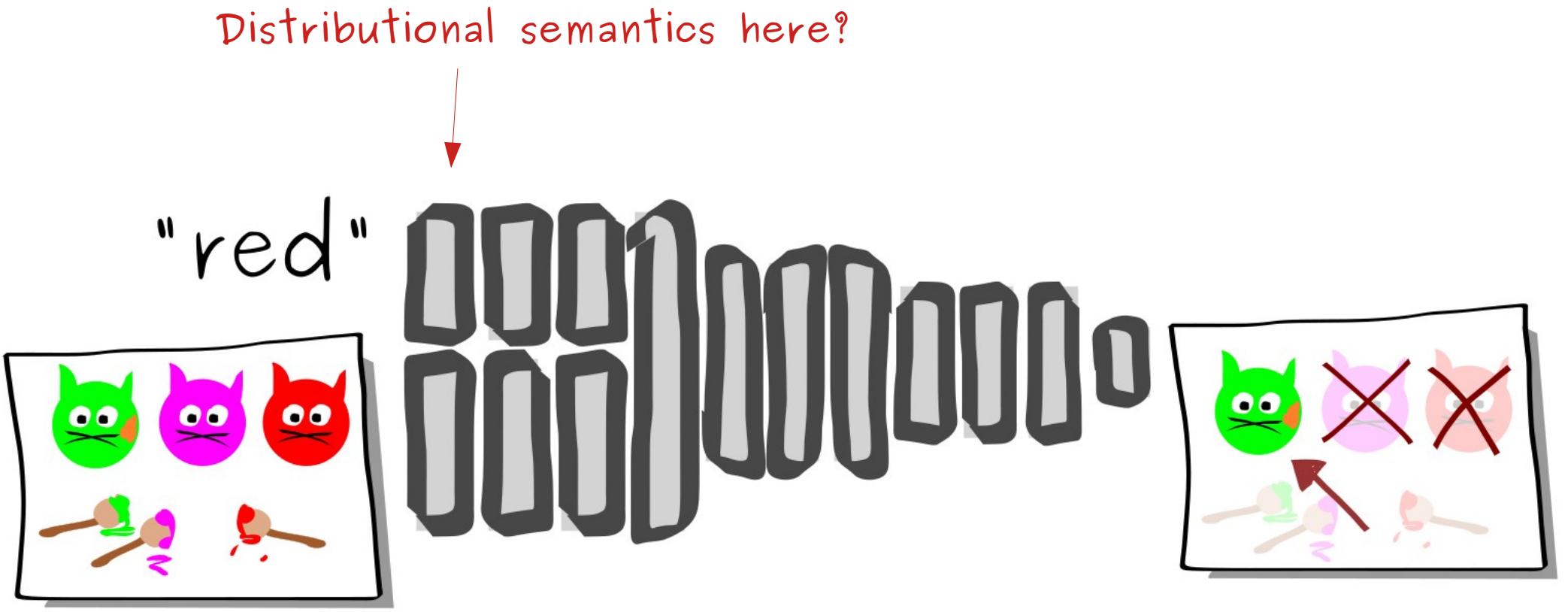
Modern intermezzo: ChatGPT



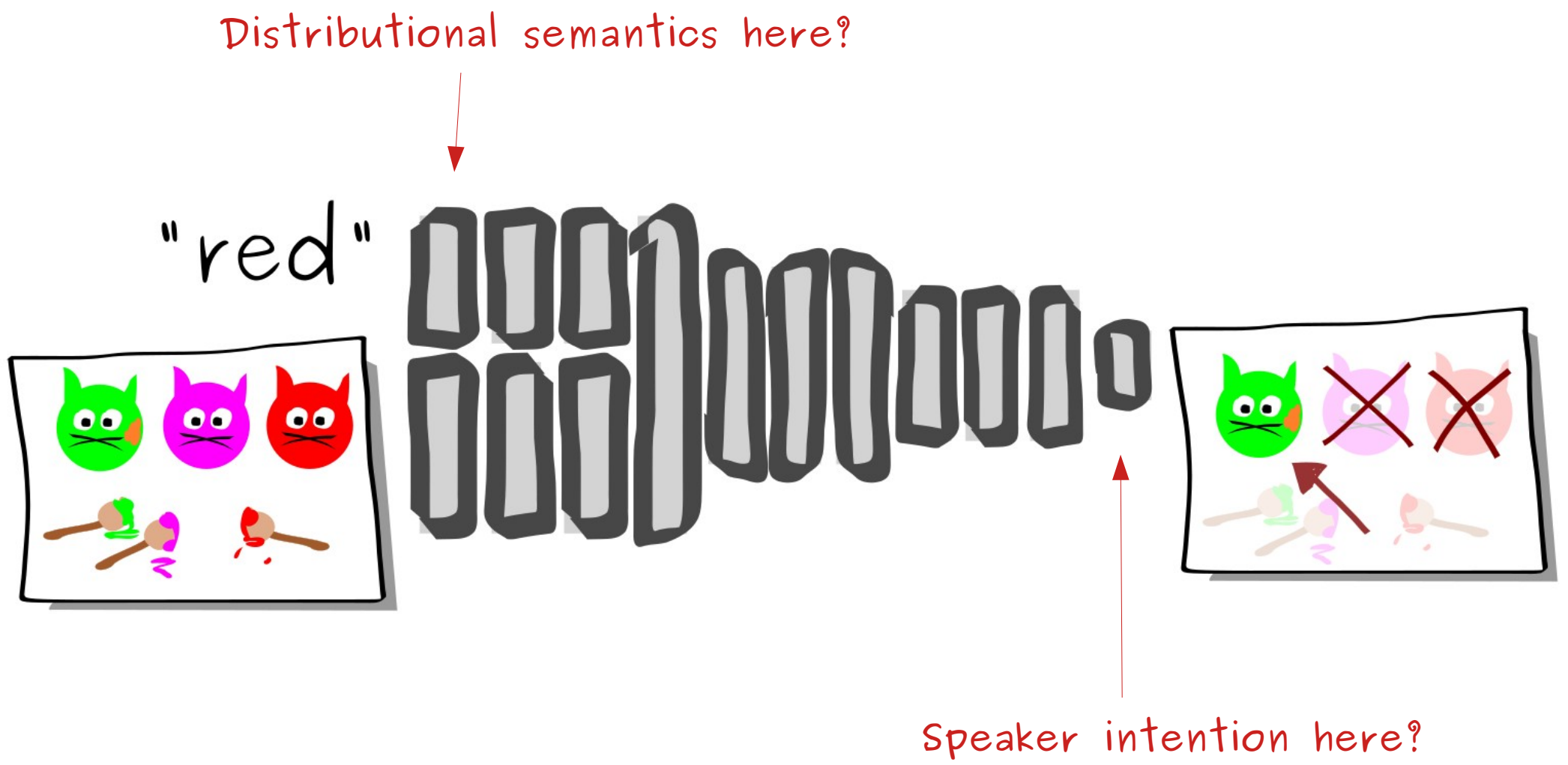
An interesting example: Visual information retrieval



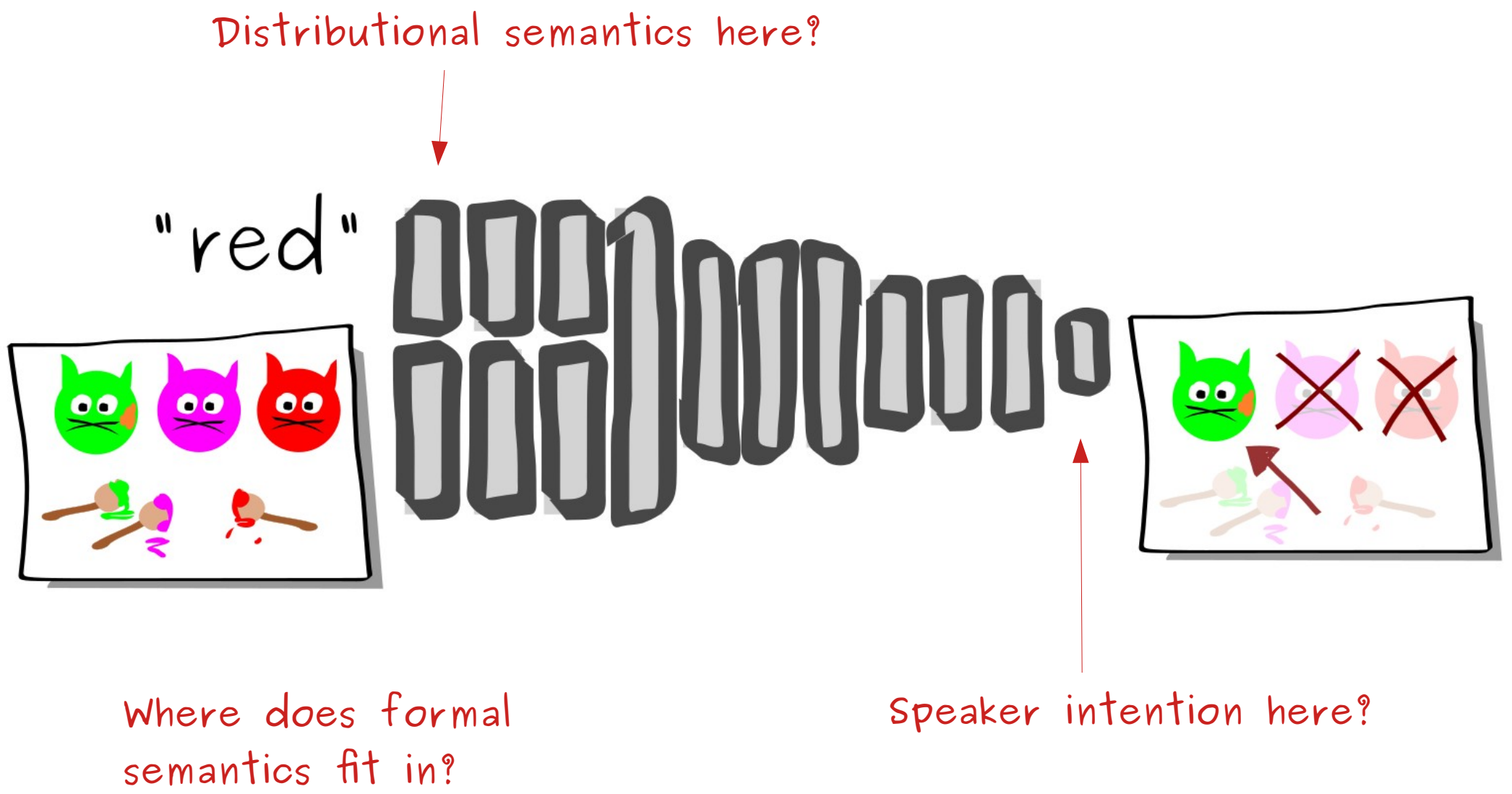
An interesting example: Visual information retrieval



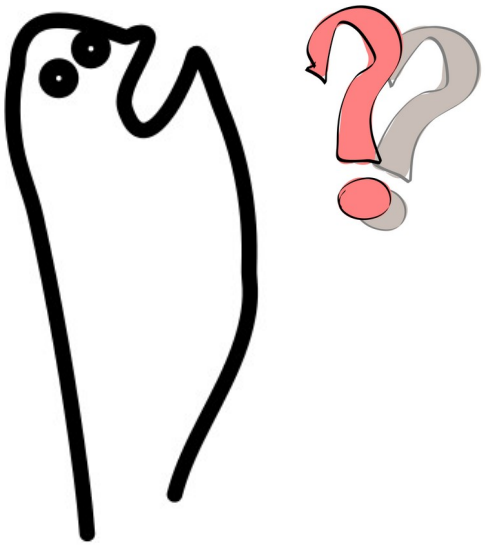
An interesting example: Visual information retrieval



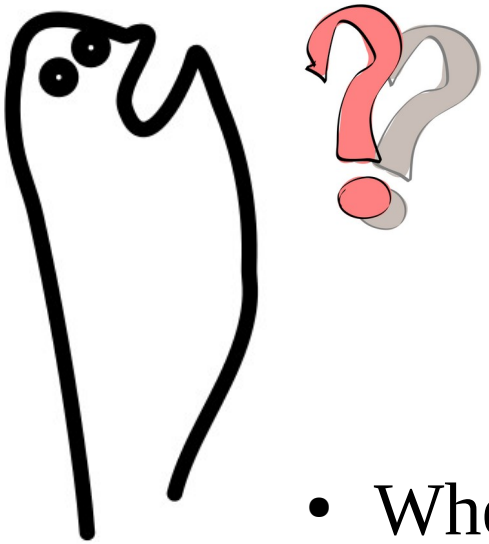
An interesting example: Visual information retrieval



Main questions at this point

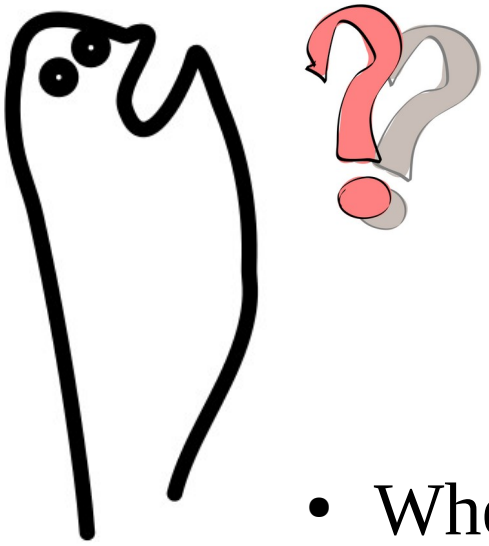


Main questions at this point



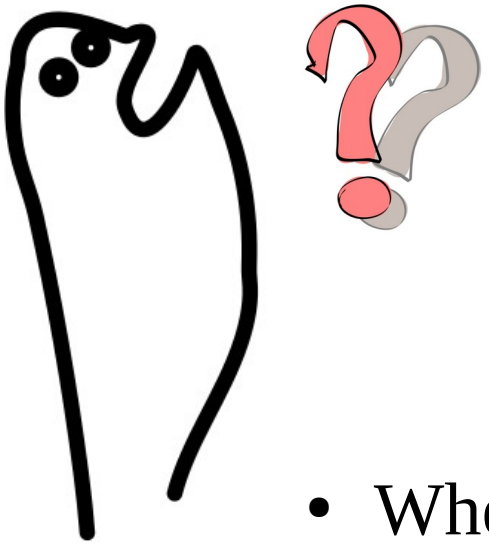
- Where does formal semantics fit in deep learning?

Main questions at this point



- Where does formal semantics fit in deep learning?
- (In what way is the ‘standard picture’ wrong?)

Main questions at this point



- Where does formal semantics fit in deep learning?
- (In what way is the ‘standard picture’ wrong?)
- (Why might Chomsky not believe in semantics?)

Part 3.
What the ‘standard picture’ gets wrong & how the two approaches may be integrated

Q: First, are there any other questions at this point?

Part 3.
What the ‘standard picture’ gets wrong & how the two approaches may be integrated

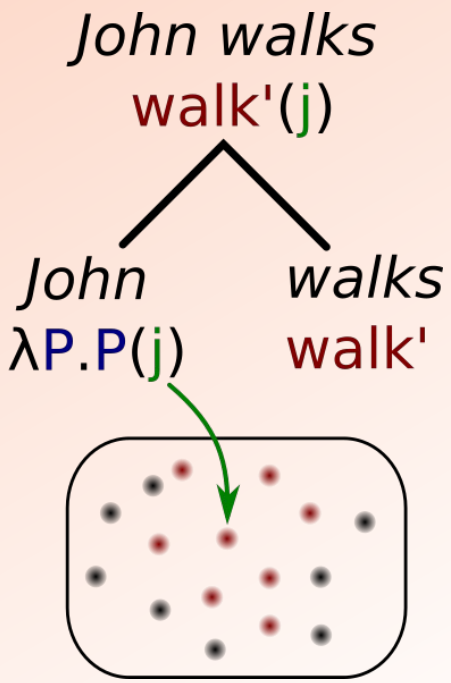
Part 3.
What the ‘standard picture’ gets wrong & how the two approaches may be integrated

Based on Westera & Boleda (2019). *Don’t Blame Distributional Semantics if it can’t do Entailment*.

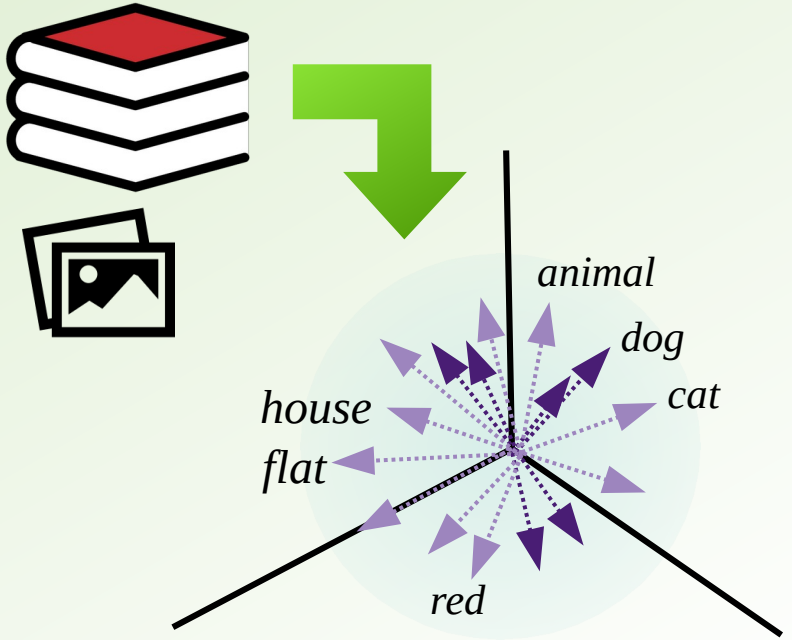
What is an adequate model of *expression meaning*?

