Concepts, Composition, and Conversational Coordination

Semantic Competence for Situated Interaction

3rd Seminar: Concepts (Part II), Composition

David Schlangen
University of Potsdam, Germany

http://clp.ling.uni-potsdam.de

https://github.com/davidschlangen/cosine-paris

plan

the seminar series:

- intro: the problem & the approach
- concepts [Mon, Sep 23]
- concepts (still!), composition [Mon, Sep 30]
- conversational coordination / dialogue [Mon, Oct 7]

functional *reprsnt.nal*

Conceptual Apparatus

Inference

- word to word
- discourse resolution

Reference

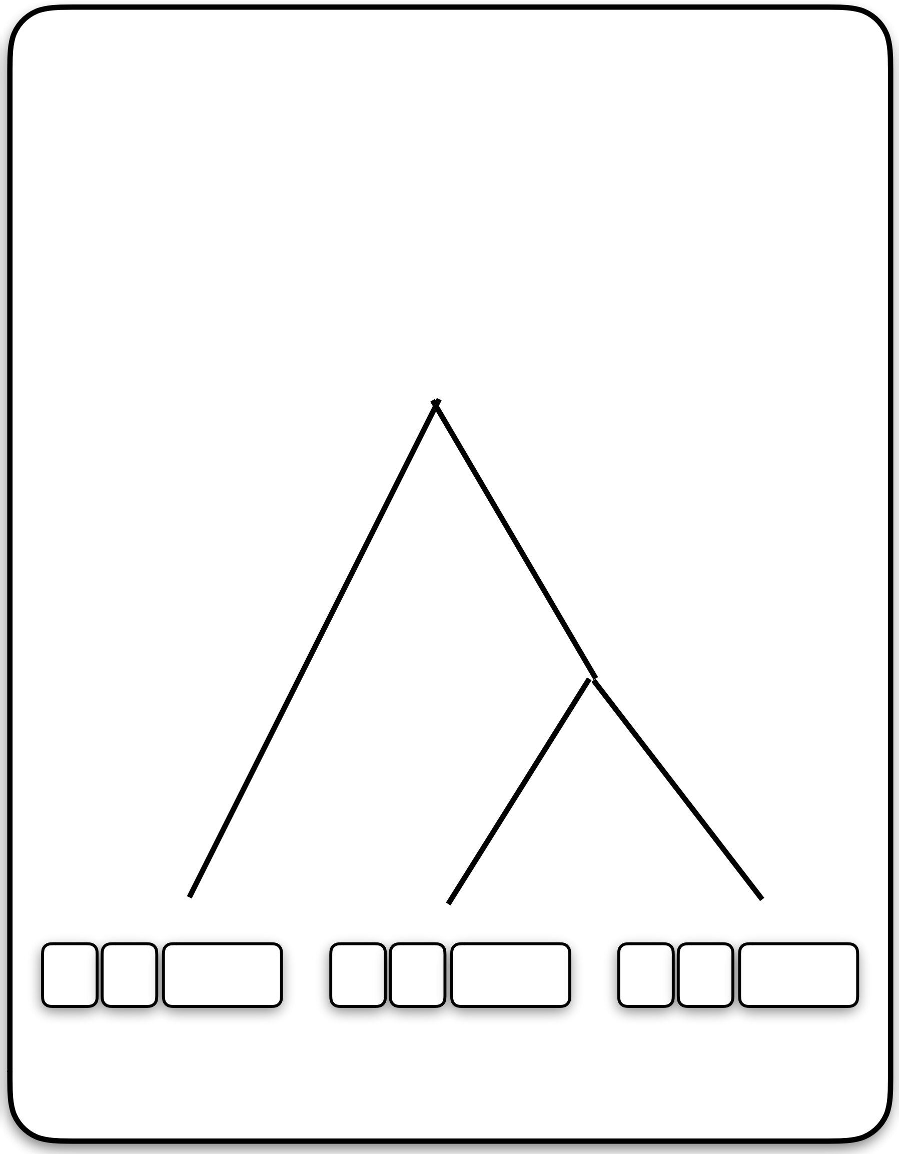
- word to world
- categorisation
- naming / resolution

symbolic repr.

