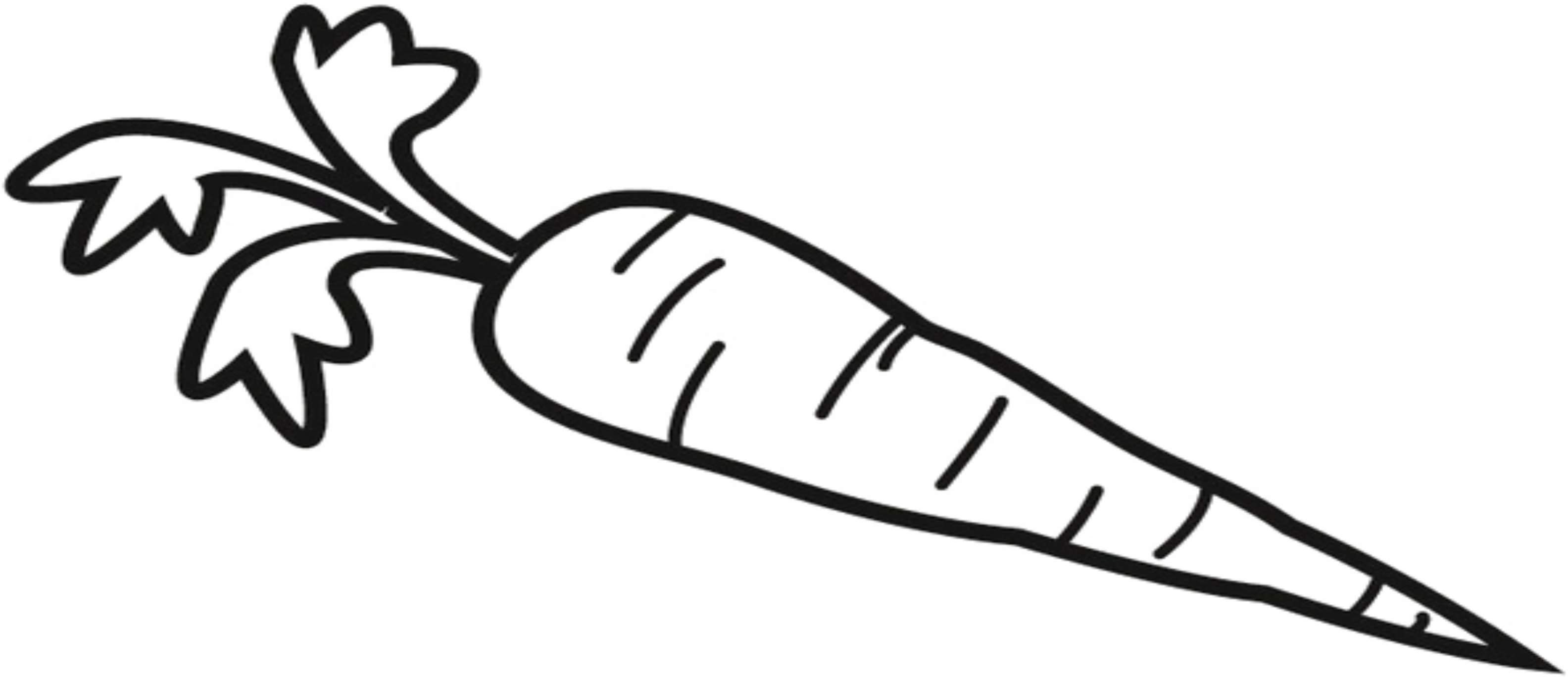
Word Representations
&
Neurocognition of Language

Unsupervised Language Learning, Spring 2018



Feature Norms

concepts are represented by features...

- Short descriptions of typical attributes for words
  - Visual Appearance
  - Function or Purpose
  - Location
  - Relationships

Feature Norms (continued)

- Human Subjects Ratings
- Useful insight into human concept acquisition
- Expensive to produce
- Available for small set of nouns
- No finite list of features that can be produced for a given concept.

Feature Norms (continued)

- One of the largest and most widely used feature-norm datasets is from McRae et al. (2005)
  - from approximately 725 participants for 541 living (*dog*) and nonliving (*chair*) basic-level concepts

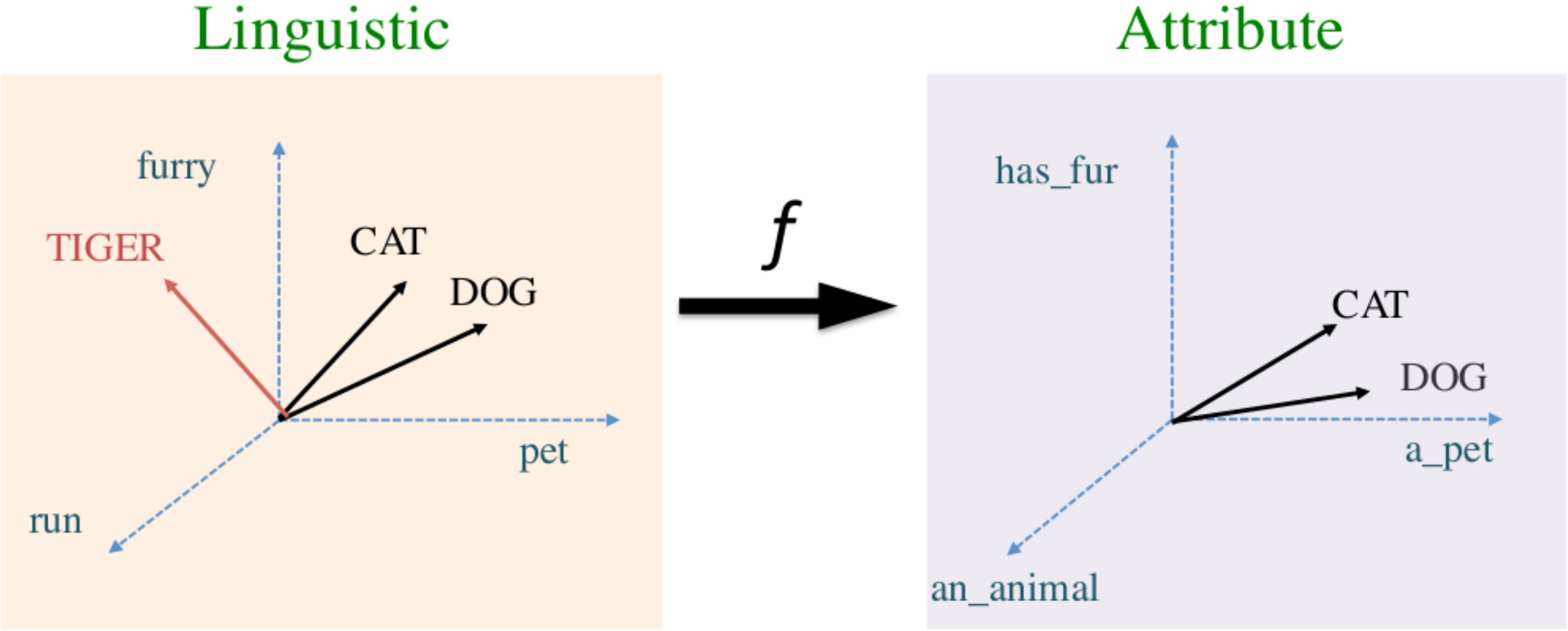
SHRIMP	CUCUMBER	DRESS
is_edible, 19	a_vegetable, 25	clothing, 21
is_small, 17	eaten_in_salads, 24	worn_by_women, 15
lives_in_water, 12	is_green, 23	is_feminine, 10
is_pink, 11	is_long, 15	is_formal, 10
tastes_good, 9	eaten_as_pickles, 12	is_long, 10
has_a_shell, 8	has_skin, 9	different_styles, 9
lives_in_oceans, 8	grows_in_gardens, 7	made_of_material, 9