What is an adequate model of *expression meaning*?

Formal semantics

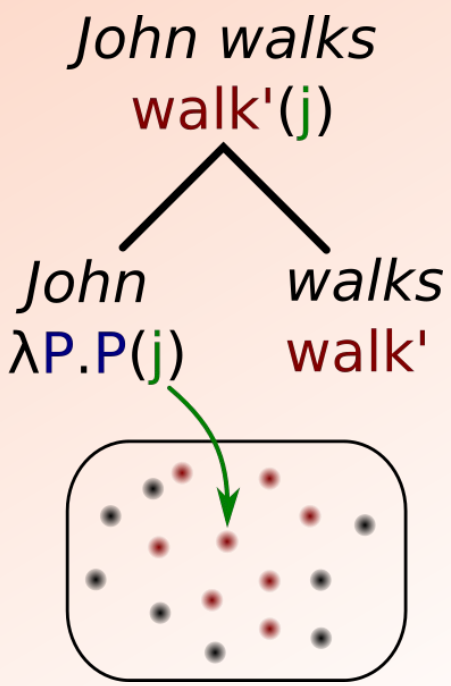


Distributional semantics

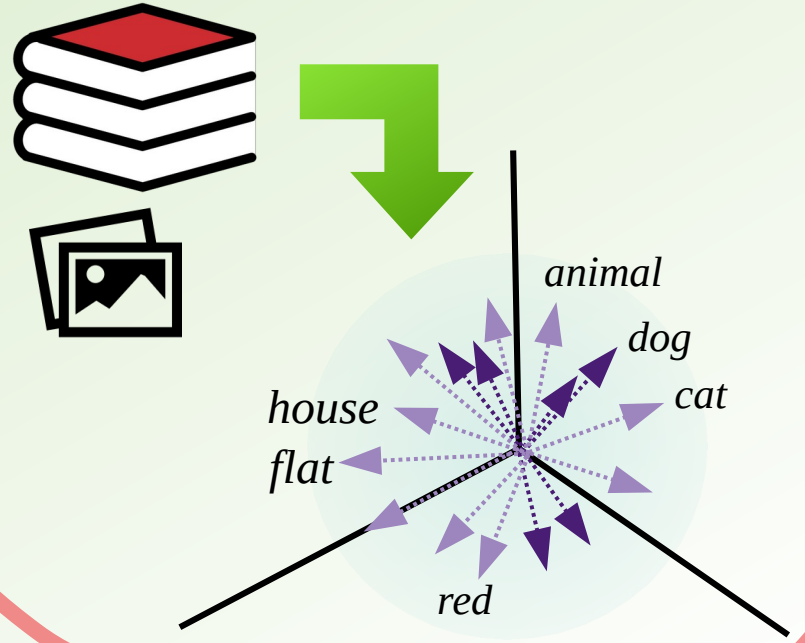


What is an adequate model of *expression meaning*?

Formal semantics

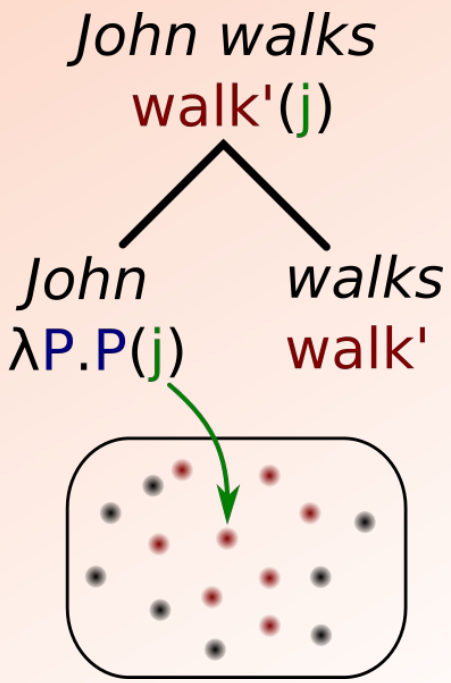


Distributional semantics

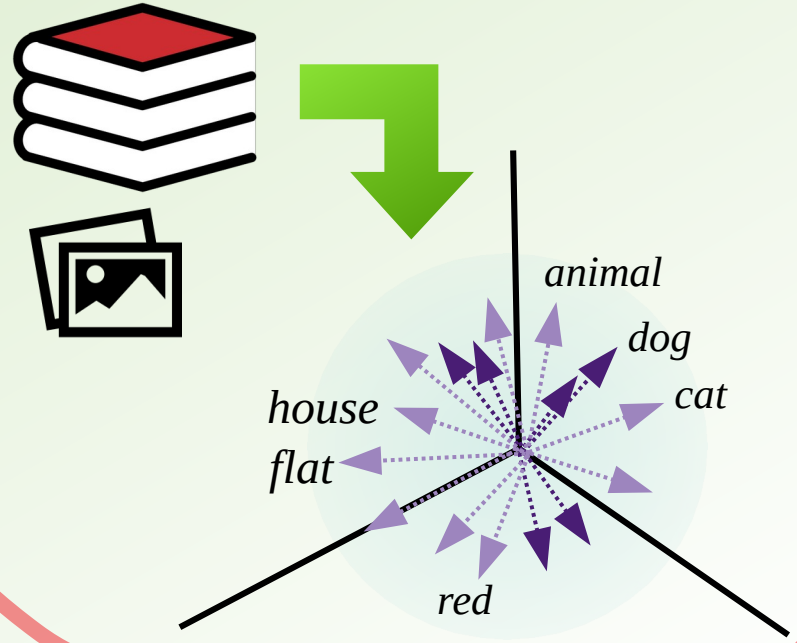


What is an adequate model of *expression meaning*?

Formal semantics

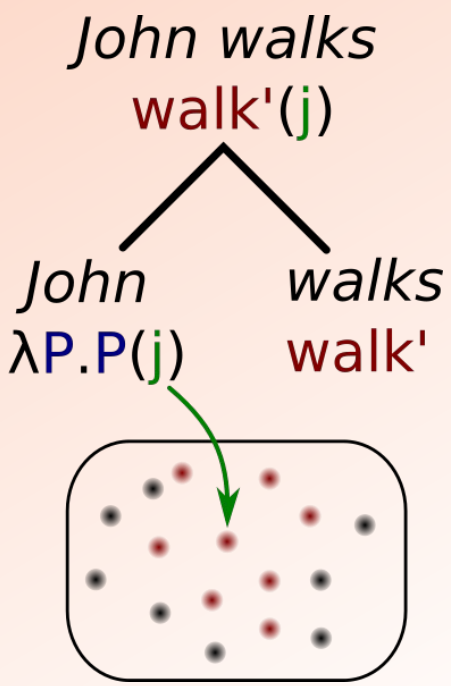


Distributional semantics

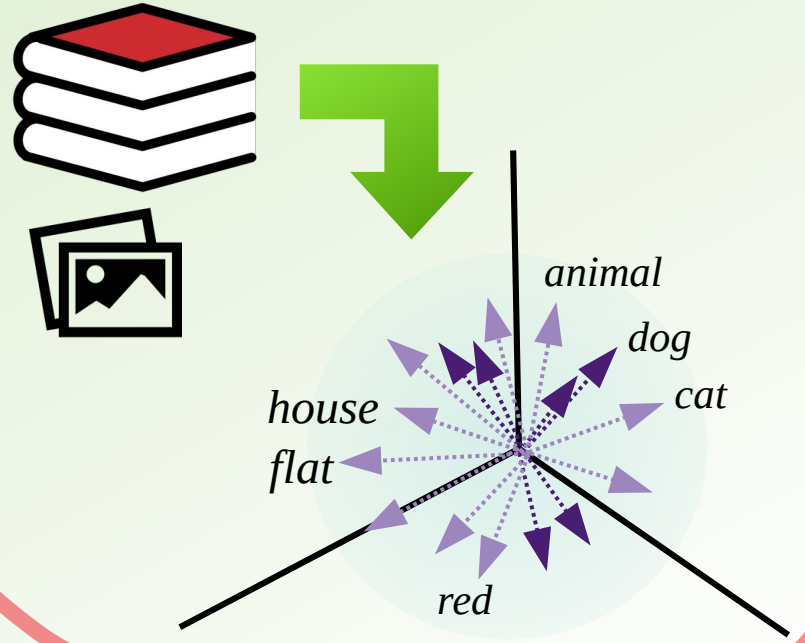


What is an adequate model of *expression meaning*?

Formal semantics



Distributional semantics



What does distributional semantics model?

Distributional semantics has emerged as a promising model of certain 'conceptual' aspects of linguistic meaning (e.g., Landauer and Dumais 1997; Turney and Pantel 2010; Baroni and Lenci 2010; Lenci 2018) and as an indispensable component of applications in Natural Language Processing (e.g., reference resolution, machine translation, image captioning; especially since Mikolov et al. 2013).

Yet its theoretical status within a general theory of meaning and of language and cognition more generally is not clear (e.g., Lenci 2008; Erk 2010; Boleda and Herbelot 2016; Lenci 2018). In particular, ***it is not clear whether distributional semantics can be understood as an actual model of expression meaning*** *– what Lenci (2008) calls the 'strong' view of distributional semantics – or merely as a model of something that correlates with expression meaning in certain partial ways – the 'weak' view.*

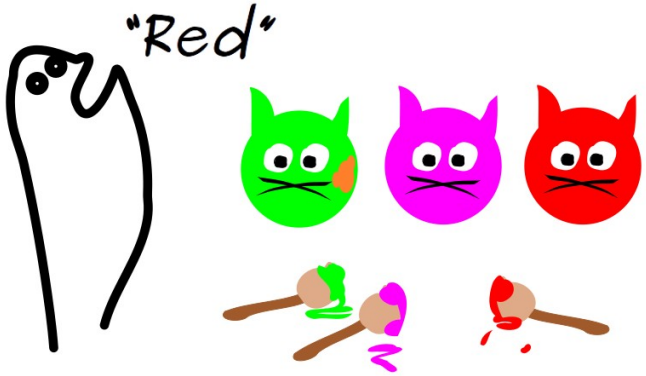
What does distributional semantics model?

Distributional semantics has emerged as a promising model of certain 'conceptual' aspects of linguistic meaning (e.g., Landauer and Dumais 1997; Turney and Pantel 2010; Baroni and Lenci 2010; Lenci 2018) and as an indispensable component of applications in Natural Language Processing (e.g., reference resolution, machine translation, image captioning; especially since Mikolov et al. 2013).

Yet its theoretical status within a general theory of meaning and of language and cognition more generally is not clear (e.g., Lenci 2008; Erk 2010; Boleda and Herbelot 2016; Lenci 2018). In particular, ***it is not clear whether distributional semantics can be understood as an actual model of expression meaning*** *– what Lenci (2008) calls the 'strong' view of distributional semantics – or merely as a model of something that correlates with expression meaning in certain partial ways – the 'weak' view.*

(Westera & Boleda 2019)

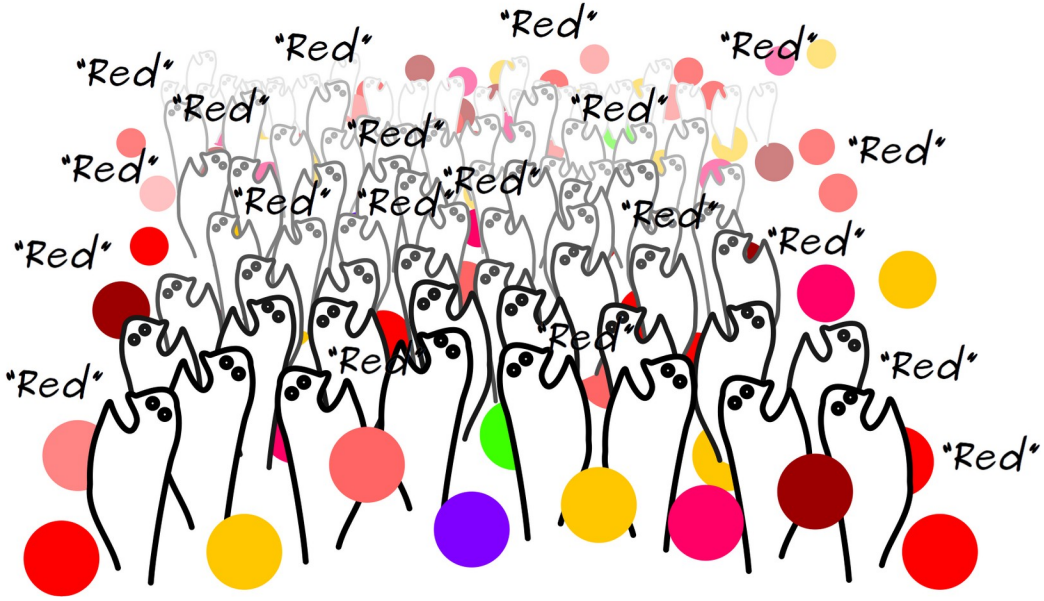
What should a model of expression meaning offer?

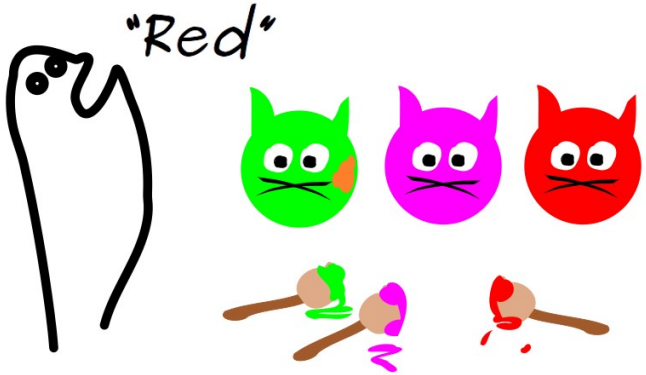


Speaker meaning:
The speaker's communicative intention (varies a lot).

Expression meaning:
Some kind of fixed backbone of all uses.

What should a model of expression meaning offer?

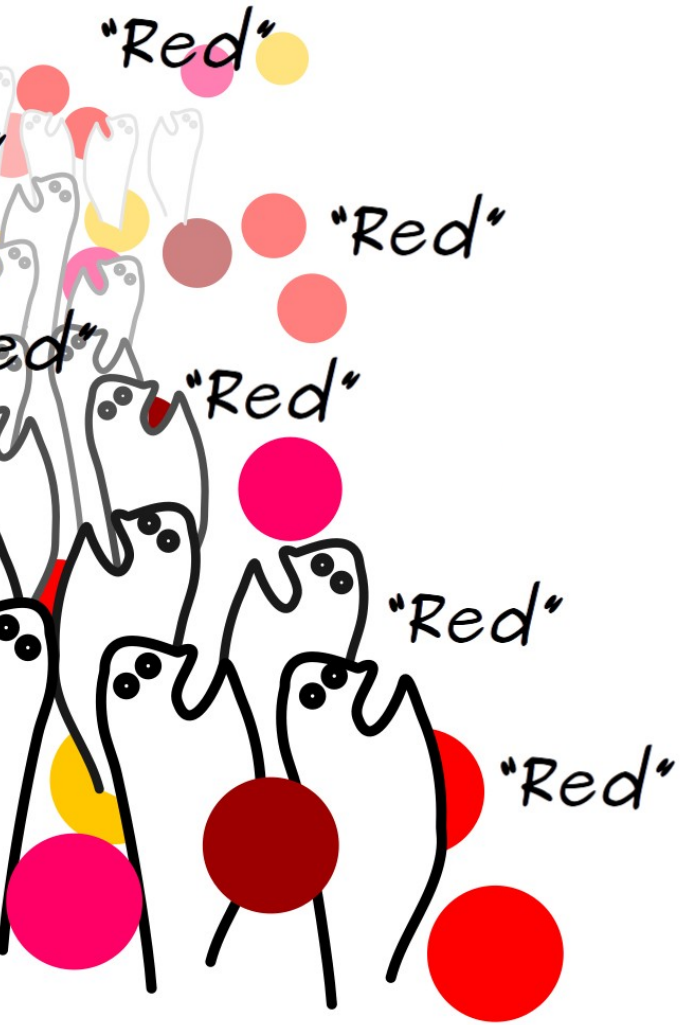




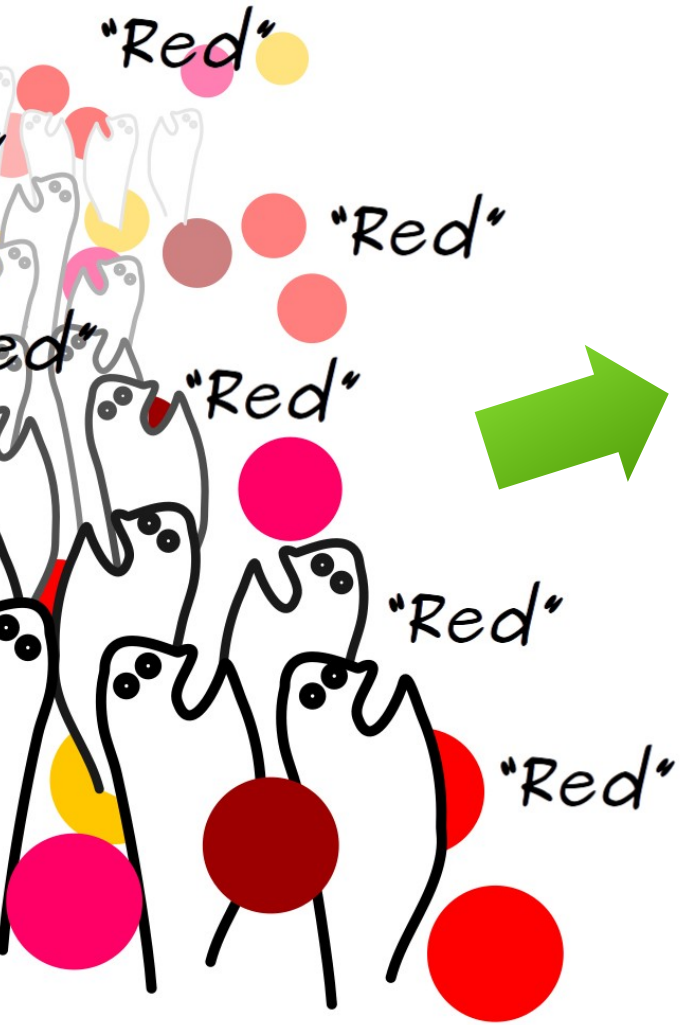
Speaker meaning:
The speaker's communicative intention (varies a lot).

Expression meaning:
Some kind of fixed backbone of all uses.

What should a model of expression meaning offer?

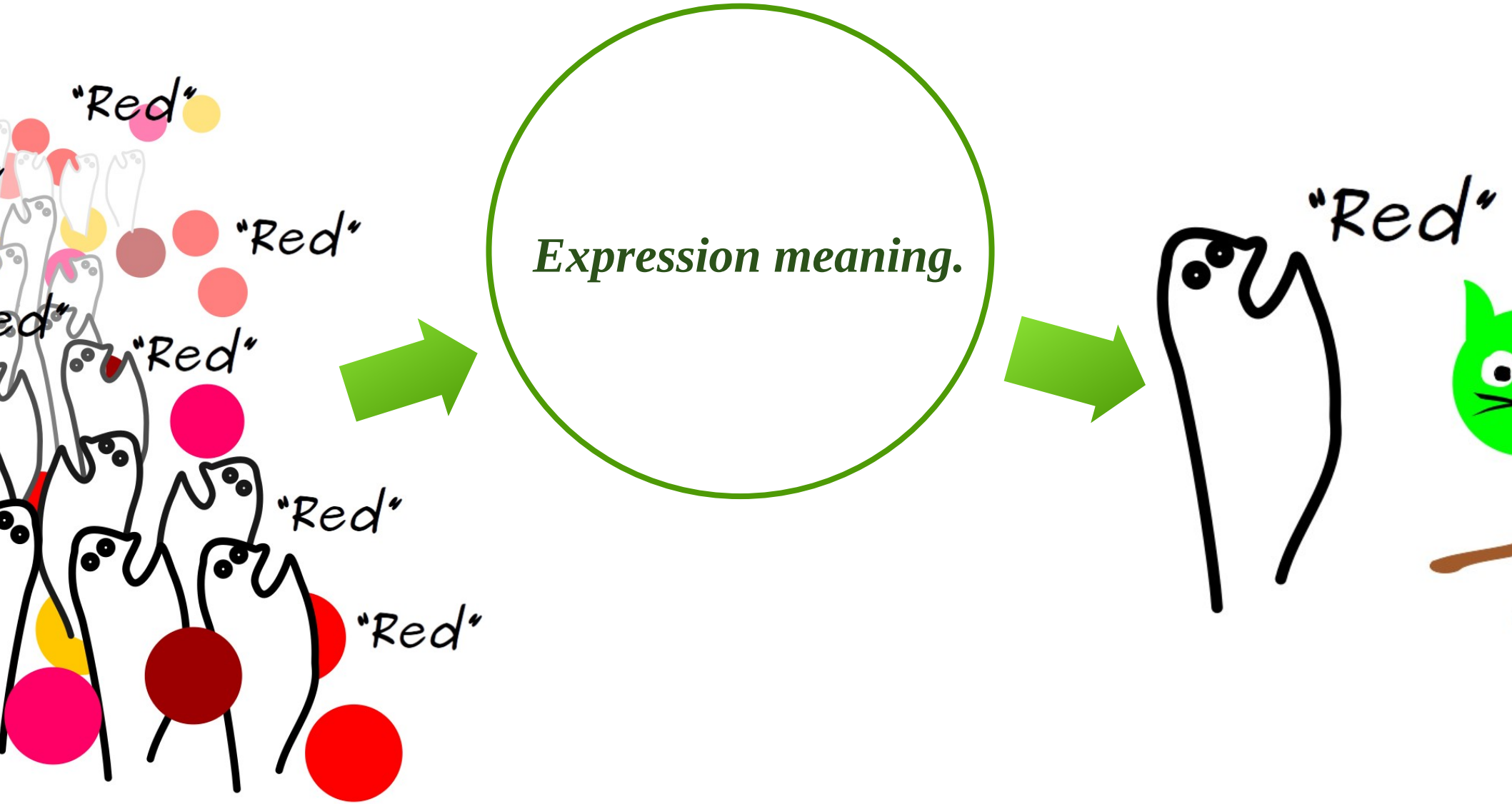


What should a model of expression meaning offer?

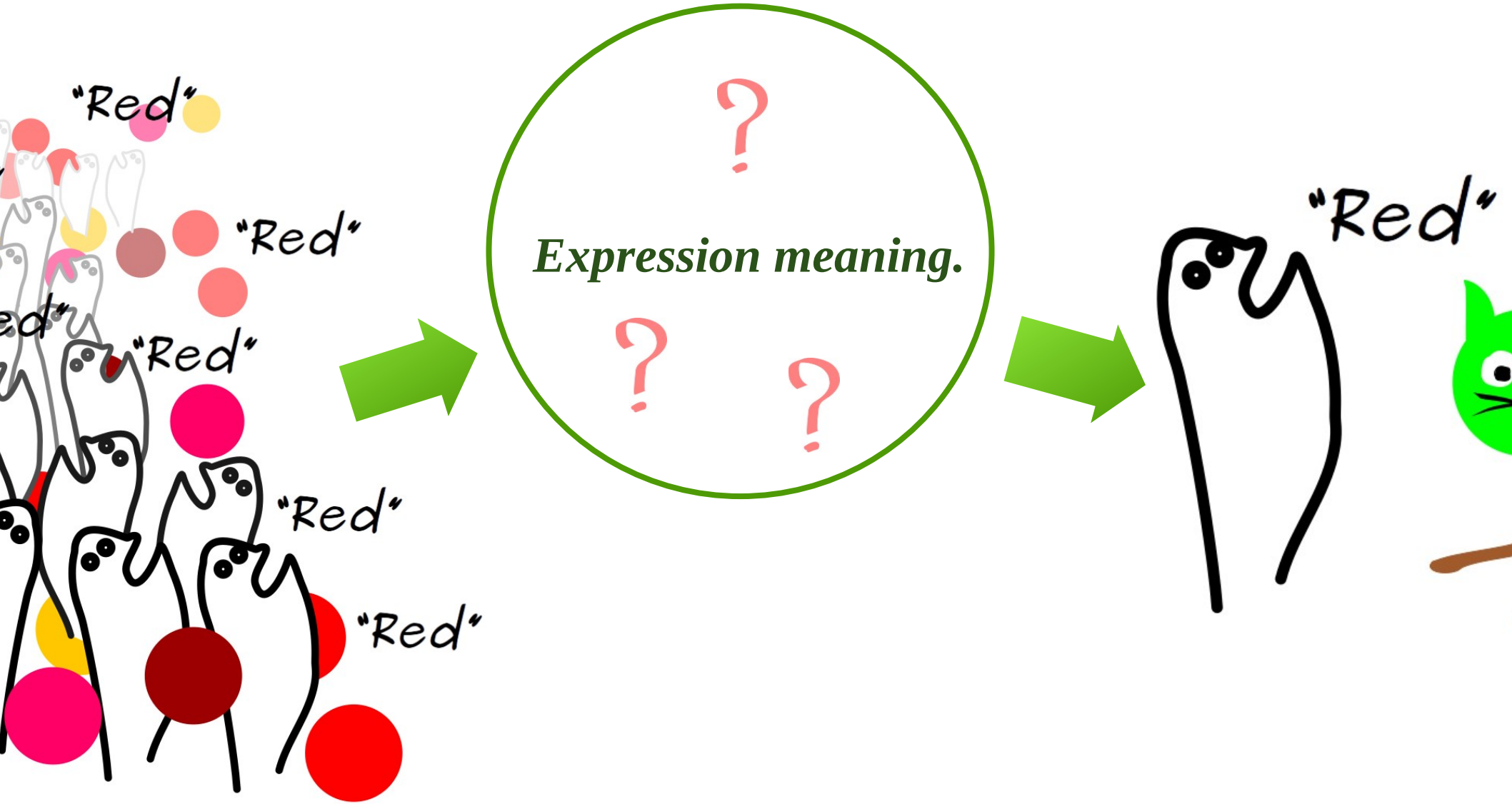


Expression meaning.

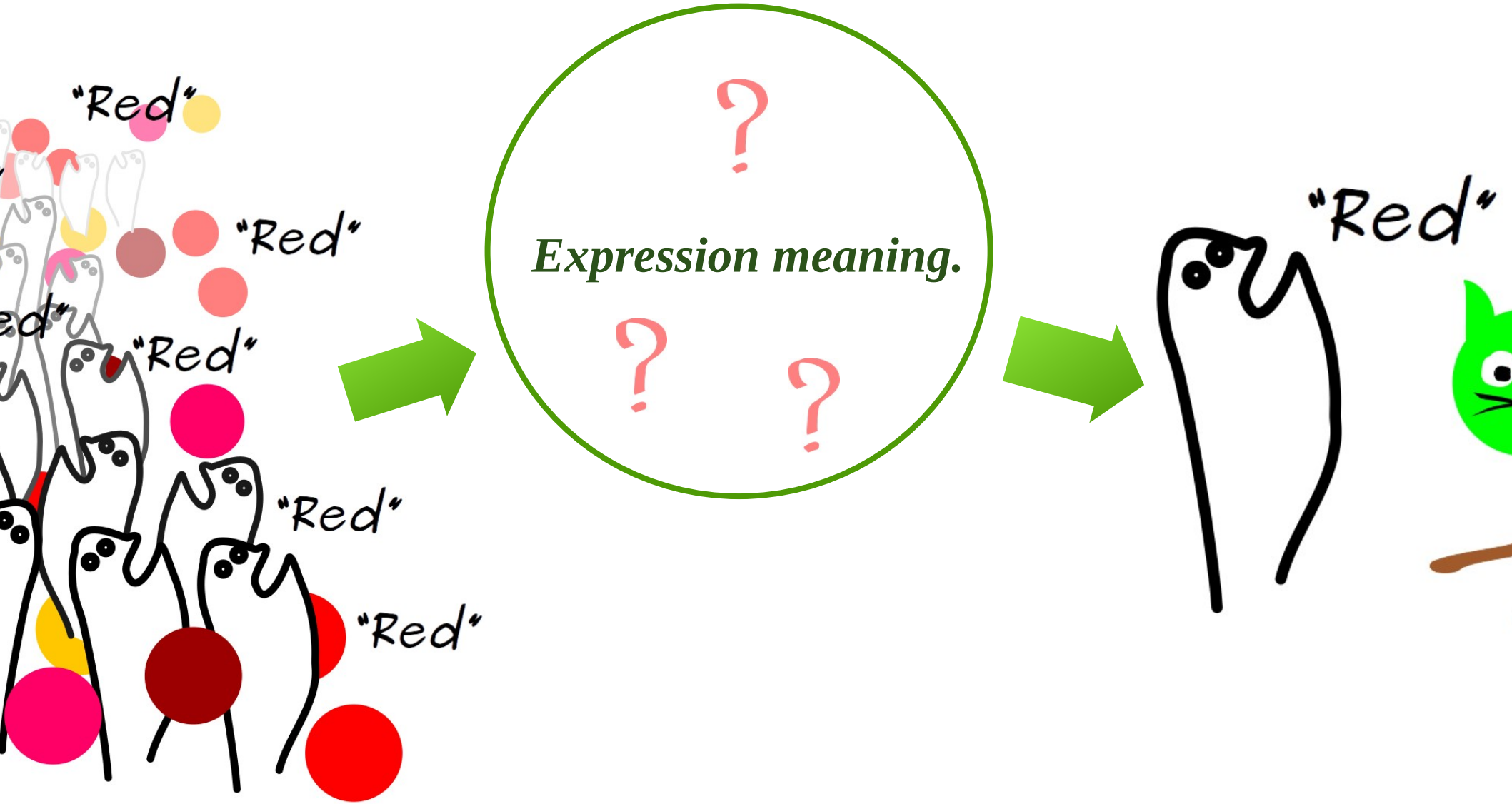
What should a model of expression meaning offer?



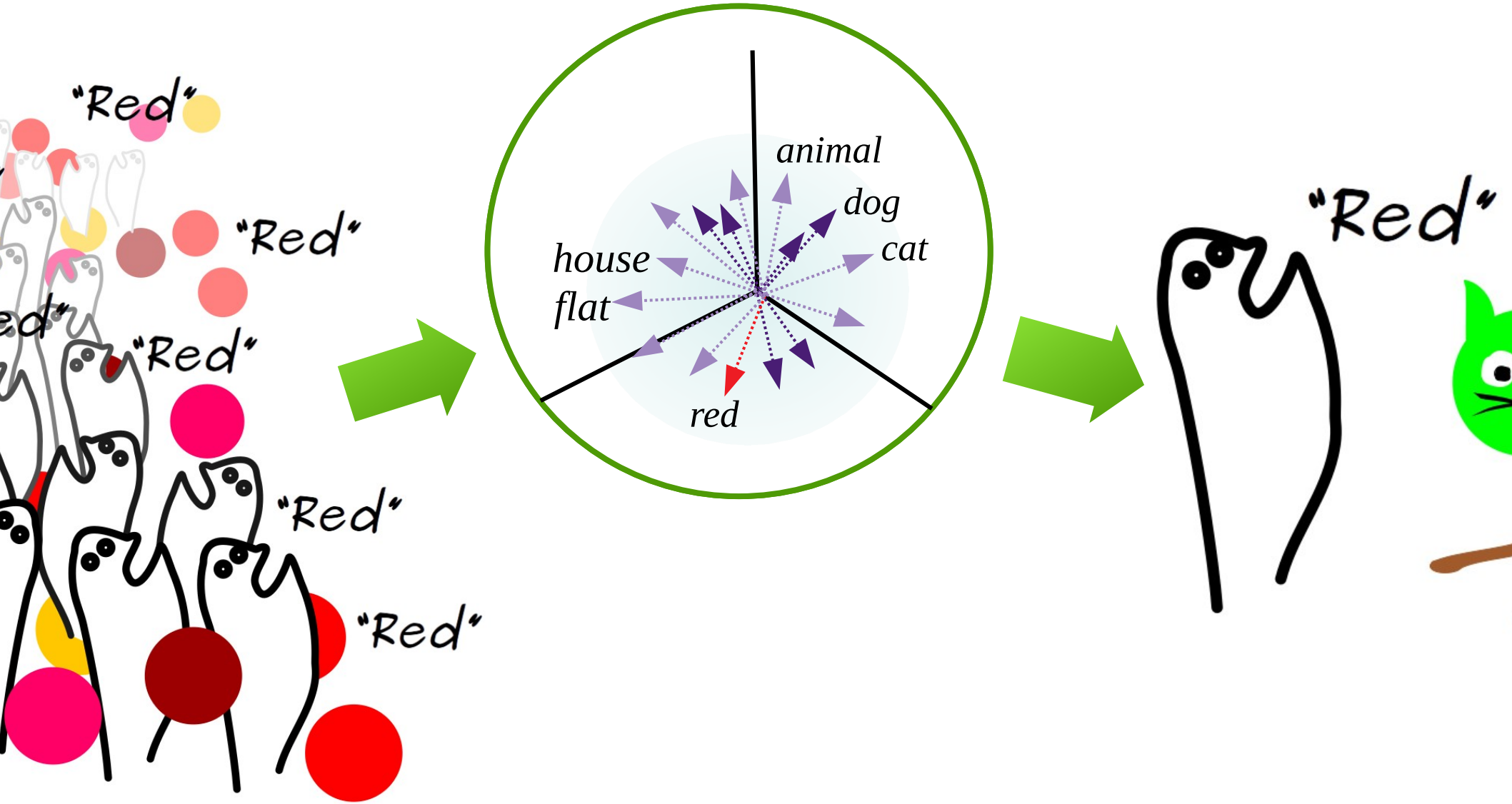
What should a model of expression meaning offer?