Automatic Computation of Feature Norms

- From distributional semantics to feature norms *(L. Fagarasan et al.)*
  - Feature-based semantic space → FS
  - Co-occurance based distributional models→DS

	has_fur	has_wheels	an_animal	a_pet
cat_FS	22	0	21	17
	dog	black	book	animal
cat_DS	4516	3124	1500	2480

- Learn a function to map the two semantic spaces: ***f: FS → DS***



Linguistic
Attribute
furry
TIGER
CAT
DOG
pet
run
f
has_fur
CAT
DOG
a_pet
an_animal

From distributional semantics to feature norms ***(L. Fagarasan et al.)***

- DS1: *contexts are the top 10K most frequent content words in Wikipedia, values are raw co- occurrence counts.*
- DS2: *same contexts as DS1, counts are re-weighted using PPMI and normalized as detailed in Polajnar and Clark (2014).*
- DS3: *perform SVD to 300 dimensions on DS2.*
- DS4: *same as DS3 but with row normalization performed after dimensionality reduction.*
- DS5: *word2vec (300d)*

From distributional semantics to feature norms *(L. Fagarasan et al.)*

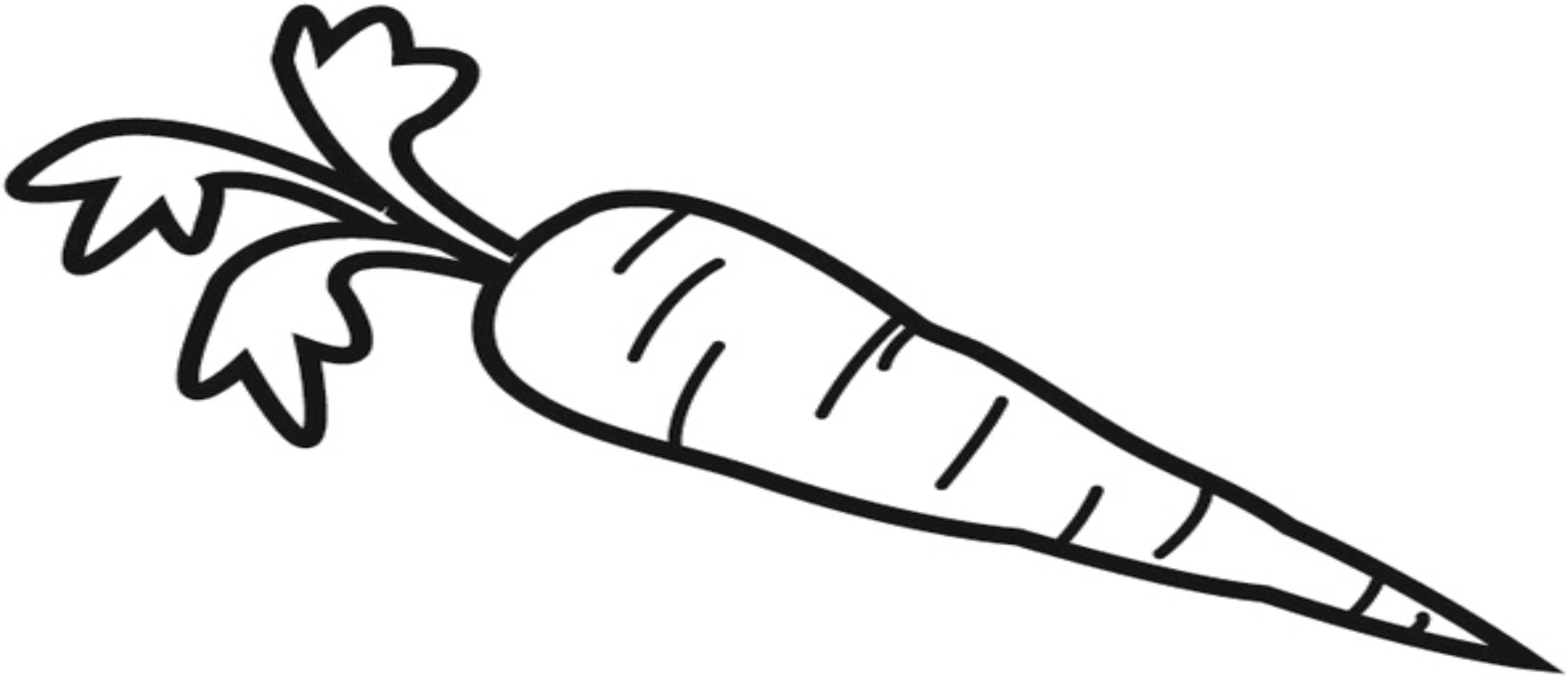
DS	FS	**top1**	**top5**	**top10**	**top20**	**MAP**
RAND	-	0.37	0.74	1.85	3.70	-
DS1	FS1	0.72	14.49	29.71	49.28	0.30
DS2	FS1	2.90	12.32	23.91	47.10	0.29
DS3	FS1	2.90	13.04	24.64	49.28	0.37
DS3	FS2	2.17	15.22	26.09	50.00	-
DS4	FS2	3.62	15.22	25.36	49.28	-
DS5	FS1	1.45	14.49	24.64	44.20	0.29
DS5	FS2	1.45	19.57	26.09	46.38	-

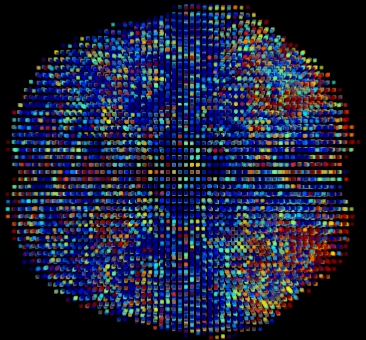
Table 3: Percentage (%) of test items that retrieve their gold-standard vector in the topN neighbours of their predicted vector.

From distributional semantics to feature norms *(L. Fagarasan et al.)*

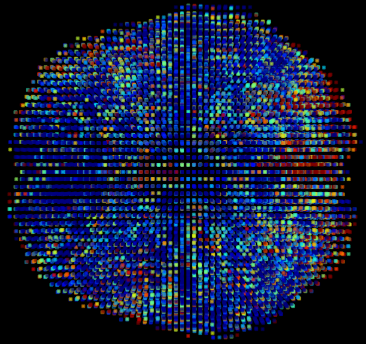
Word	Nearest neighbours of predicted vector	Result	Top weighted predicted features
JAR	bucket, strainer, pot, spatula	not top20	made_of_plastic, is_round*, made_of_metal, found_in_kitchens*
JEANS	shawl, shirt, blouse, sweater	not top20	clothing, different_colours, worn_by_women*
BUGGY	skateboard, truck, scooter, cart	in top20	has_wheels, made_of_wood*, is_large*, used_for_transportation
SEAWEED	shrimp, perch, trout, salmon	in top20	is_edible, lives_in_water*, is_green, swims*, is_small*
HORSE	cow, ox, sheep, donkey	in top10	an_animal, has_4_legs, is_large, has_legs, lives_on_farms
PLATYPUS	otter, salamander, turtle, walrus	in top10	an_animal, is_small*, lives_in_water, is_long*,
SPARROW	starling, finch, partridge, sparrow	in top5	a_bird, flies, has_feathers, has_a_beak, has_wings
SPATULA	strainer, spatula, grater, colander	in top5	made_of_metal, found_in_kitchens, made_of_plastic
HATCHET	hatchet, machete, sword, dagger	in top1	made_of_metal, is_sharp, has_a_handle, a_tool, a_weapon*
GUN	gun, rifle, bazooka, shotgun	in top1	used_for_killing, a_weapon, made_of_metal, is_dangerous

Table 4: Qualitative analysis of predicted vectors (obtained by mapping from DS3 to FS1) for 10 concepts in the test set. Features annotated with an asterix(*) are not listed in the gold standard feature vector for the given concepts.