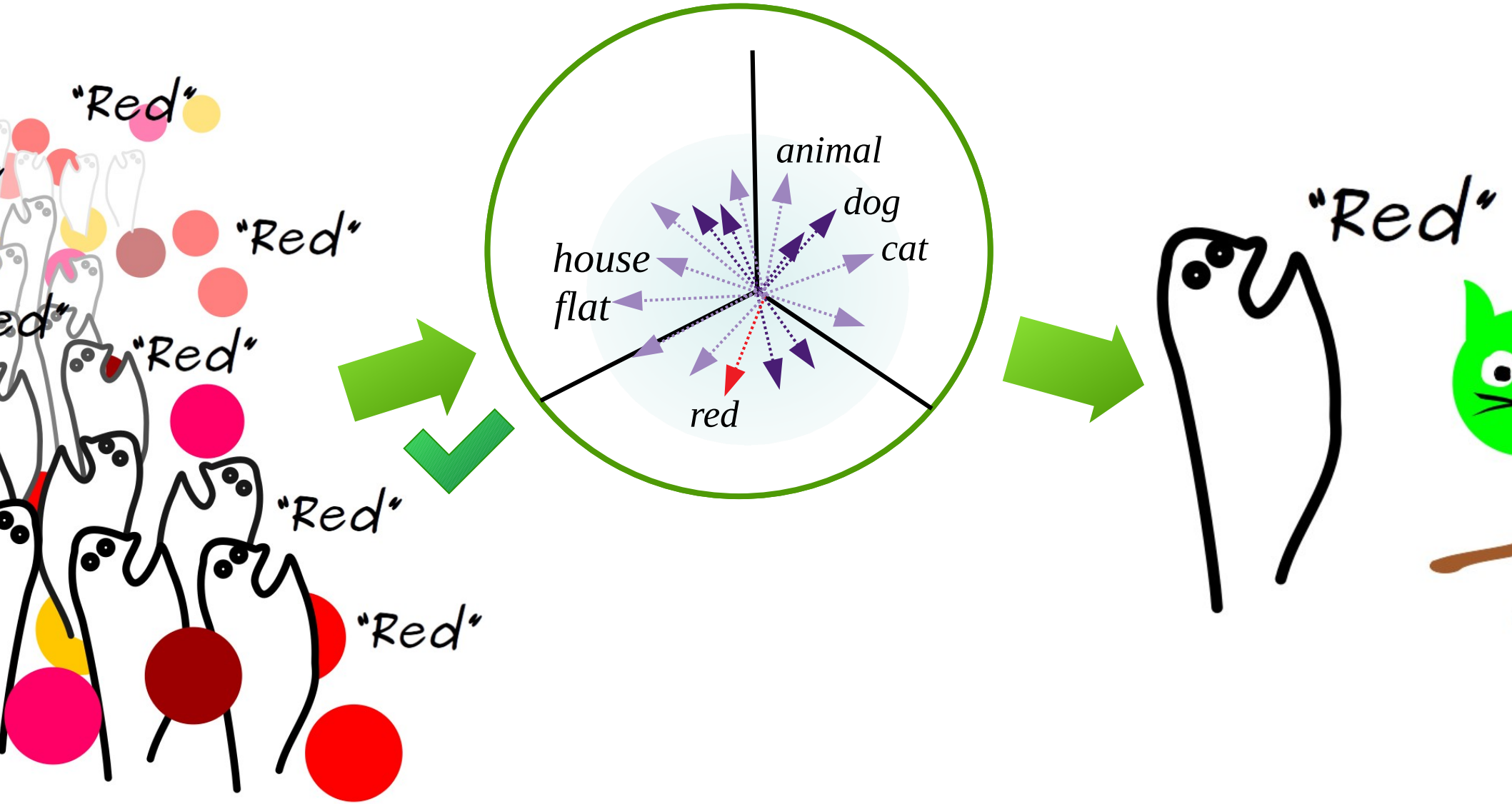
DS as a model of expression meaning



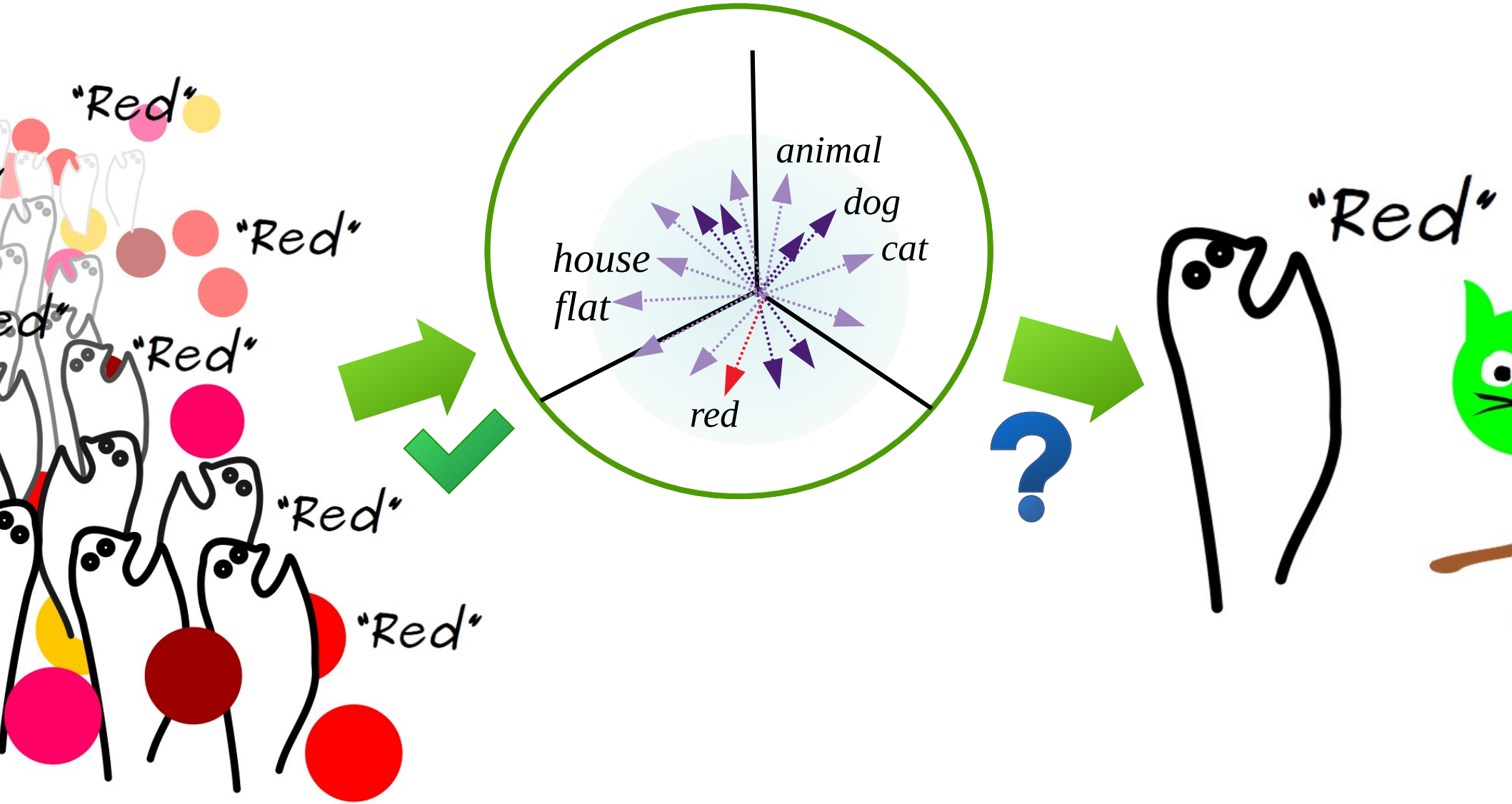
DS as a model of expression meaning



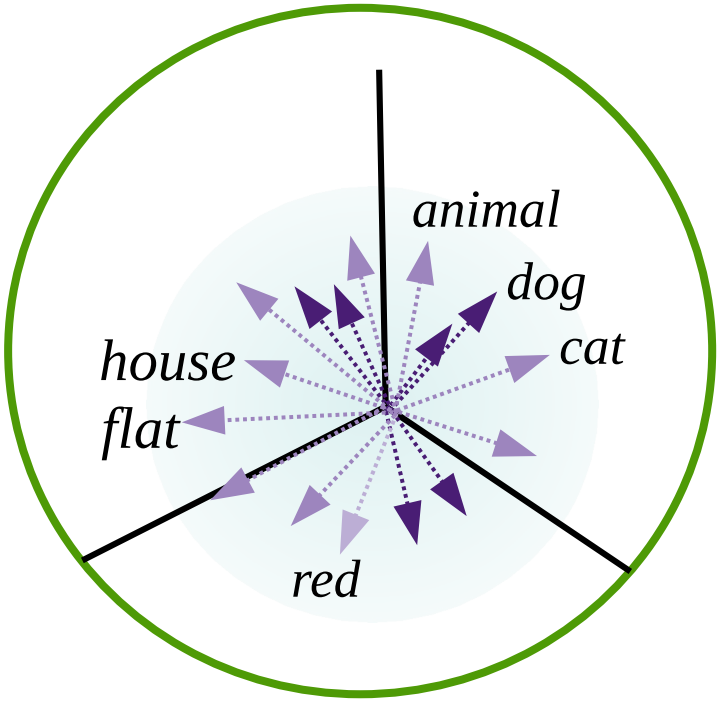
DS as a model of expression meaning



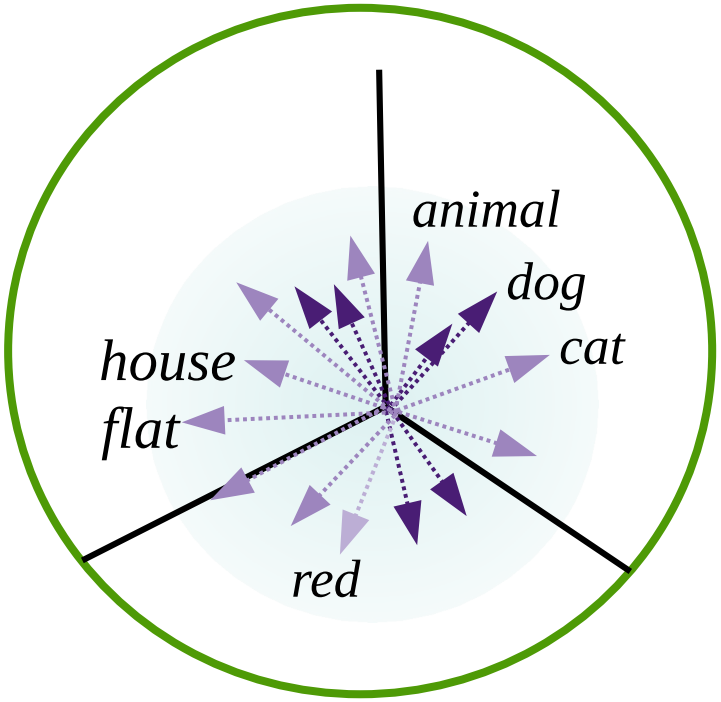
DS as a model of expression meaning



DS as a model of concepts

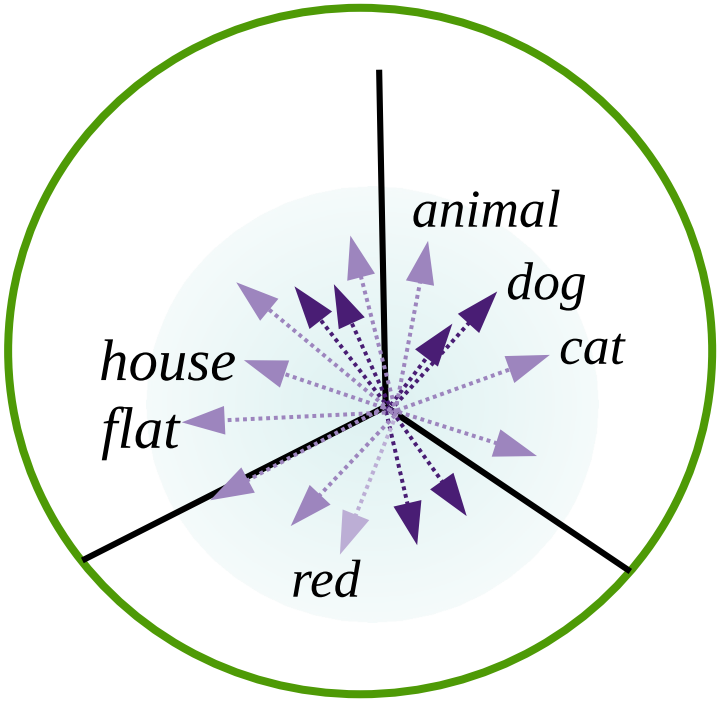


DS as a model of concepts



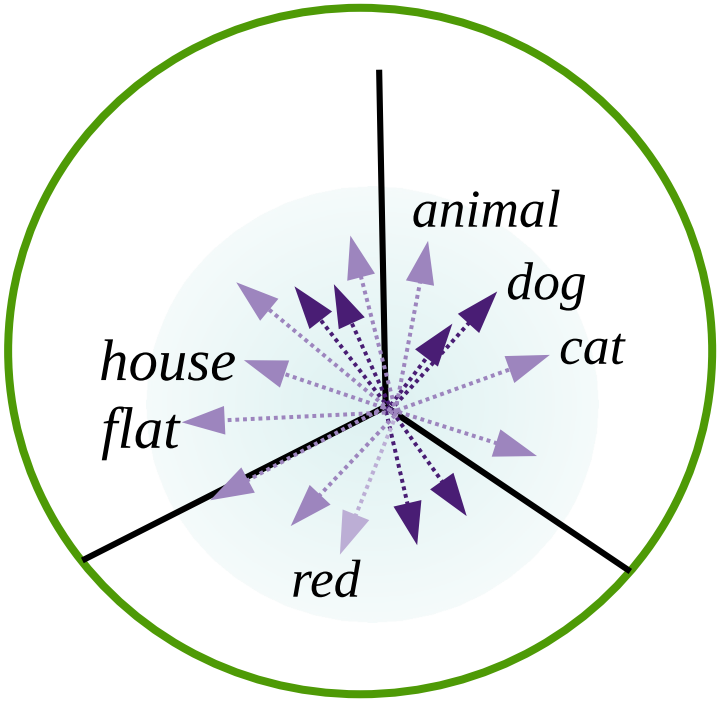
- The vectors of DS are *abstractions over occurrences*.

DS as a model of concepts



- The vectors of DS are *abstractions over occurrences*.
- And so are *concepts* (e.g., Piaget).

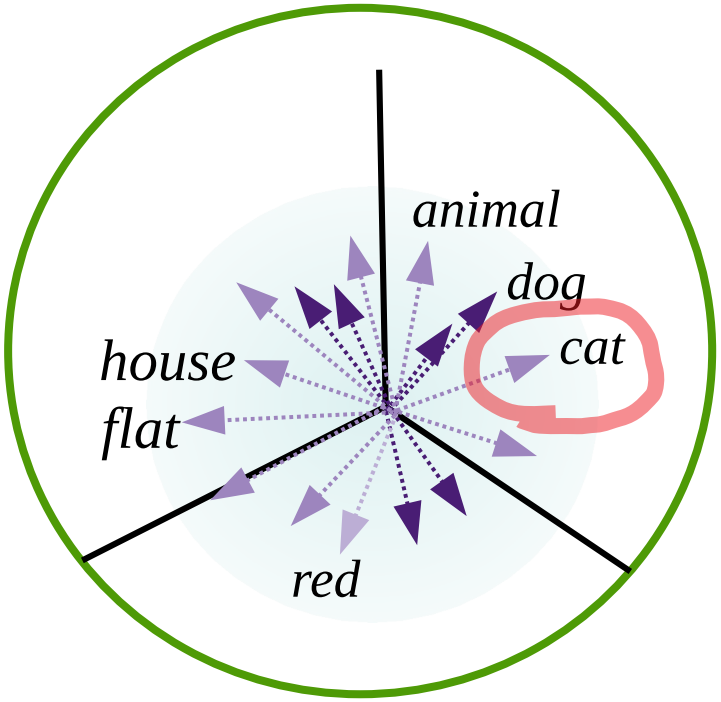
DS as a model of concepts



- The vectors of DS are *abstractions over occurrences*.
- And so are *concepts* (e.g., Piaget).

But what sort of concepts does DS model?