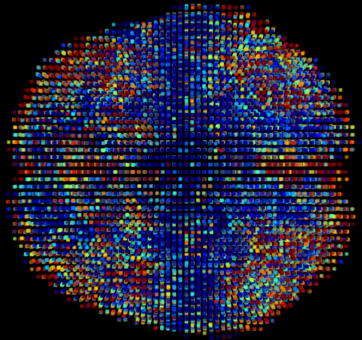




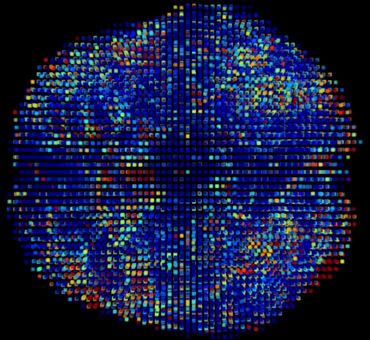
bicycle



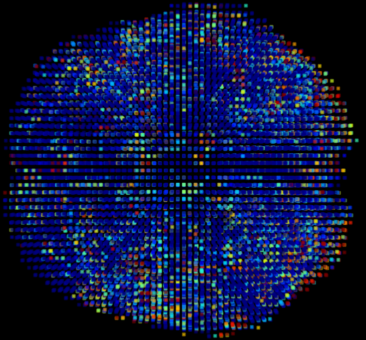
car



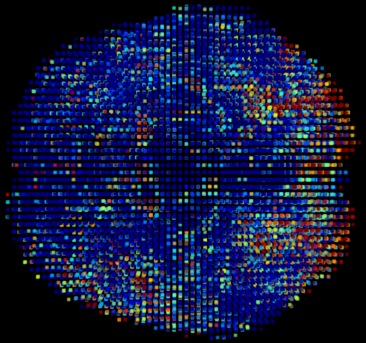
airplane



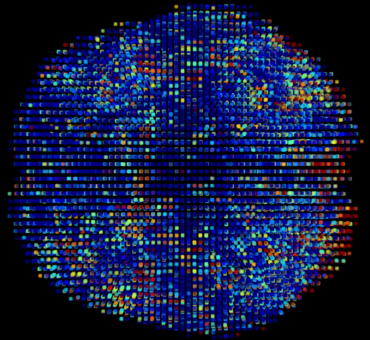
bottle



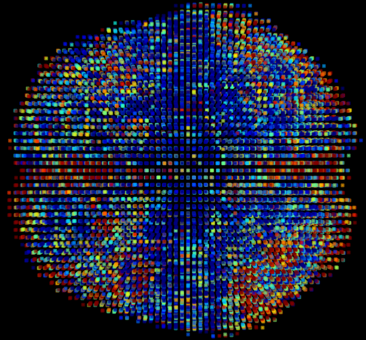
ant



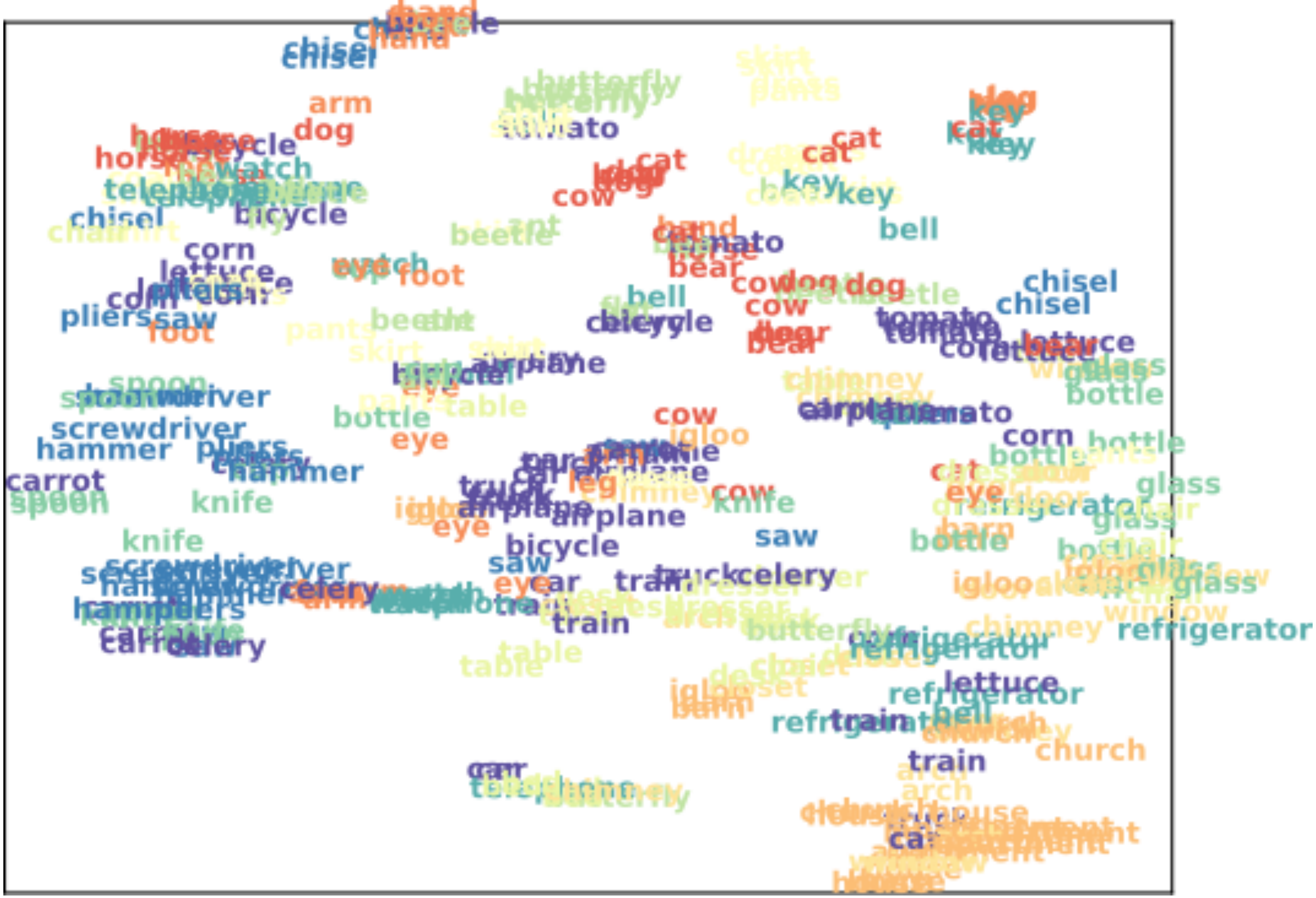
bee



cat



fly


chisel
arm
dog
horse
bicycle
watch
telephone
chisel
corn
lettuce
pliers
saw
foot
eye
beetle
butterfly
tomato
cat
cow
key
bell
hand
bear
skirt
pants
airplane
spoon
screwdriver
hammer
carrot
knife
bottle
table
igloo
truck
leg
chimney
celery
train
car
dresser
refrigerator
glass
window
chair
door
barn
closet
arch
church
house

Semantic Organization in the Human Brain

- ❖ Region Based
  - ➢ certain categories (such as faces and body parts) are strongly represented in discrete, highly localized regions of cortex
- ❖ Distributed
  - ➢ certain categories (such as household objects) are represented by distributed patterns of activity

A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain (A. G. Huth et al.)