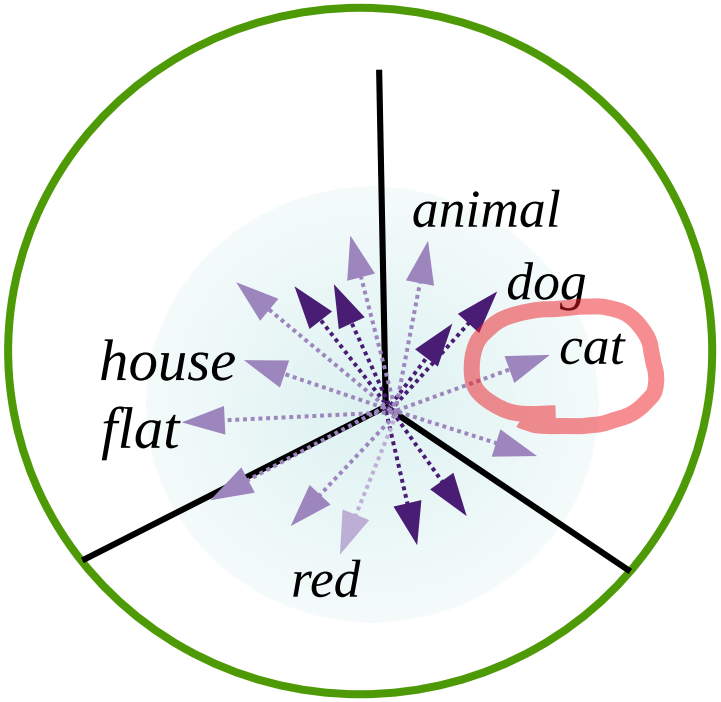
DS as a model of concepts



- The vectors of DS are *abstractions over occurrences*.
- And so are *concepts* (e.g., Piaget).

But what sort of concepts does DS model?

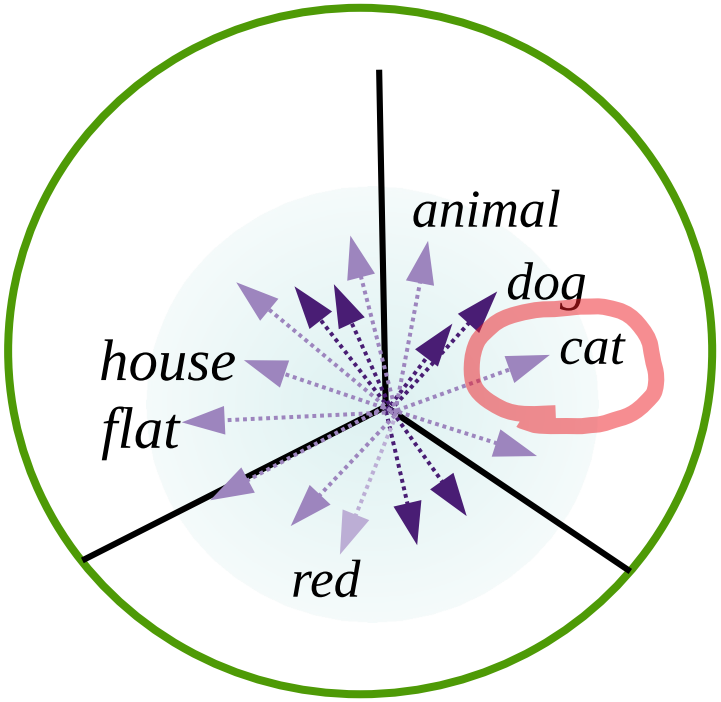
DS as a model of concepts



- The vectors of DS are *abstractions over occurrences*.
- And so are *concepts* (e.g., Piaget).

But what sort of concepts does DS model?

DS as a model of concepts

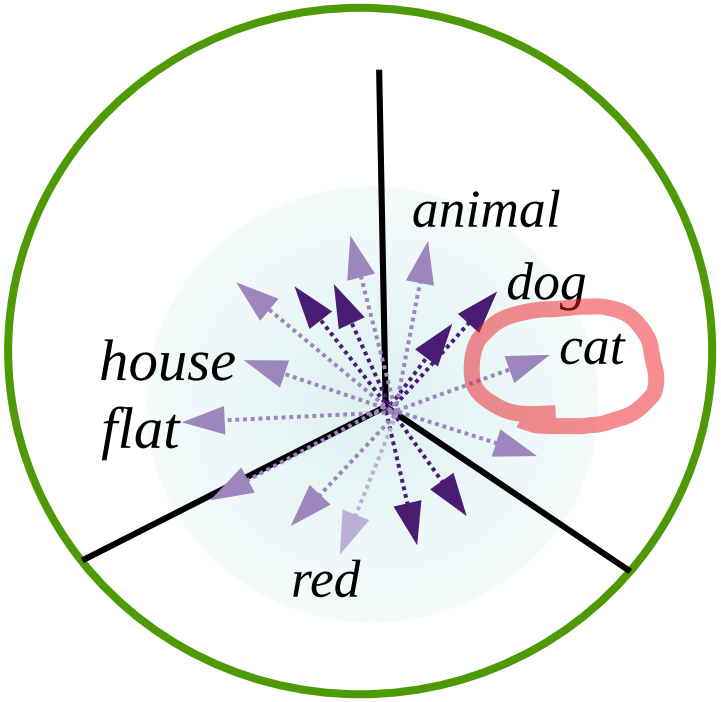


- The vectors of DS are *abstractions over occurrences*.
- And so are *concepts* (e.g., Piaget).

But what sort of concepts does DS model?

“cat”

DS as a model of concepts



- The vectors of DS are *abstractions over occurrences*.
- And so are *concepts* (e.g., Piaget).

But what sort of concepts does DS model?

“cat”

DS as a model of concepts *of expressions*

DS as a model of concepts *of expressions*

- Concepts represent our ability to make sense of the world (e.g., Dummett ‘93).

DS as a model of concepts *of expressions*

- Concepts represent our ability to make sense of the world (e.g., Dummett ‘93).

DS as a model of concepts *of expressions*

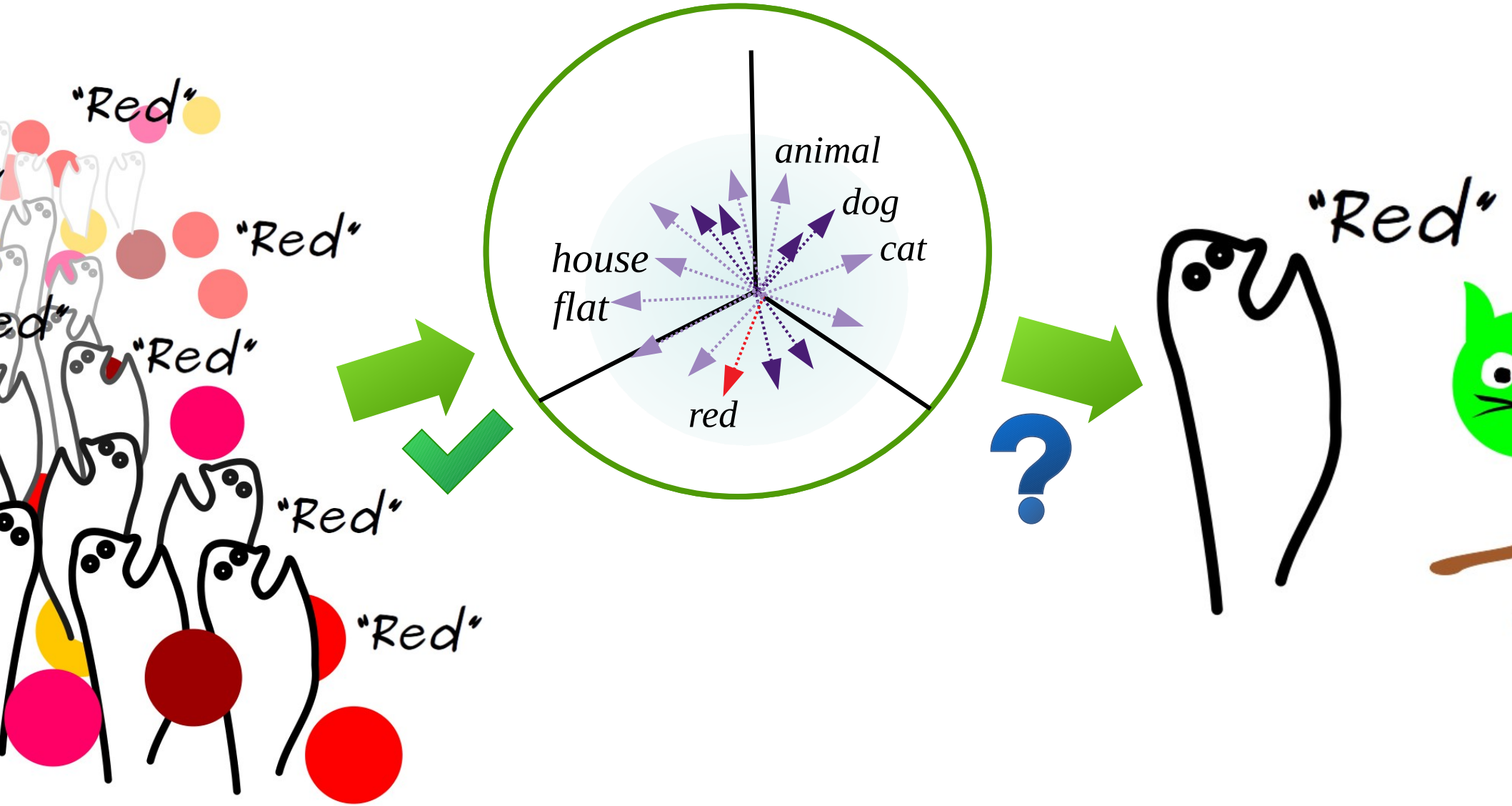
- Concepts represent our ability to make sense of the world (e.g., Dummett ‘93).

DS as a model of concepts *of expressions*

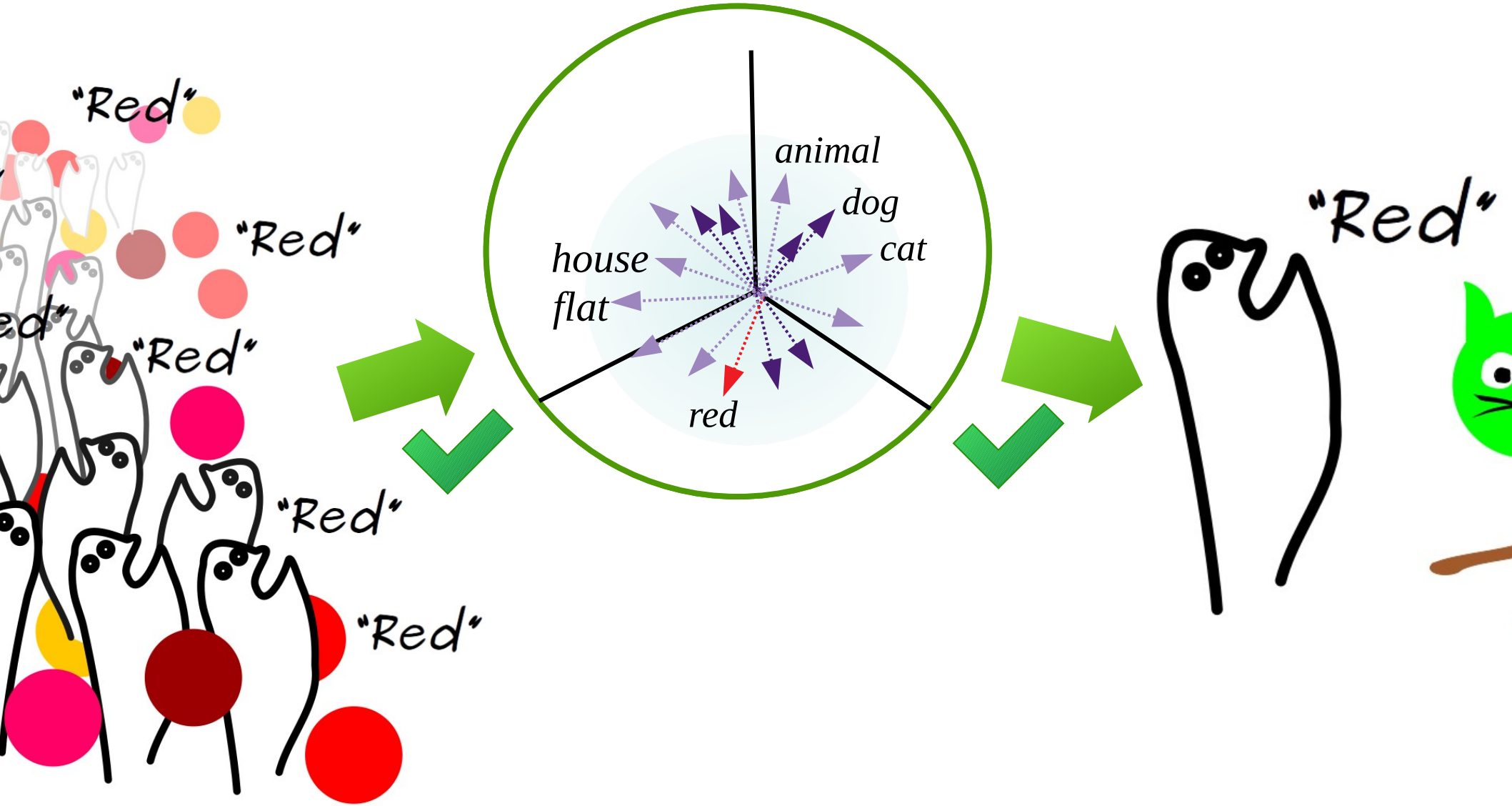
- Concepts represent our ability to make sense of the world (e.g., Dummett ‘93).

- So: concepts *of expressions* represent our ability to make sense of *expressions*.

DS as a model of expression meaning



DS as a model of expression meaning



“But distributional semantics cannot do X!”

- Logical entailment
- Truth conditions
- Denotations/reference
- ‘Semantic intuitions’
- Compositionality

“But distributional semantics cannot do X!”

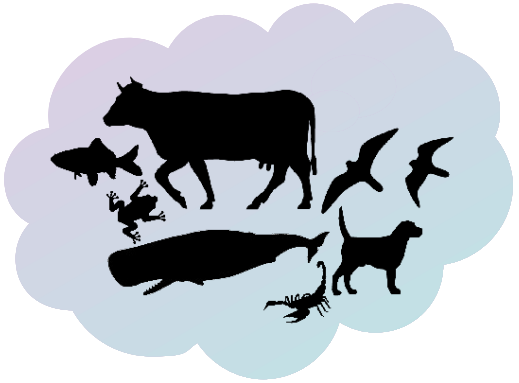
- Logical entailment
- Truth conditions
- Denotations/reference
- ‘Semantic intuitions’
- Compositionality

But are these even things that a theory of expression meaning should do??

Re: Entailment

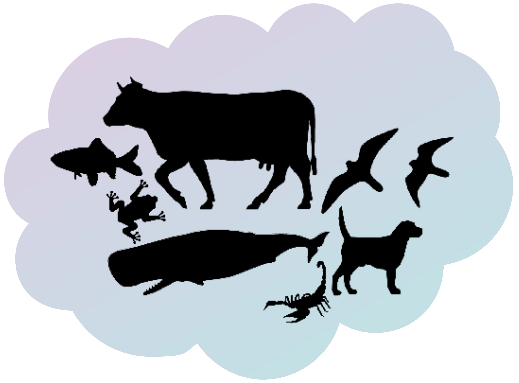
Re: Entailment

“cat”



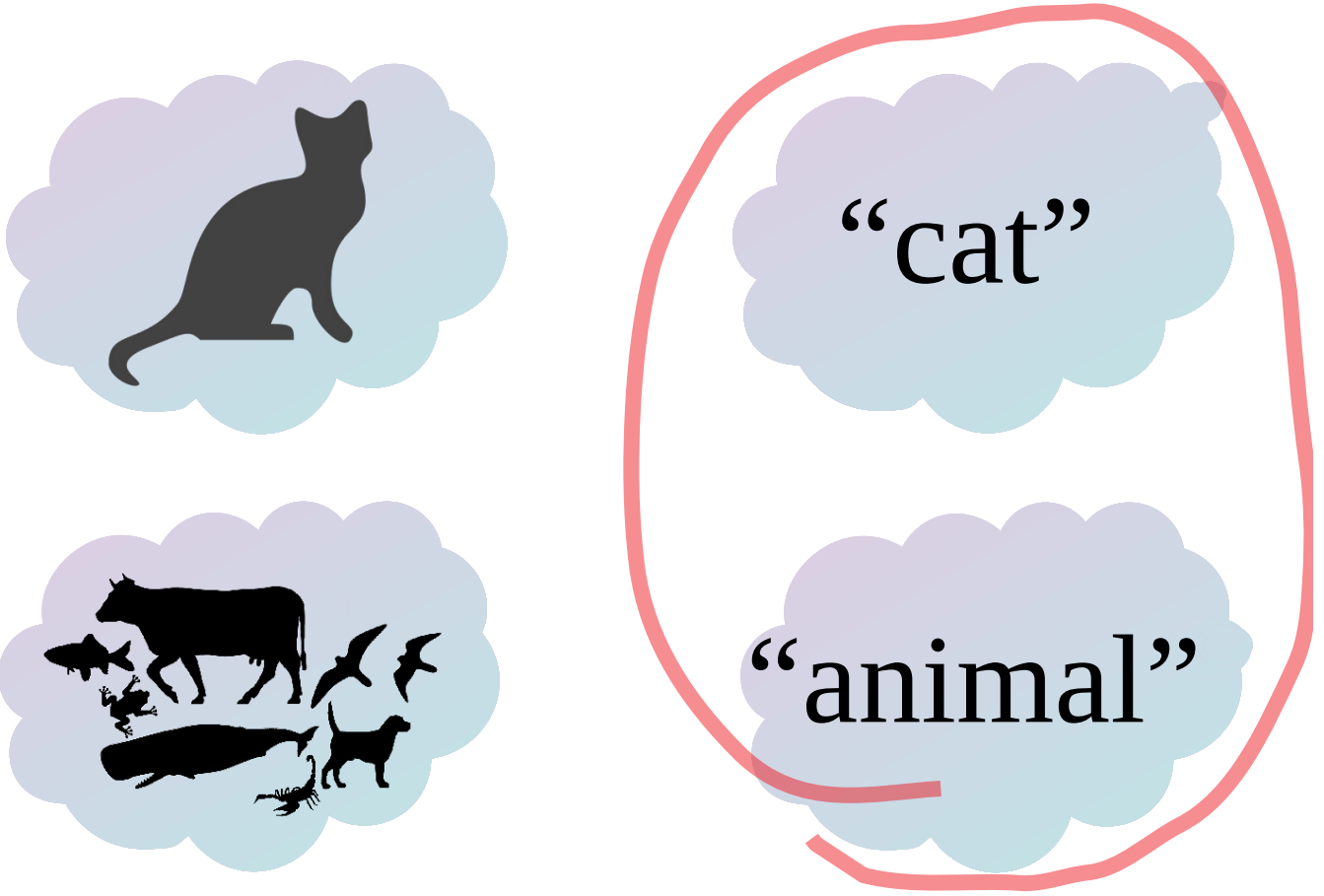
Re: Entailment

“cat”

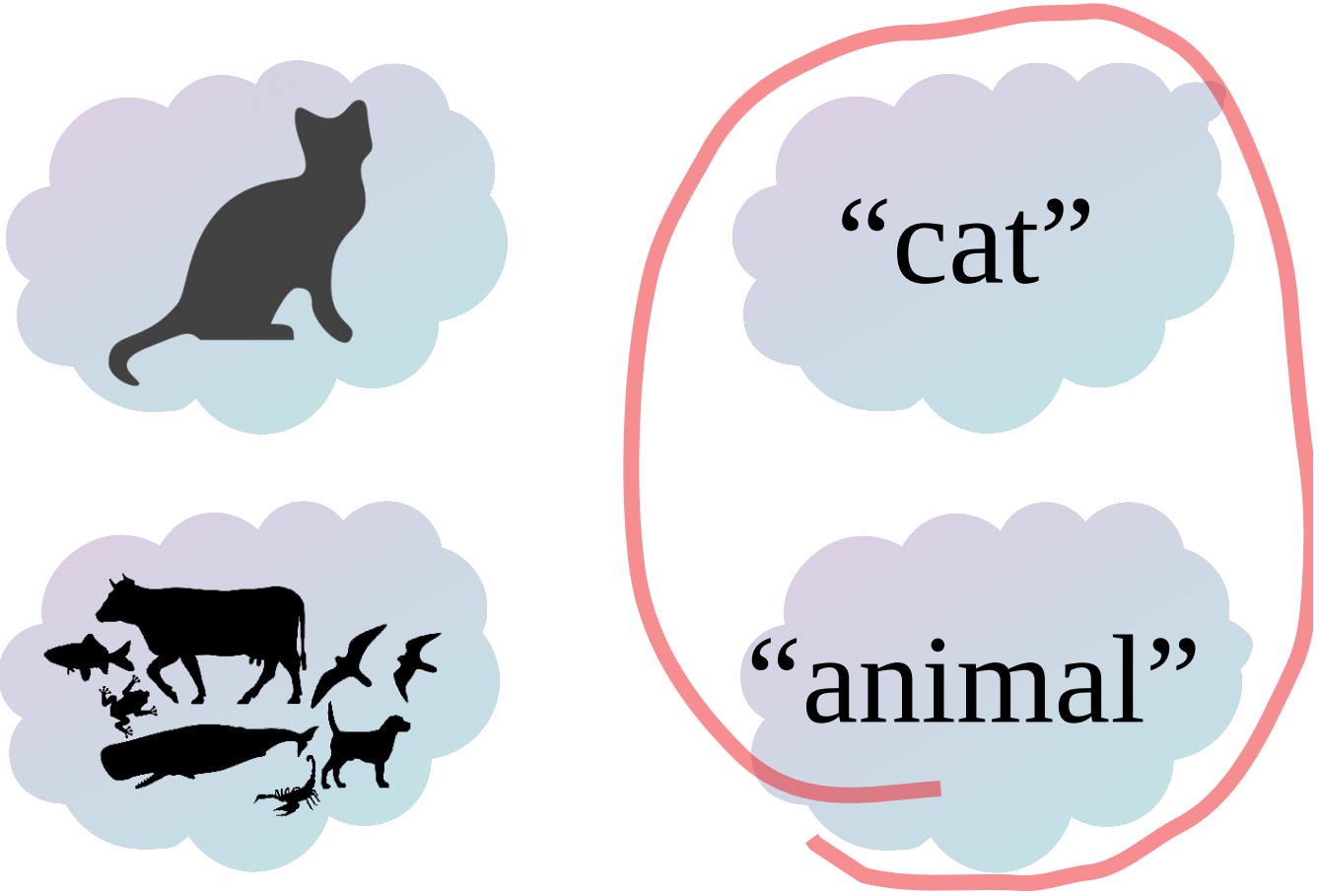


“animal”

Re: Entailment

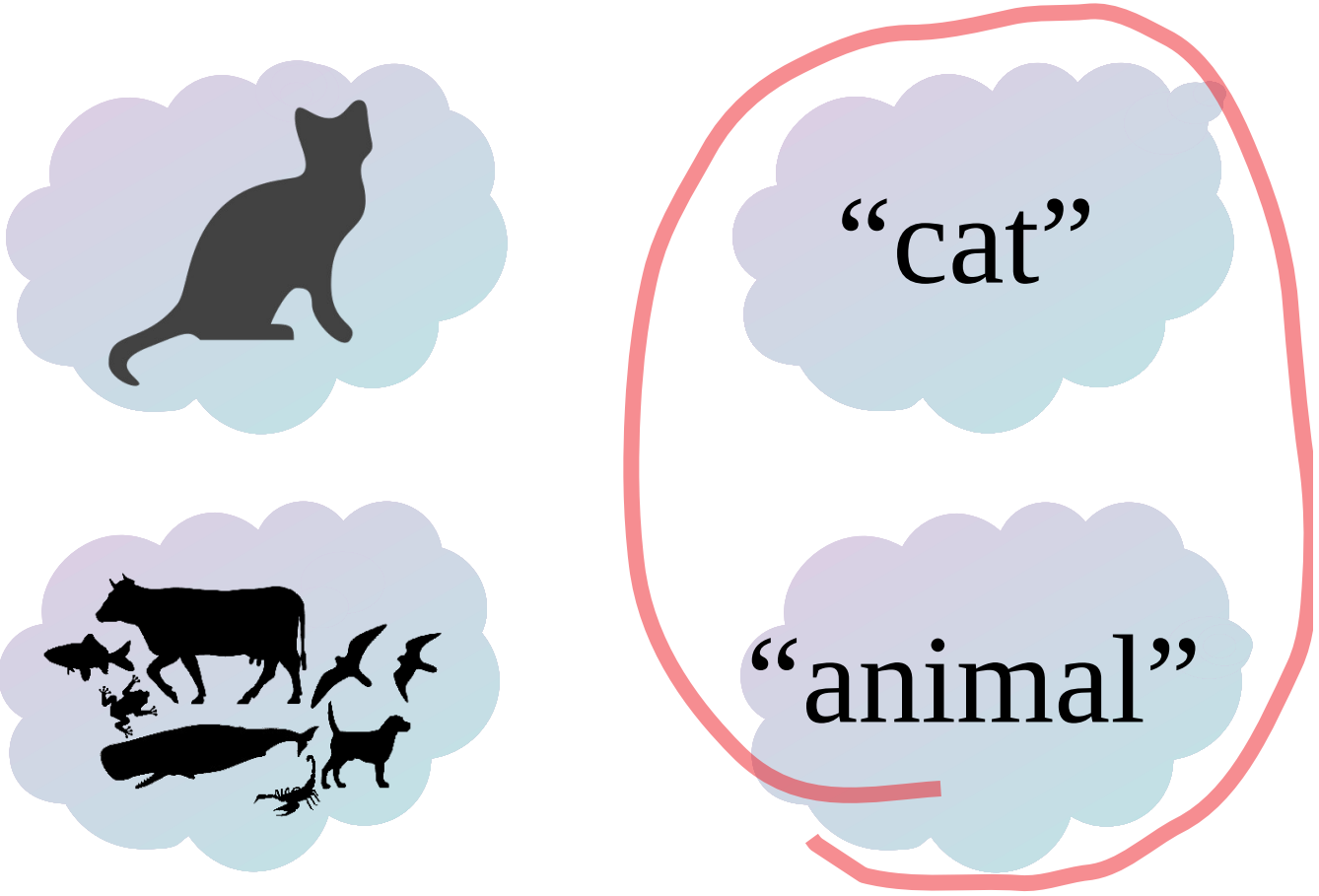


Re: Entailment



Entailment is not a job for expression meaning!

Re: Entailment



Entailment is not a job for expression meaning!

Re: Truth conditions

Re: Truth conditions

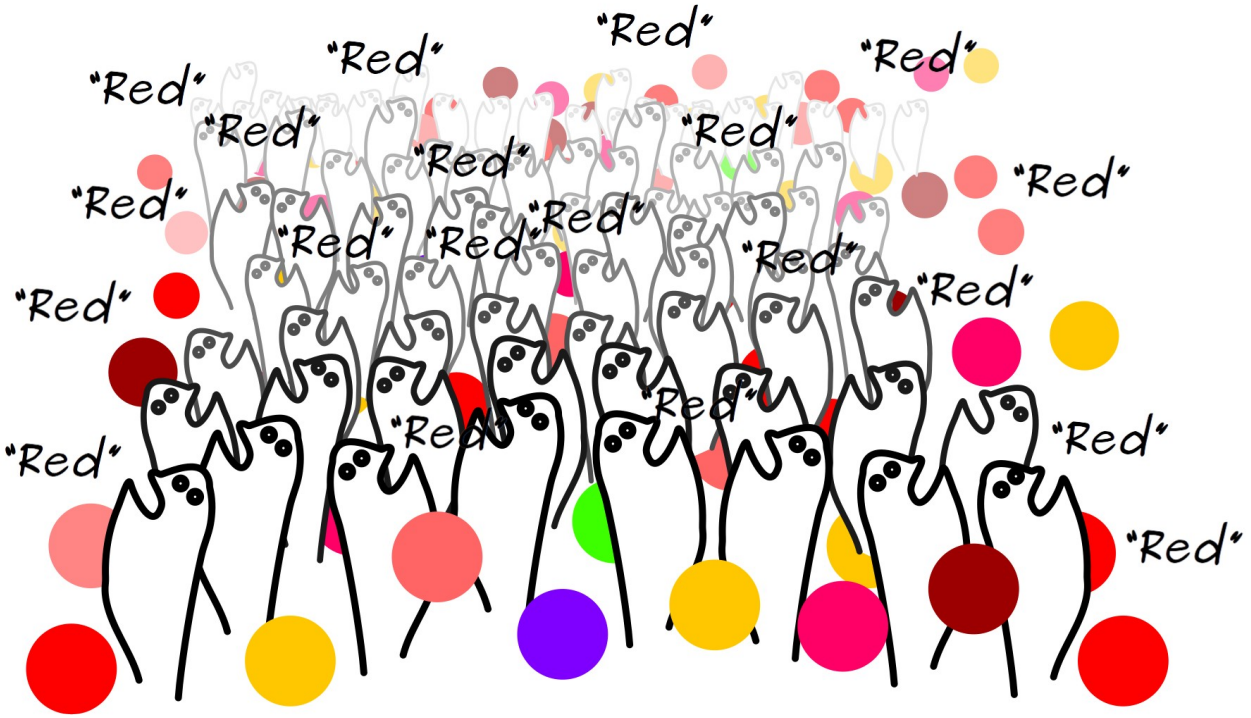
The set of uses of a given expression resists logical definition.

(e.g., Wittgenstein; Rosch; Fodor; Kilgarriff)

Re: Truth conditions

The set of uses of a given expression resists logical definition.

(e.g., Wittgenstein; Rosch; Fodor; Kilgarriff)



Re: Truth conditions

In philosophy, the classical theory has been subjected to a number of criticisms but perhaps the most fundamental is that attempts to specify definitions for concepts have a poor track record. Quite simply, there are too few examples of successful definitional analyses, and certainly none that are uncontroversial.

It could be that the problem is that definitions are hard to come by. But another possibility—one that many philosophers are now taking seriously—is that our concepts lack definitional structure.

(Margolis & Lawrence 2022)

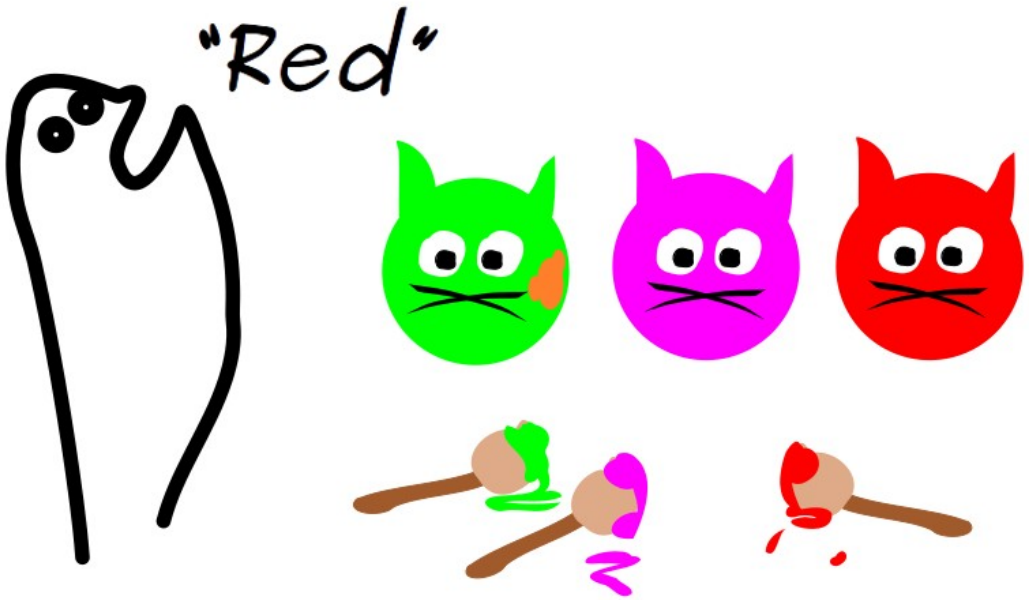
Re: Denotation

Re: Denotation

Words don’t refer, *speakers* do.

(e.g., Strawson, Bach)

Re: Denotation



Words don't refer, *speakers* do.

(e.g., Strawson, Bach)

Re: Denotation

Every one knows that the sentence, "The table is covered with books", is significant, and every one knows what it means. But if I ask, "What object is that sentence about?" I am asking an absurd question – a question which cannot be asked about the sentence, but only about some use of the sentence: and in this case the sentence hasn't been used, it has only been taken as an example.

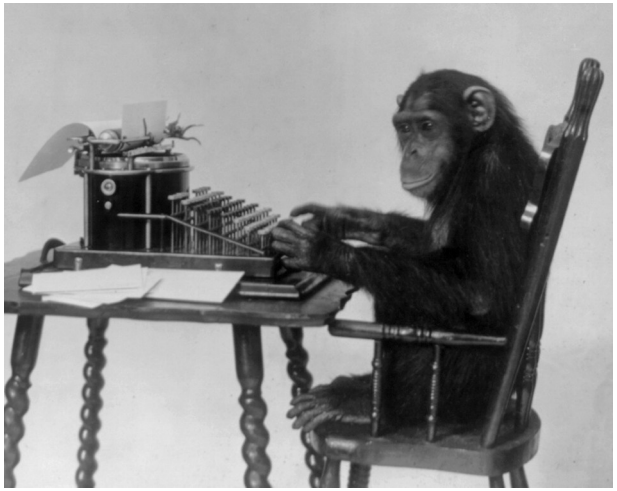
(Strawson 1950)

Re: Denotation

Re: Denotation

The red cat chased a mouse.

Re: Denotation



The red cat chased a mouse.