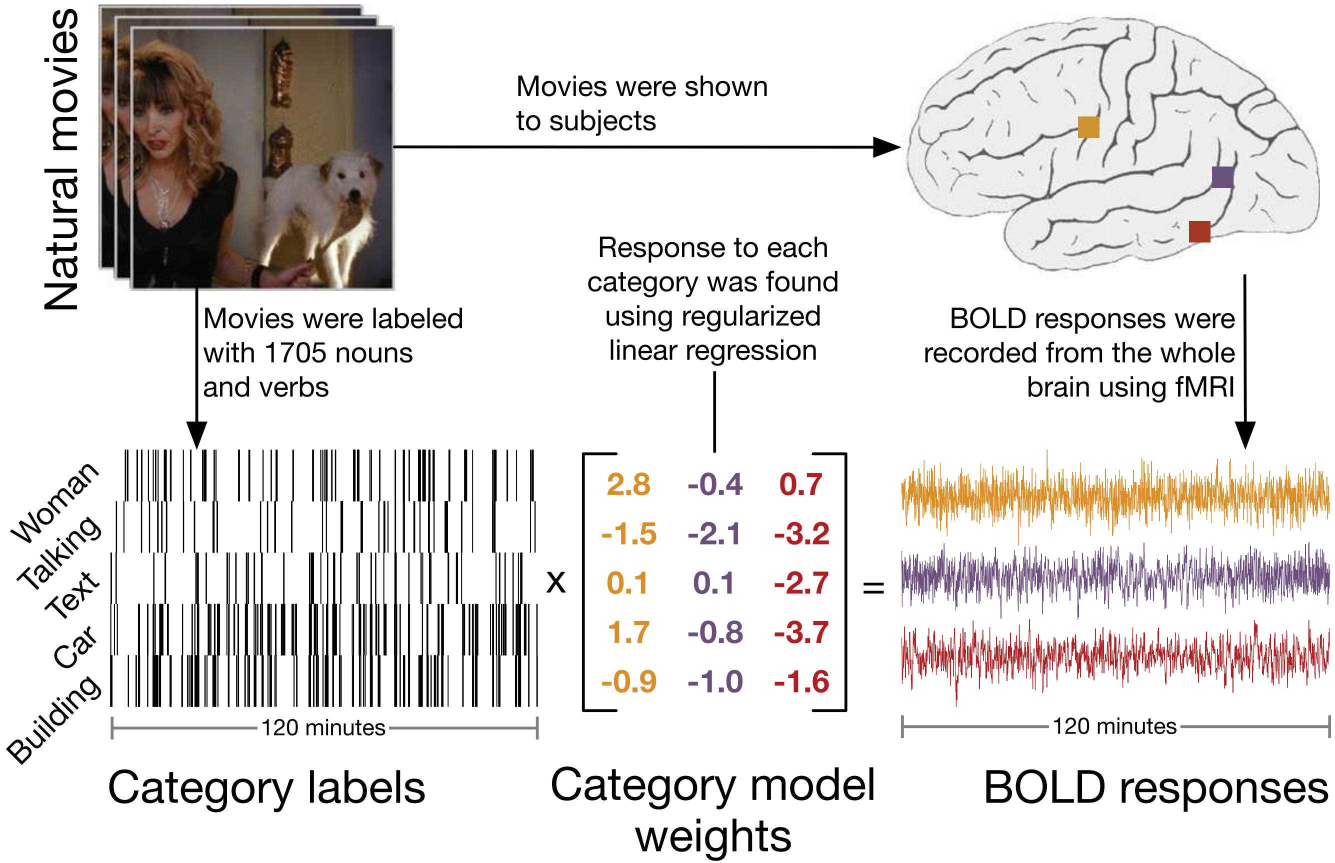
- Humans can see and name thousands of distinct object and action categories, so it is unlikely that each category is represented in a distinct brain area.

- fMRI: measure human brain activity evoked by natural movies

    - fMRI measures brain activity by detecting changes associated with blood flow. This technique relies on the fact that cerebral blood flow and neuronal activation are coupled.

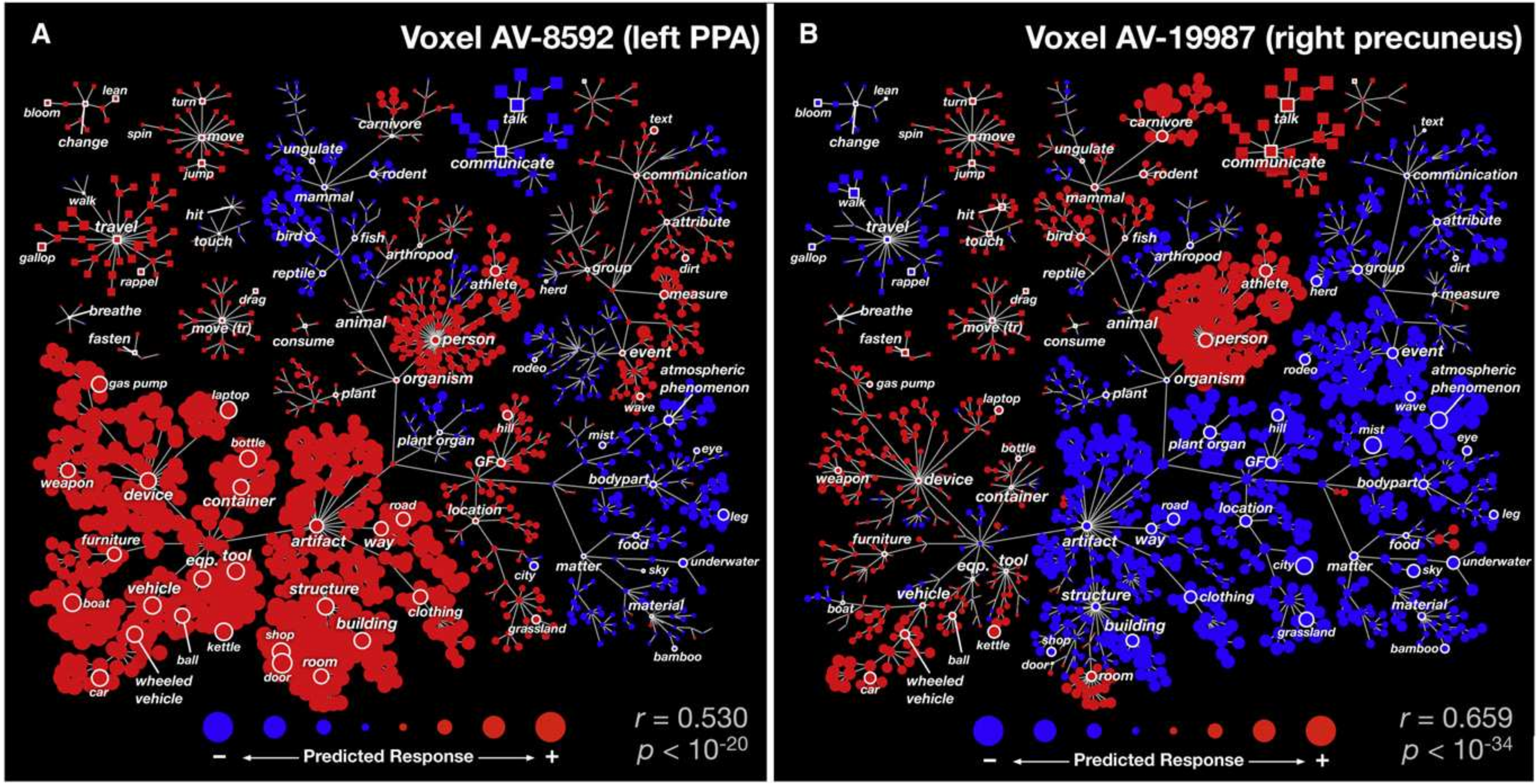
- Build voxel wise models: examine the cortical representation of 1,705 object and action categories