Re: Semantic intuitions

Re: Semantic intuitions

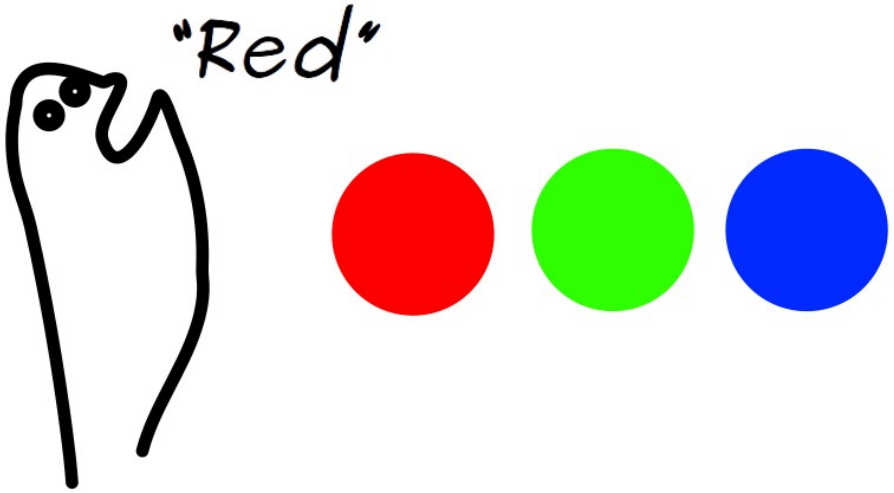
Supposed intuitions about expression meaning
in fact reflect stereotypical speaker meaning.

(e.g., Grice, Bach, Schwarz)

Re: Semantic intuitions

Supposed intuitions about expression meaning
in fact reflect stereotypical speaker meaning.

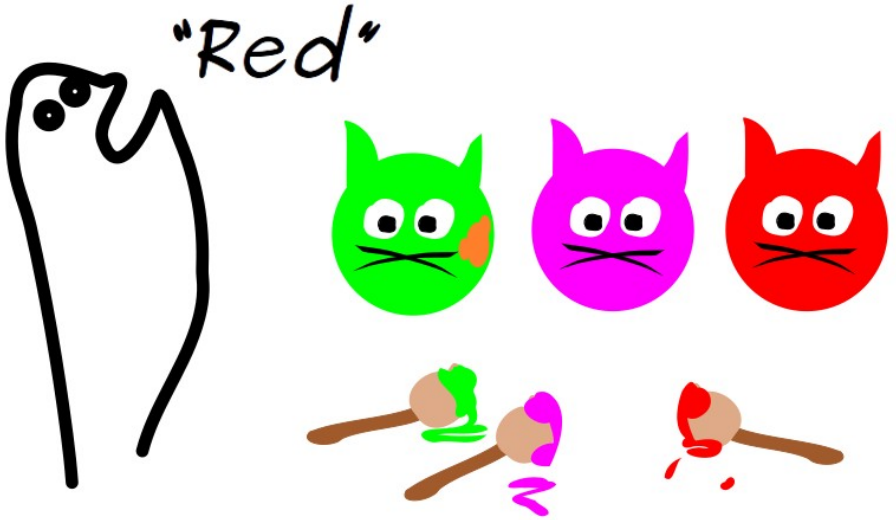
(e.g., Grice, Bach, Schwarz)



Re: Semantic intuitions

Supposed intuitions about expression meaning
in fact reflect stereotypical speaker meaning.

(e.g., Grice, Bach, Schwarz)



Re: Semantic intuitions

Supposed intuitions about expression meaning
in fact reflect stereotypical speaker meaning.

(e.g., Grice, Bach, Schwarz)

Re: Compositionality

Re: Compositionality

- What was the argument for compositionality again?

Re: Compositionality

- What was the argument for compositionality again?

The 'productivity' argument for compositionality does not tell us *where* to put it: expression or speaker meaning.

“But DS cannot do X!”

- Logical entailment
- Truth conditions
- Denotations/reference
- ‘Semantic intuitions’
- Compositionality

But are these even things that a theory of expression meaning should do?

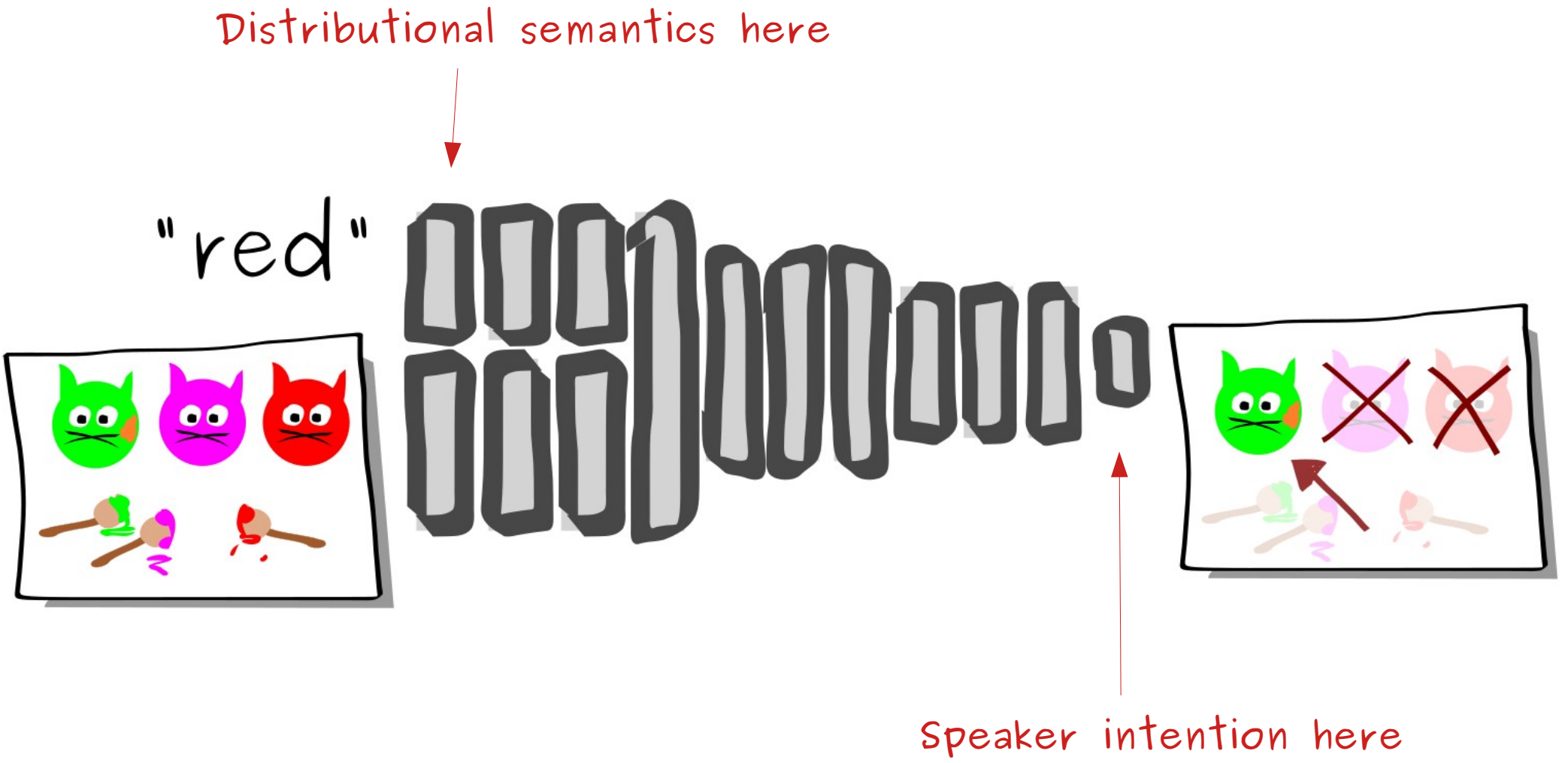
“But DS cannot do X!”

- Logical entailment
- Truth conditions
- Denotations/reference
- ‘Semantic intuitions’
- Compositionality

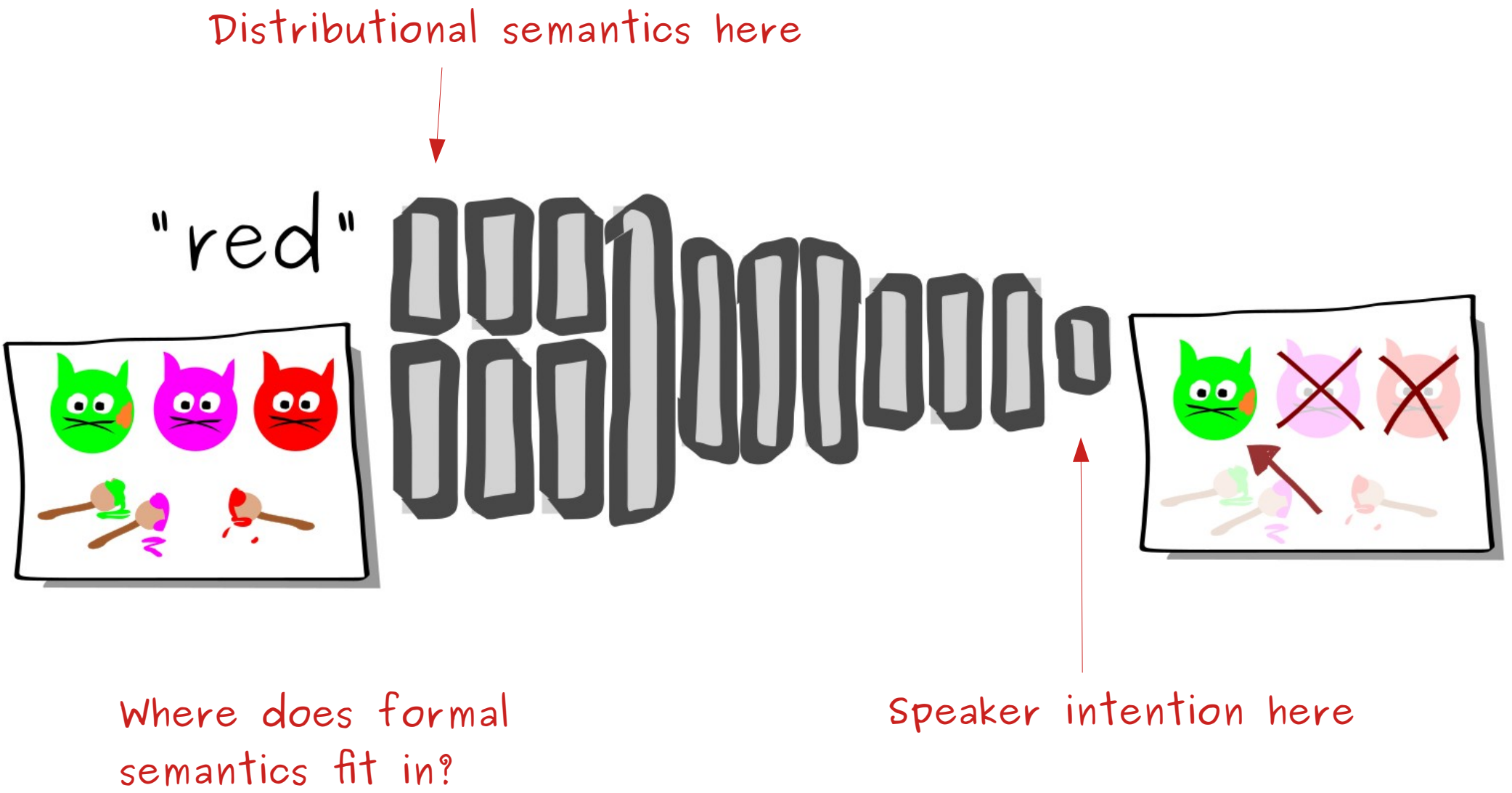
But are these even things that a theory of expression meaning should do?

…let alone on its own??

Recall: visual information retrieval



Recall: visual information retrieval



A perspective on formal semantics

A perspective on formal semantics

- Formal semantics is not necessarily *wrong*...

A perspective on formal semantics

- Formal semantics is not necessarily *wrong*...
- But it isn’t a theory of what we thought it was a theory of.

A perspective on formal semantics

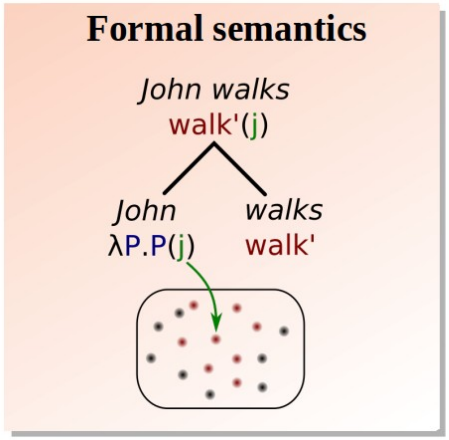
- Formal semantics is not necessarily *wrong*...
- But it isn’t a theory of what we thought it was a theory of.
- Entailment, compositionality, truth conditions, etc., reside at the level of speaker meaning, not expression meaning.

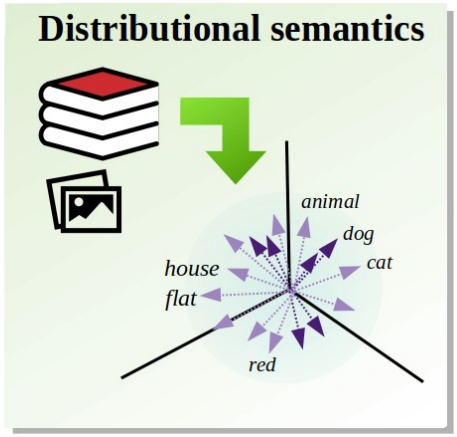
A perspective on formal semantics

- Formal semantics is not necessarily *wrong*...
- But it isn’t a theory of what we thought it was a theory of.
- Entailment, compositionality, truth conditions, etc., reside at the level of speaker meaning, not expression meaning.
- Distributional and formal semantics are not contenders; they model different notions of meaning.

A perspective on formal semantics

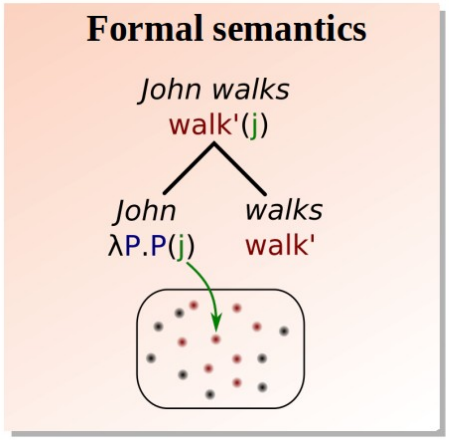
- Formal semantics is not necessarily *wrong*...
- But it isn't a theory of what we thought it was a theory of.
- Entailment, compositionality, truth conditions, etc., reside at the level of speaker meaning, not expression meaning.
- Distributional and formal semantics are not contenders; they model different notions of meaning.

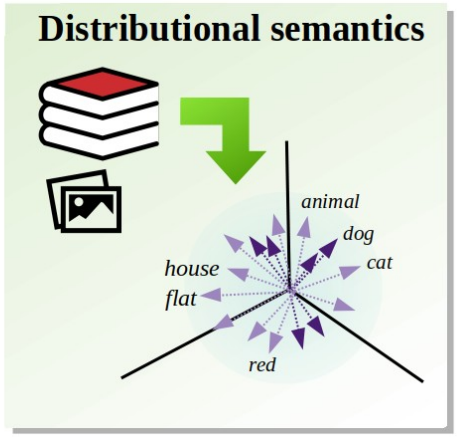




A perspective on formal semantics

- Formal semantics is not necessarily *wrong*...
- But it isn't a theory of what we thought it was a theory of.
- Entailment, compositionality, truth conditions, etc., reside at the level of speaker meaning, not expression meaning.
- Distributional and formal semantics are not contenders; they model different notions of meaning.

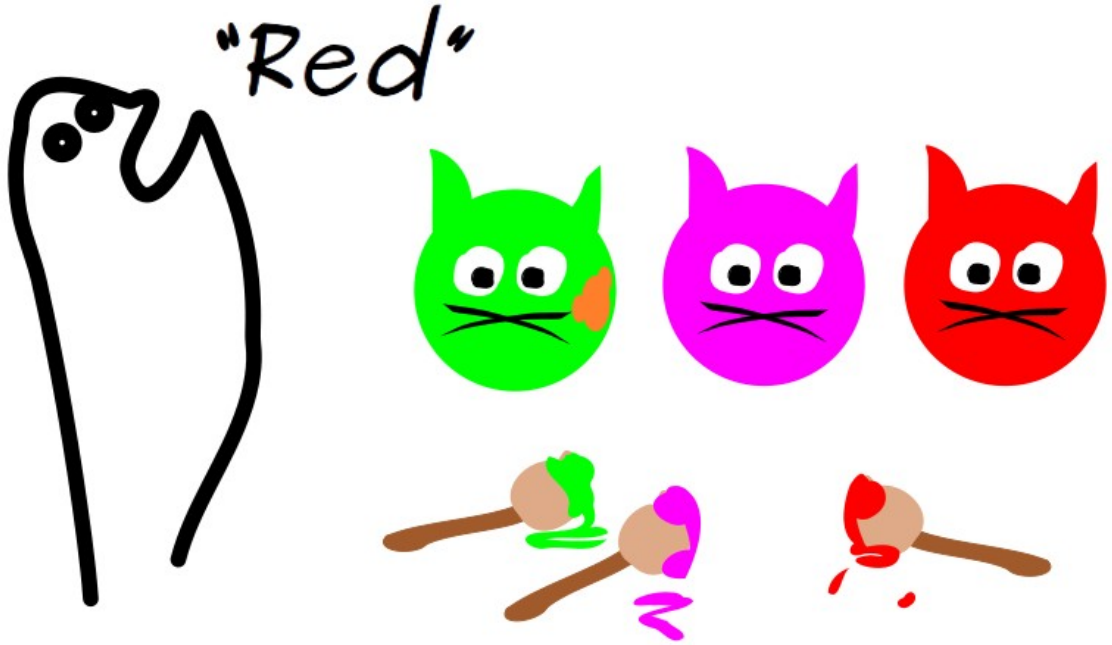




Chomsky?

Conclusions

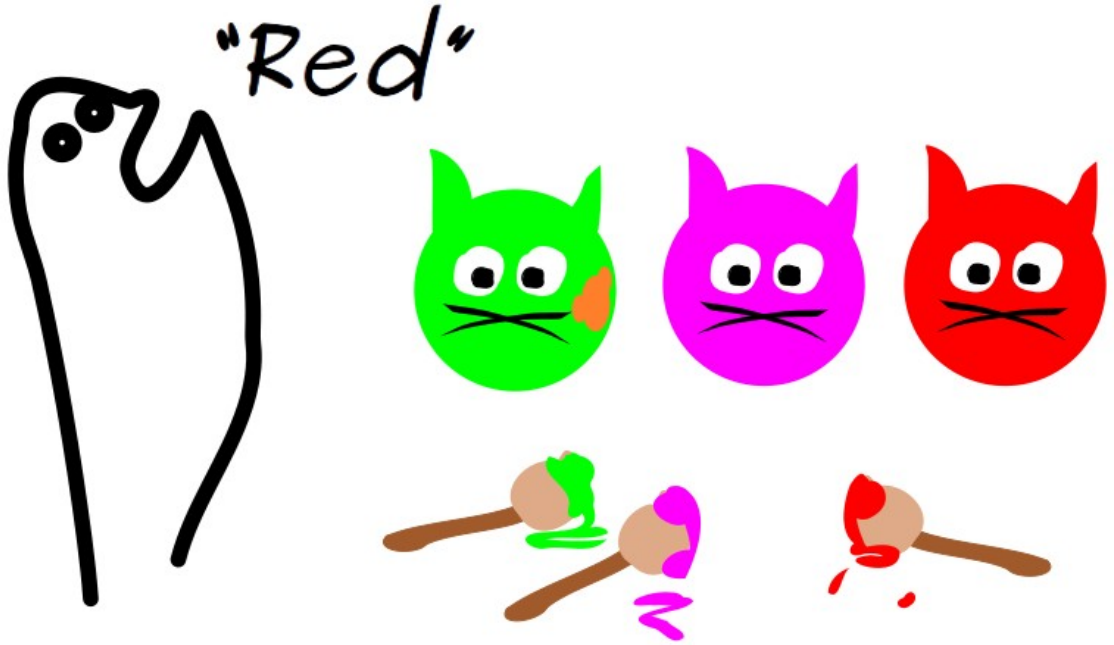
Revising the standard picture



Pragmatics: **Speaker meaning**:
The speaker's communicative intention (varies a lot).

FORMAL Semantics: **Expression meaning**:
Some kind of fixed backbone of all uses.

Revising the standard picture



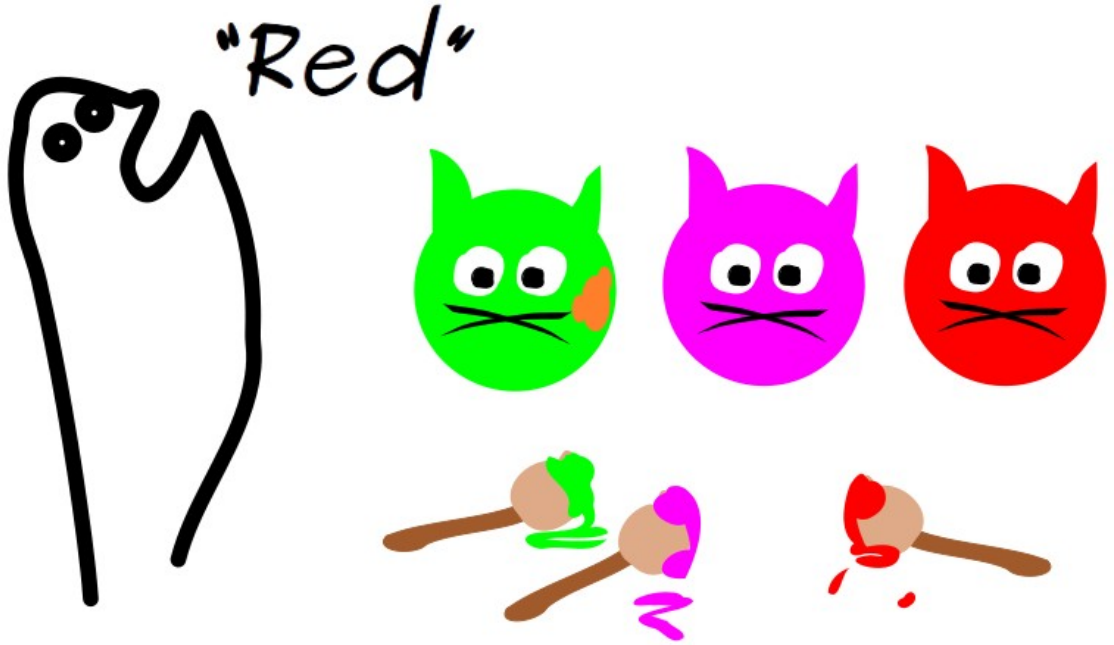
Pragmatics: +formal sem **Speaker meaning**:

The speaker's communicative intention (varies a lot).

FORMAL Semantics: **Expression meaning**:

Some kind of fixed backbone of all uses.

Revising the standard picture



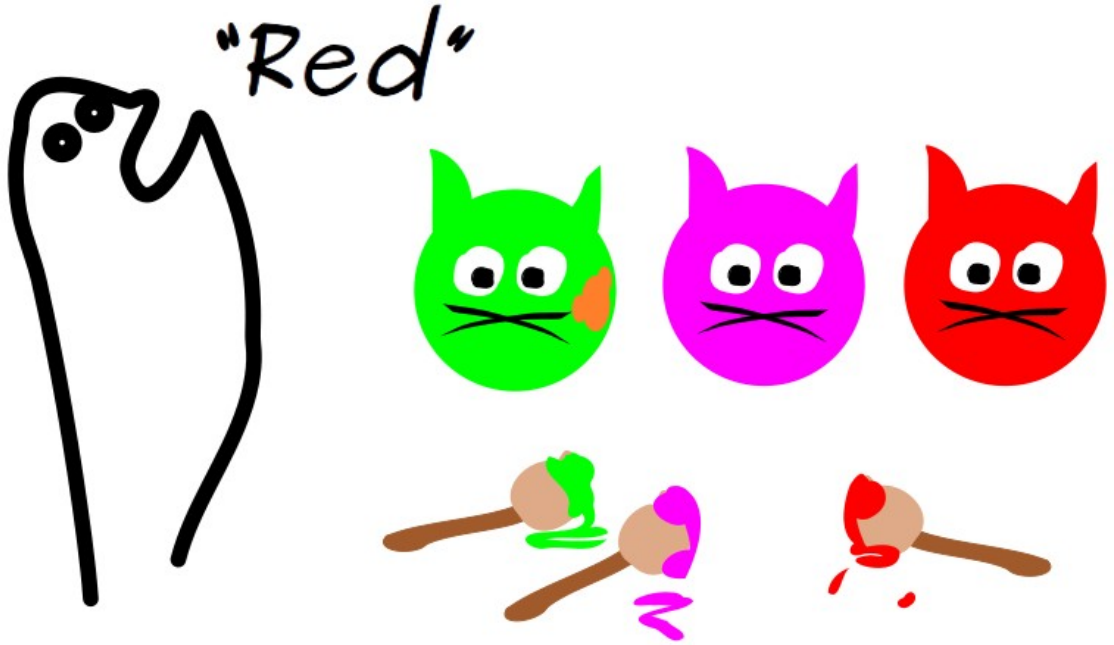
Pragmatics: +formal sem **Speaker meaning**:

The speaker's communicative intention (varies a lot).

~~FORMAL~~ Semantics: **Expression meaning**:

Some kind of fixed backbone of all uses.

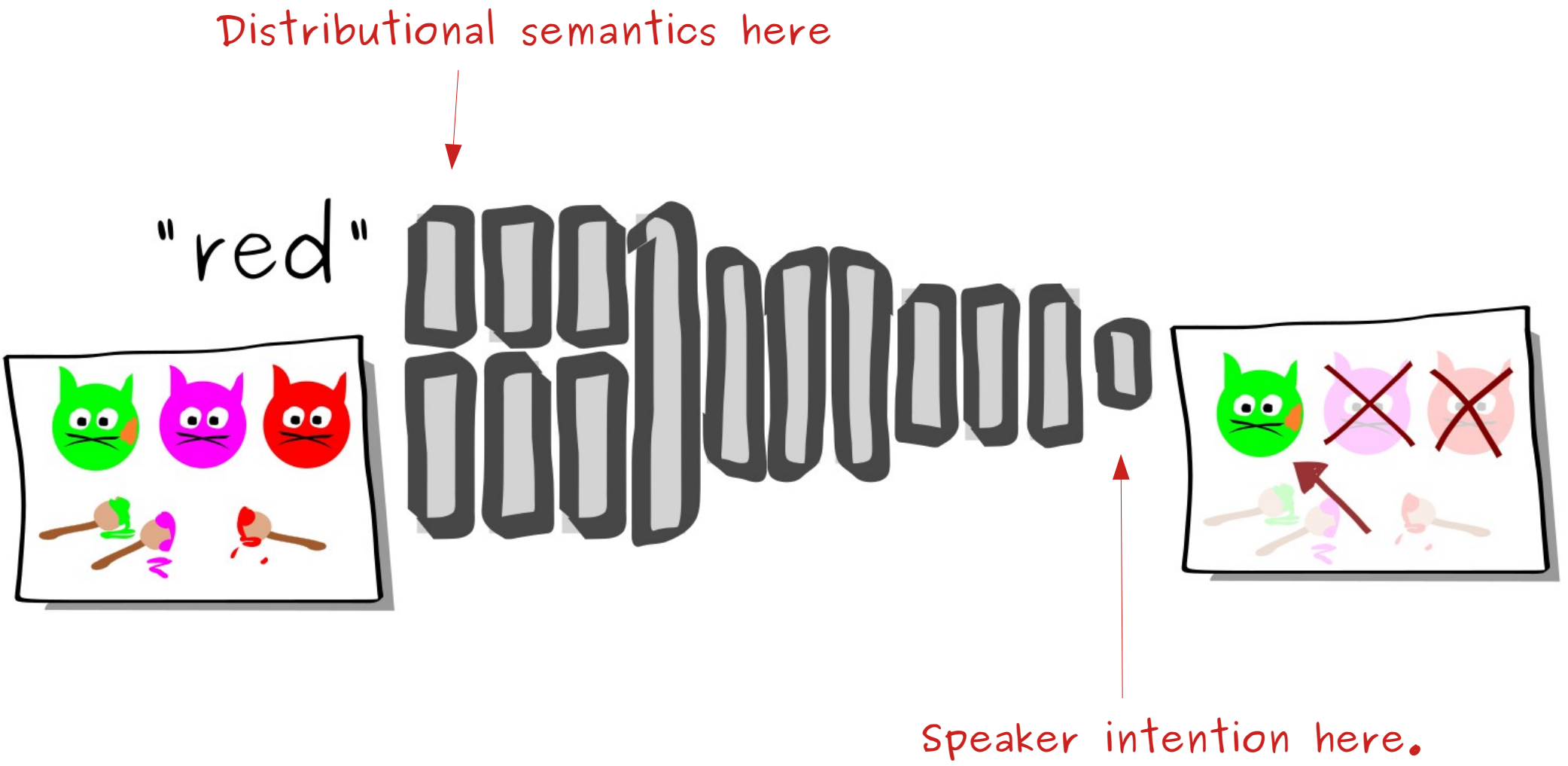
Revising the standard picture



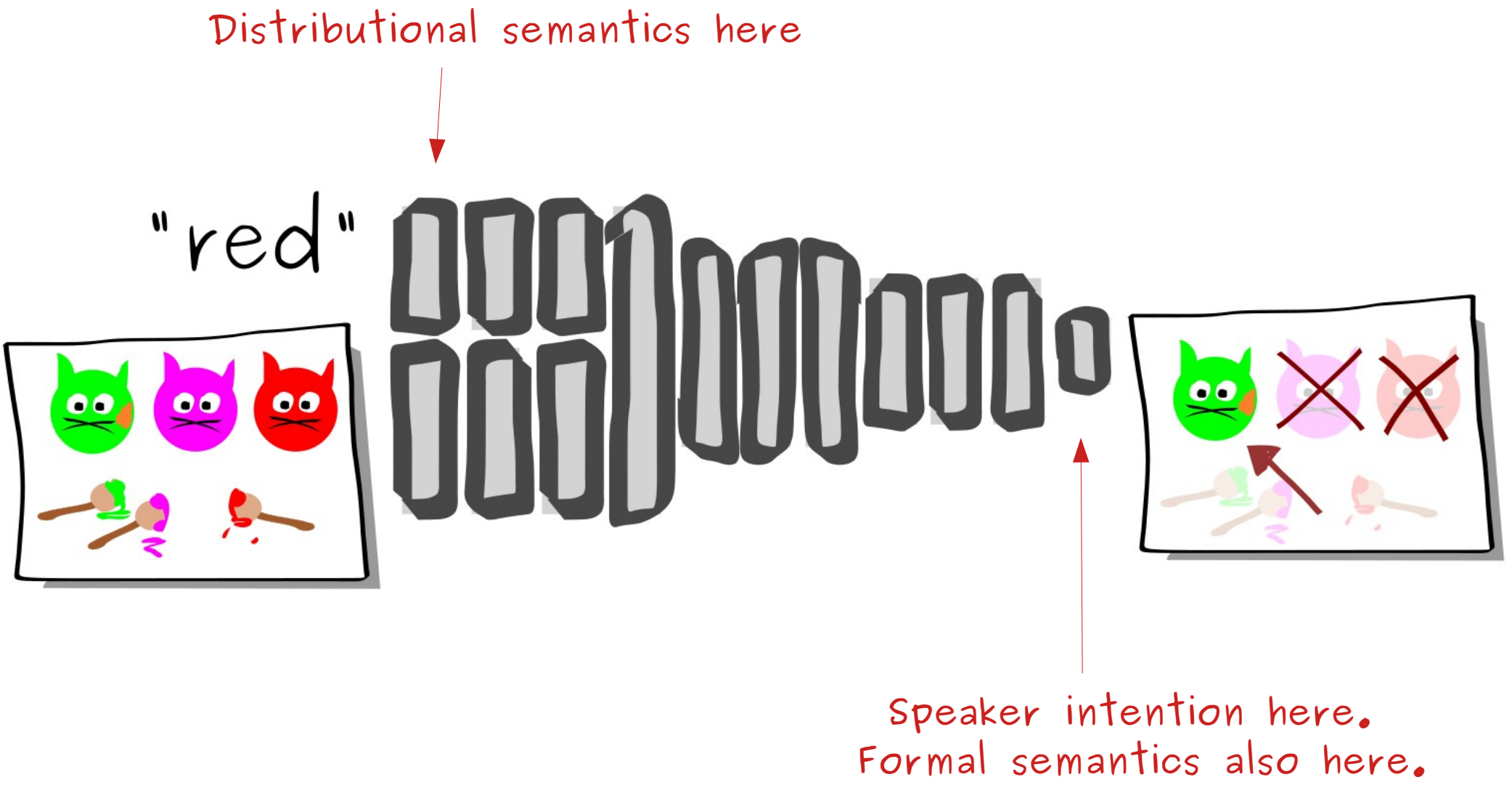
+formal Pragmatics: **Speaker meaning**:
The speaker's communicative intention (varies a lot).

~~FORMAL~~ distr. Semantics: **Expression meaning**:
Some kind of fixed backbone of all uses.

The revised picture in the age of deep learning



The revised picture in the age of deep learning



Semantics and pragmatics ...in the age of deep learning

Matthijs Westera

Image sources

https://ui-ex.com/explore/whale-transparent-dark/

https://commons.wikimedia.org/wiki/File:Cowicon.svg

https://commons.wikimedia.org/wiki/File:Bird_1010720_drawing.svg

https://commons.wikimedia.org/wiki/File:Dog_silhouette.svg

https://commons.wikimedia.org/wiki/File:Cat_silhouette_darkgray.svg

https://commons.wikimedia.org/wiki/File:Frog_(example).svg

https://commons.wikimedia.org/wiki/File:PeregrineFalconSilhouettes.svg

https://commons.wikimedia.org/wiki/File:Common_goldfish_silhouette.svg

https://en.wikipedia.org/wiki/Infinite_monkey_theorem#/media/File:Chimpanzee_seated_at_typewriter.jpg

http://chittagongit.com/icon/gallery-icon-png-25.html