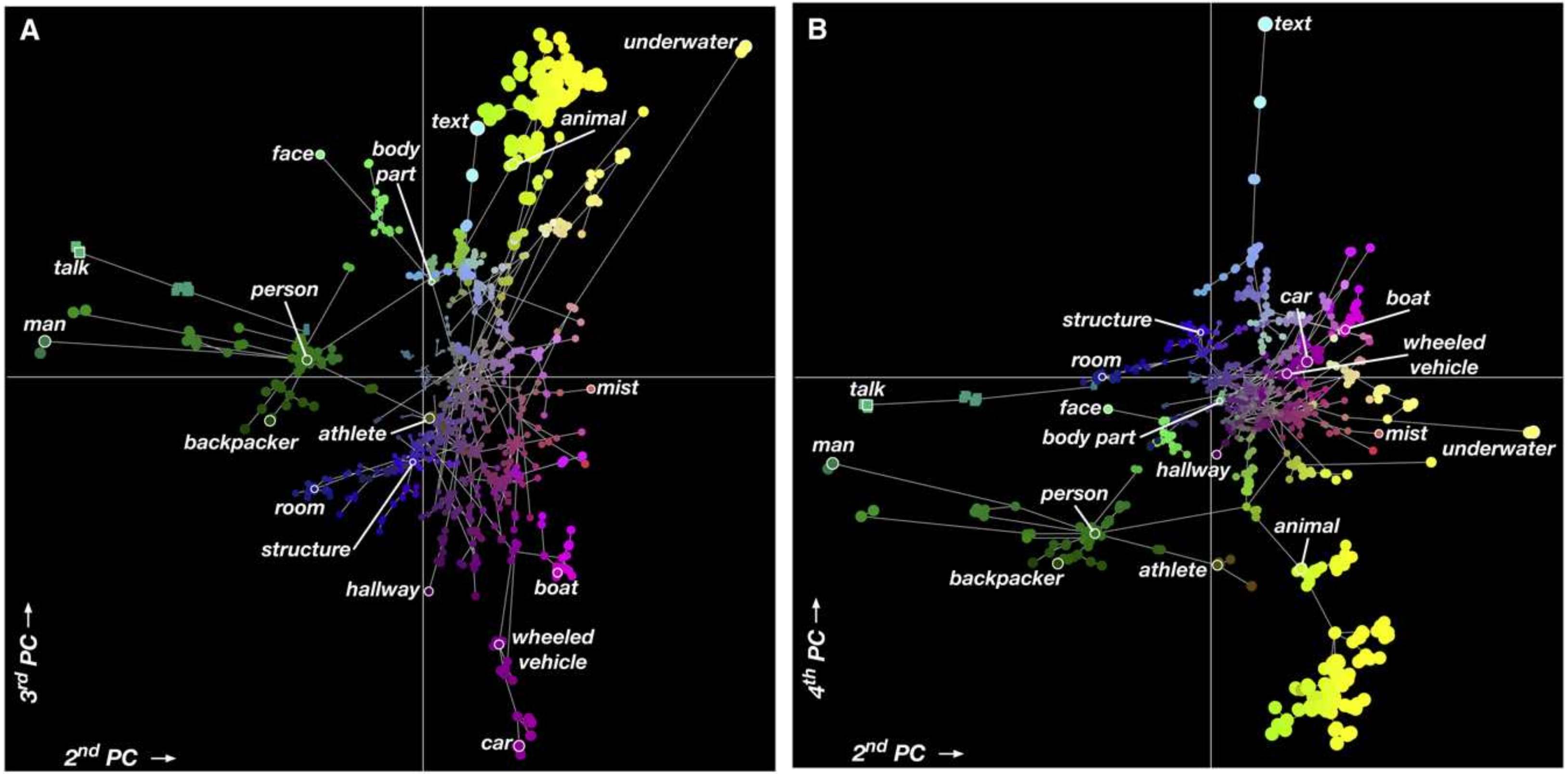
    - Each **voxel** can represent a million or so brain cells.

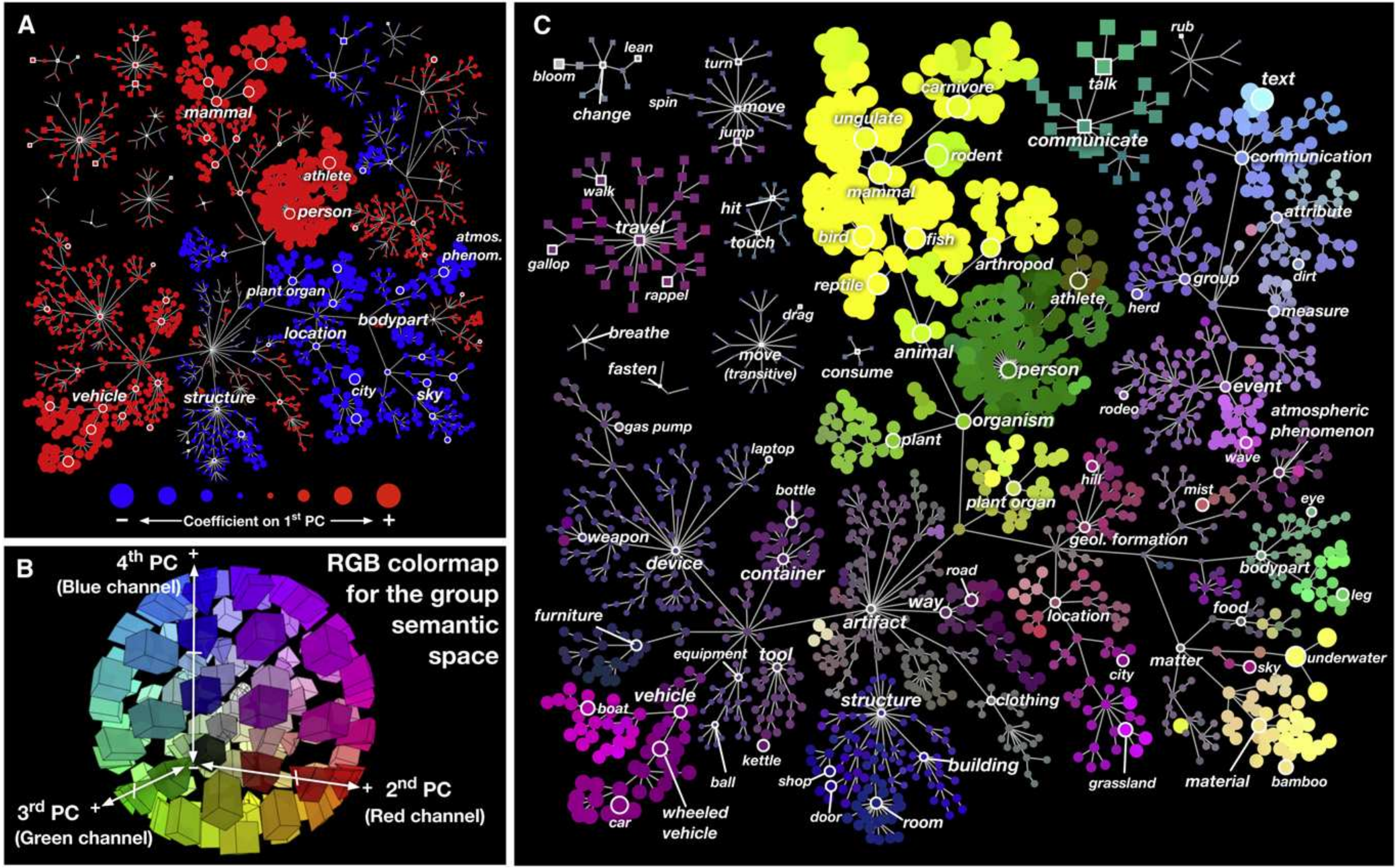


Natural movies
Movies were shown to subjects
Movies were labeled with 1705 nouns and verbs
Response to each category was found using regularized linear regression
BOLD responses were recorded from the whole brain using fMRI
Woman
Talking
Text
Car
Building
120 minutes
x
2.8 -0.4 0.7
-1.5 -2.1 -3.2
0.1 0.1 -2.7
1.7 -0.8 -3.7
-0.9 -1.0 -1.6
=
120 minutes
Category labels
Category model weights
BOLD responses

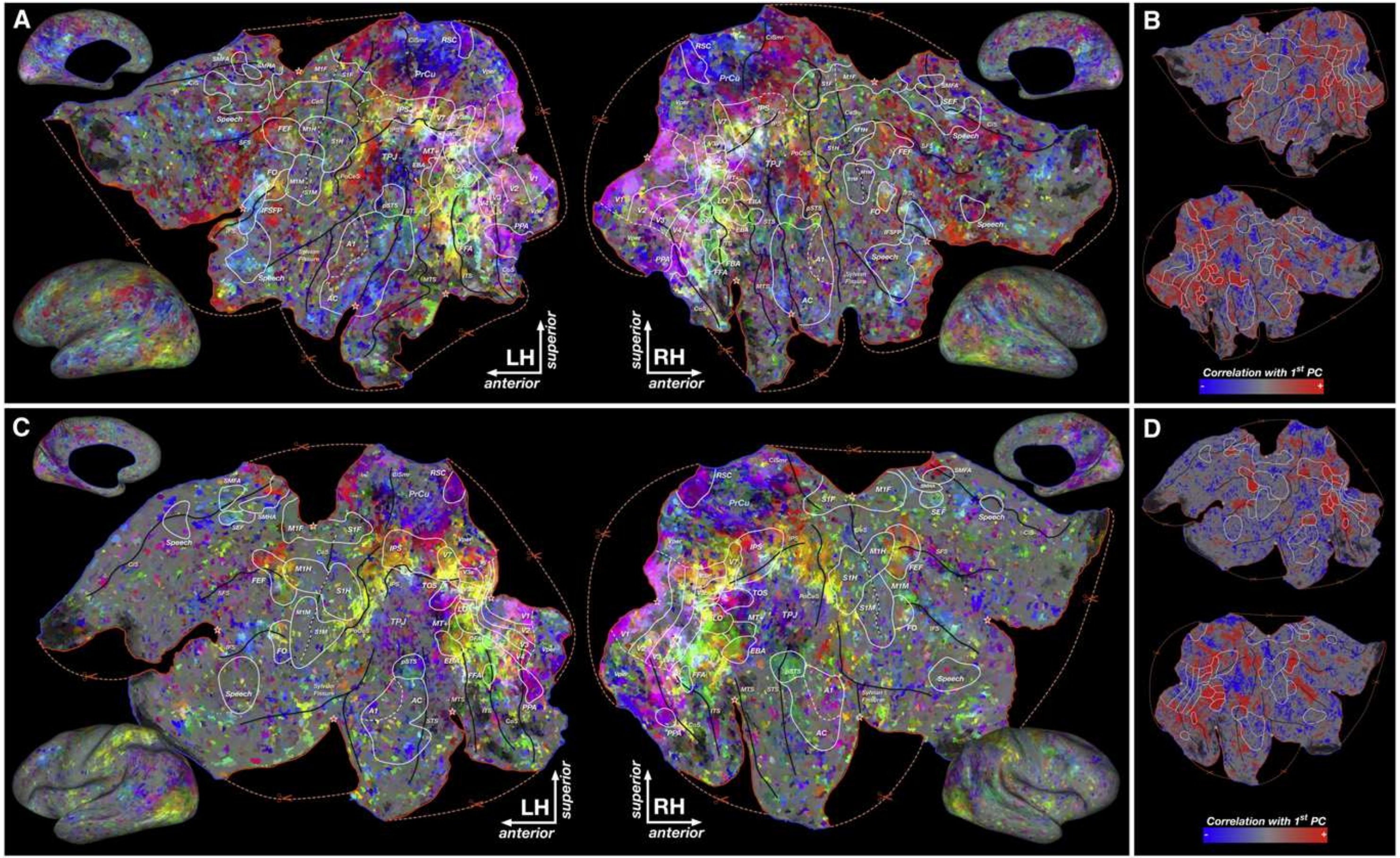
Movie Clip	Labels	Movie Clip	Labels
	butte.n.01 desert.n.01 sky.n.01 cloud.n.01 brush.n.01		city.n.01 expressway.n.01 skyscraper.n.01 traffic.n.01 sky.n.01
	woman.n.01 talk.v.02 gesticulate.v.01 book.n.01		bison.n.01 walk.v.01 grass.n.01 stream.n.01
	hammerhead.n.01 swim.v.01 water.n.01		woman.n.01 man.n.01 talk.v.02


A
Voxel AV-8592 (left PPA)
r = 0.530
p < 10^-20
Predicted Response
B
Voxel AV-19987 (right precuneus)
r = 0.659
p < 10^-34
Predicted Response


A
underwater
text
animal
face
body
part
talk
person
man
mist
backpacker
athlete
room
structure
hallway
boat
wheeled
vehicle
car
3rd PC →
2nd PC →
B
text
car
boat
structure
wheeled
vehicle
room
talk
face
mist
body part
underwater
man
hallway
person
animal
backpacker
athlete
4th PC →
2nd PC →



A
mammal
athlete
person
atmos. phenom.
plant organ
bodypart
location
vehicle
structure
city
sky
– Coefficient on 1st PC +
B
4th PC (Blue channel)
RGB colormap for the group semantic space
3rd PC (Green channel)
2nd PC (Red channel)
C
bloom
lean
change
turn
spin
move
jump
rub
talk
communicate
text
communication
carnivore
ungulate
rodent
mammal
bird
fish
reptile
arthropod
attribute
dirt
group
herd
measure
walk
travel
gallop
rappel
hit
touch
drag
breathe
fasten
move (transitive)
consume
animal
athlete
person
organism
plant
event
rodeo
atmospheric phenomenon
wave
gas pump
laptop
bottle
plant organ
hill
mist
eye
weapon
device
container
geol. formation
bodypart
leg
road
way
furniture
artifact
location
food
equipment
tool
matter
underwater
sky
city
vehicle
boat
structure
clothing
kettle
ball
shop
door
building
room
grassland
material
bamboo
car
wheeled vehicle



A
B
C
D
LH
RH
superior
anterior
Correlation with 1st PC
-
+

A Continuous Semantic Space Describes the Representation of Thousands of Object and Action Categories across the Human Brain (A. G. Huth et al.)

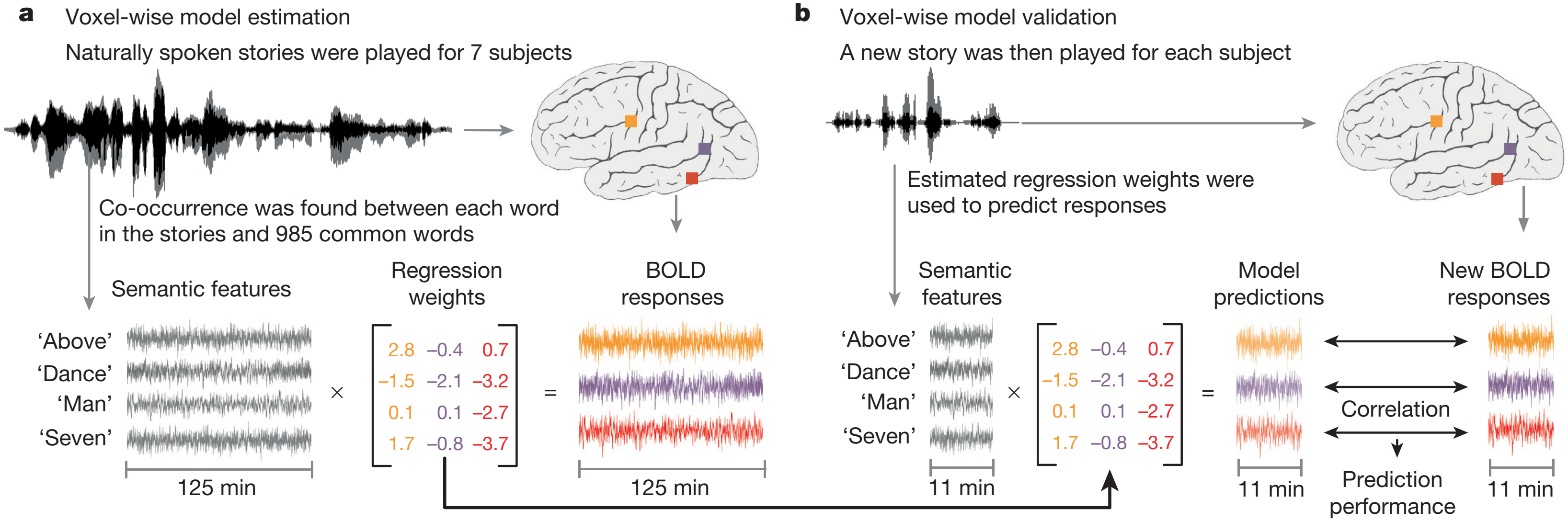
- Projection of the recovered semantic space onto cortical flat maps shows that semantic selectivity is organized into smooth gradients that cover much of visual and nonvisual cortex

- Both the recovered semantic space and the cortical organization of the space are shared across different individuals

Natural speech reveals the semantic maps that tile human cerebral cortex (A. G. Huth et al.)

- The semantic system is organized into intricate patterns that seem to be consistent across individuals.
- Generative model to create a detailed semantic atlas.
- Most areas within the semantic system represent information about specific semantic domains, or groups of related concepts, and their atlas shows which domains are represented in each area.

Natural speech reveals the semantic maps that tile human cerebral cortex (A. G. Huth et al.)



Natural speech reveals the semantic maps that tile human cerebral cortex (A. G. Huth et al.)

- **The embedding space:**
  - Normalized co-occurrence between each word and a set of 985 common English words across a large corpus of English text.
    - such as ‘above’, ‘worry’ and ‘mother’
  - Words related to the same semantic domain tend to occur in similar contexts, and so have similar co-occurrence values.
    - For example, the words ‘month’ and ‘week’ are very similar (the correlation between the two is 0.74), while the words ‘month’ and ‘tall’ are not (correlation −0.22).

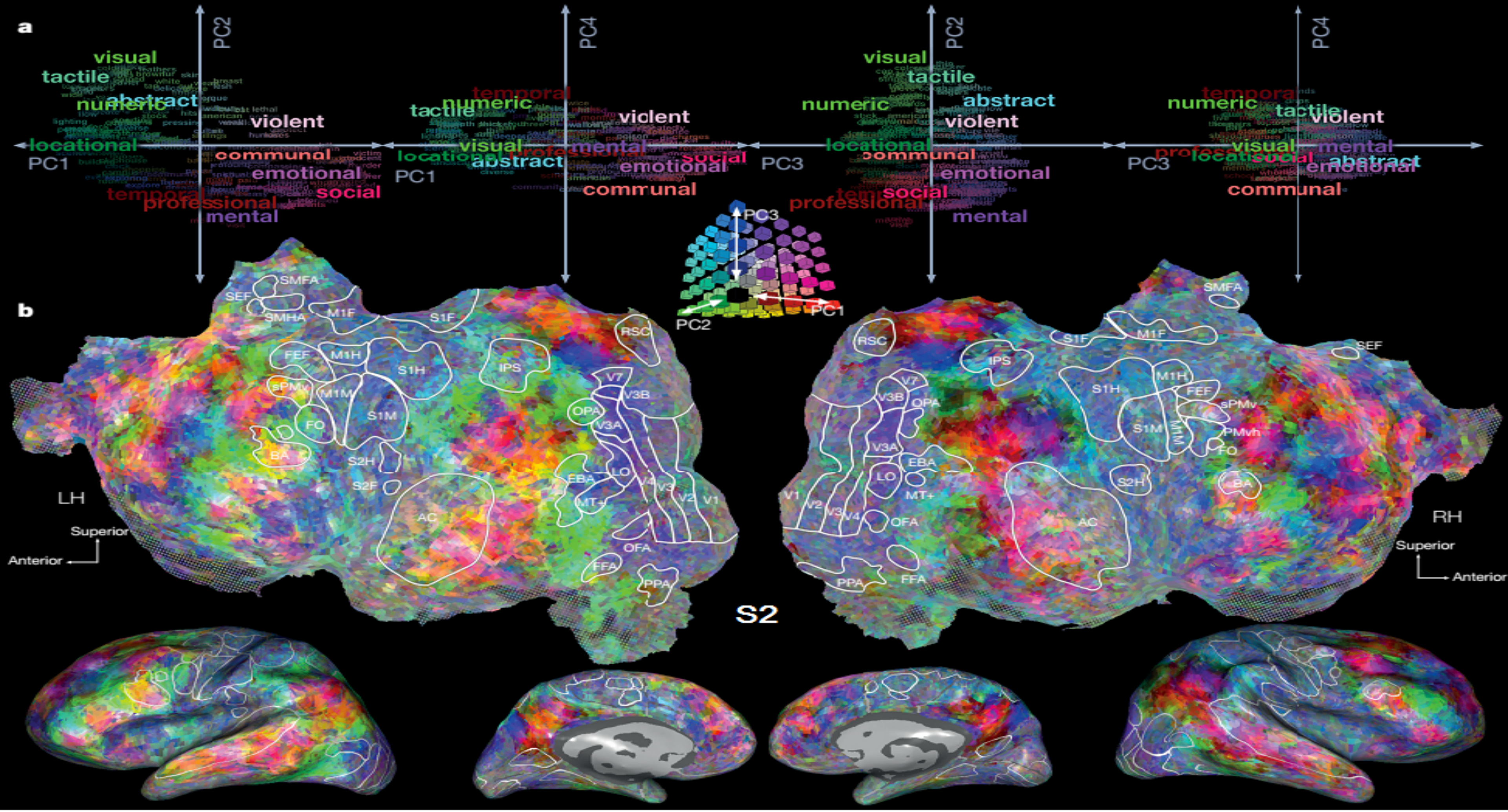
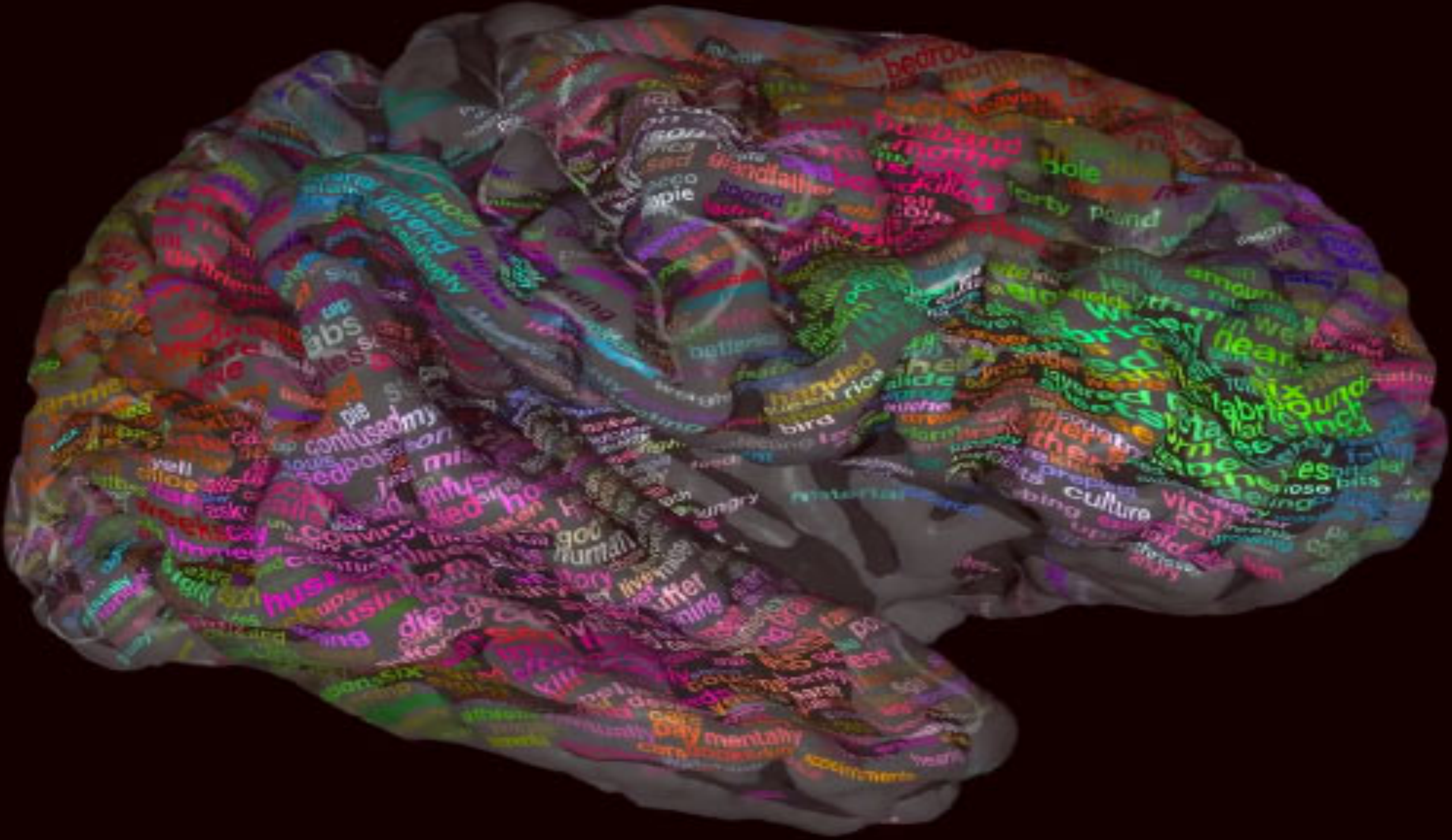
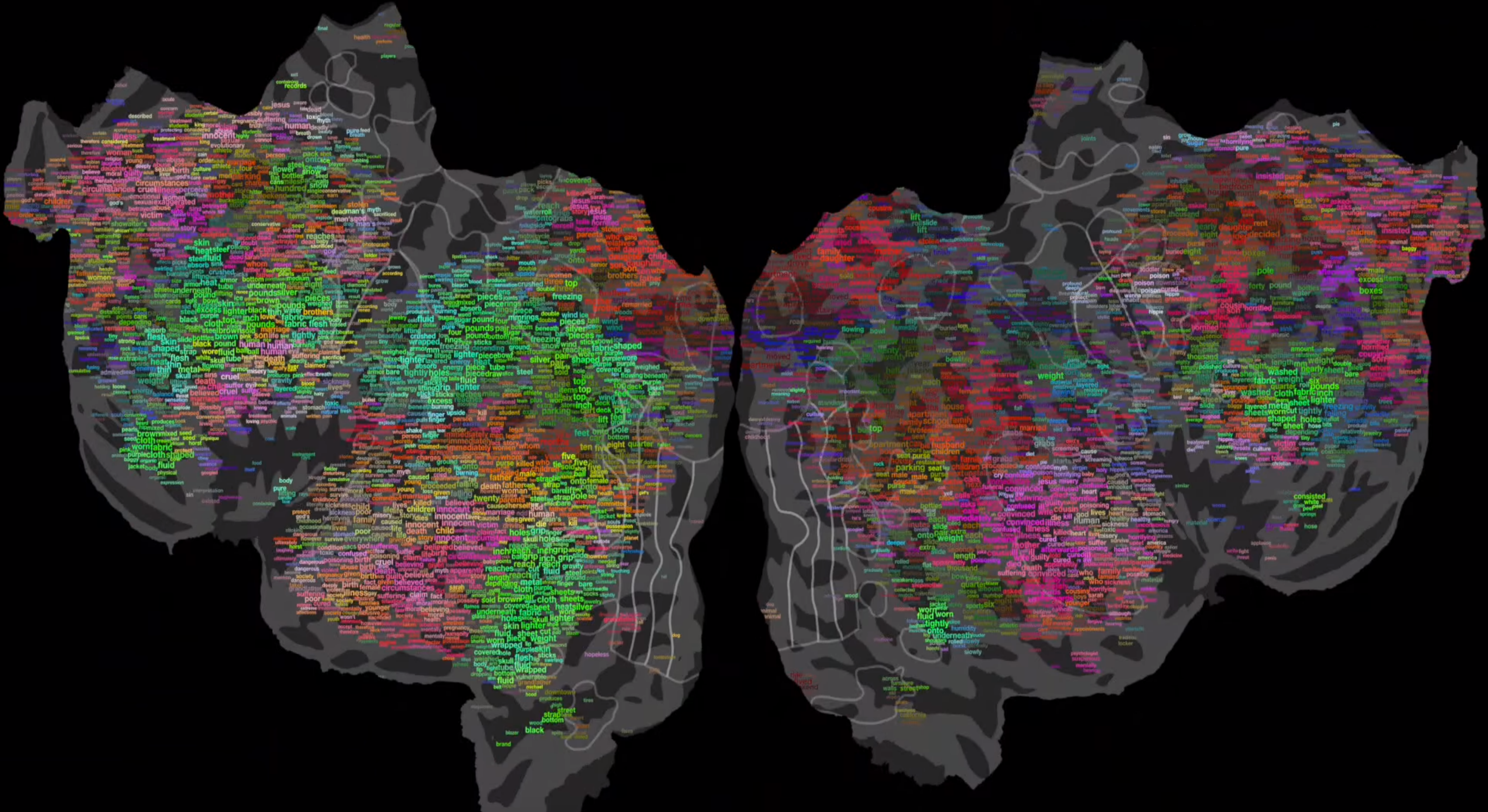


Figure 2 | Principal components of voxel-wise semantic models.


husband
hole
forty
pound
grandfather
pie
up
confused
yell
rice
bird
hungry
culture
god
died
six
inch
mentally



Natural speech reveals the semantic maps that tile human cerebral cortex (A. G. Huth et al.)

- The distribution of semantically selective areas is relatively symmetrical across the two cerebral hemispheres.
- Organization of semantically selective brain areas seems to be highly consistent across individuals.
- Demonstrates the power and efficiency of data-driven approaches for functional mapping of the human brain.
    - Although the experiment used a simple design in which subjects only listened to stories, the data were rich enough to produce a comprehensive atlas of semantically selective areas.
- General data-driven framework:
    - Other properties of language can be mapped (even in this same data set) by using feature spaces that reflect phonemes, syntax and so on.


Perception
Aphasia
Word
ERP
Bilingualism
Imaging
Brain
Sentence
Cognitive
Lexicon
Comprehension
Writing
Bases
Processing
Neuroscience
Recognition
Neurocognition
Biological
Production
Evolution
Development
Semantics
Reading
Sign
Syntax
Speech
Language

Questions?