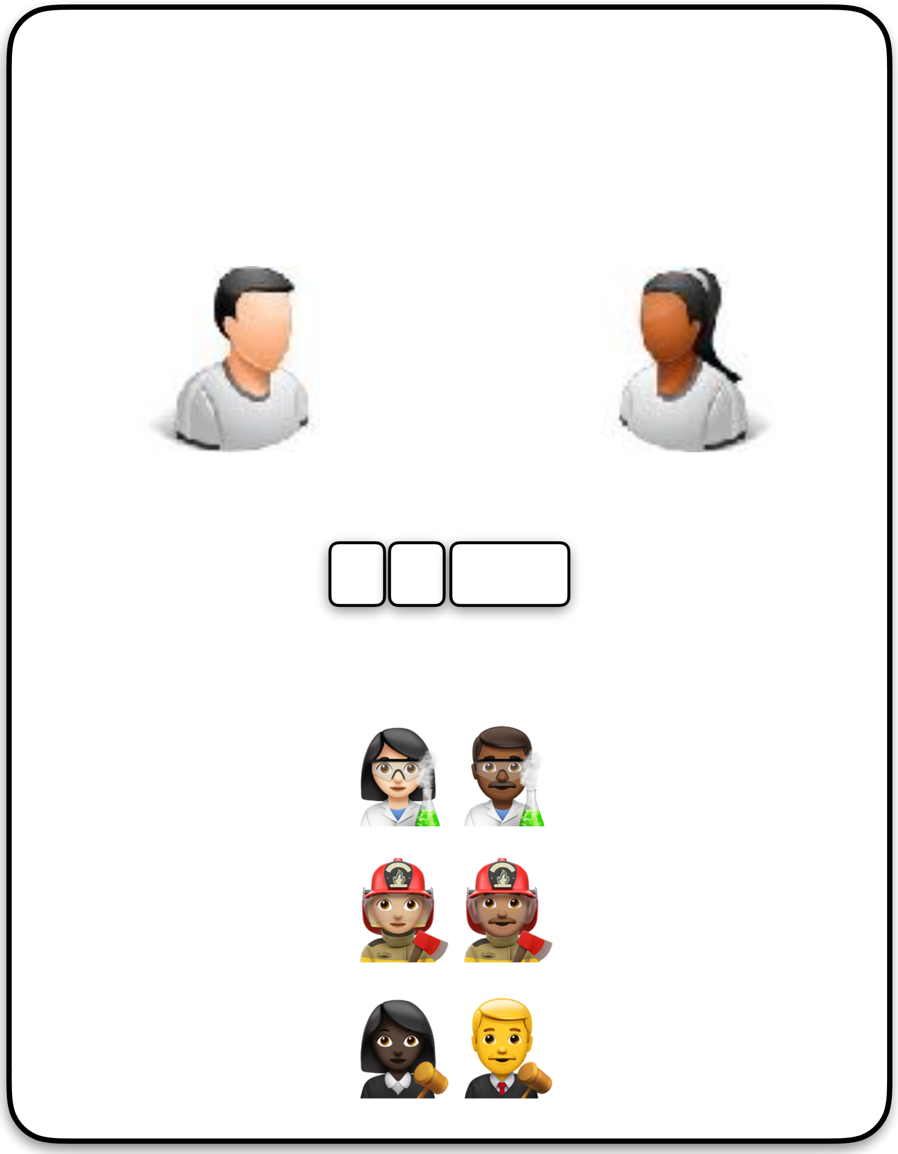
continuous repr.

classifiers on perceptual input

composition



coordination



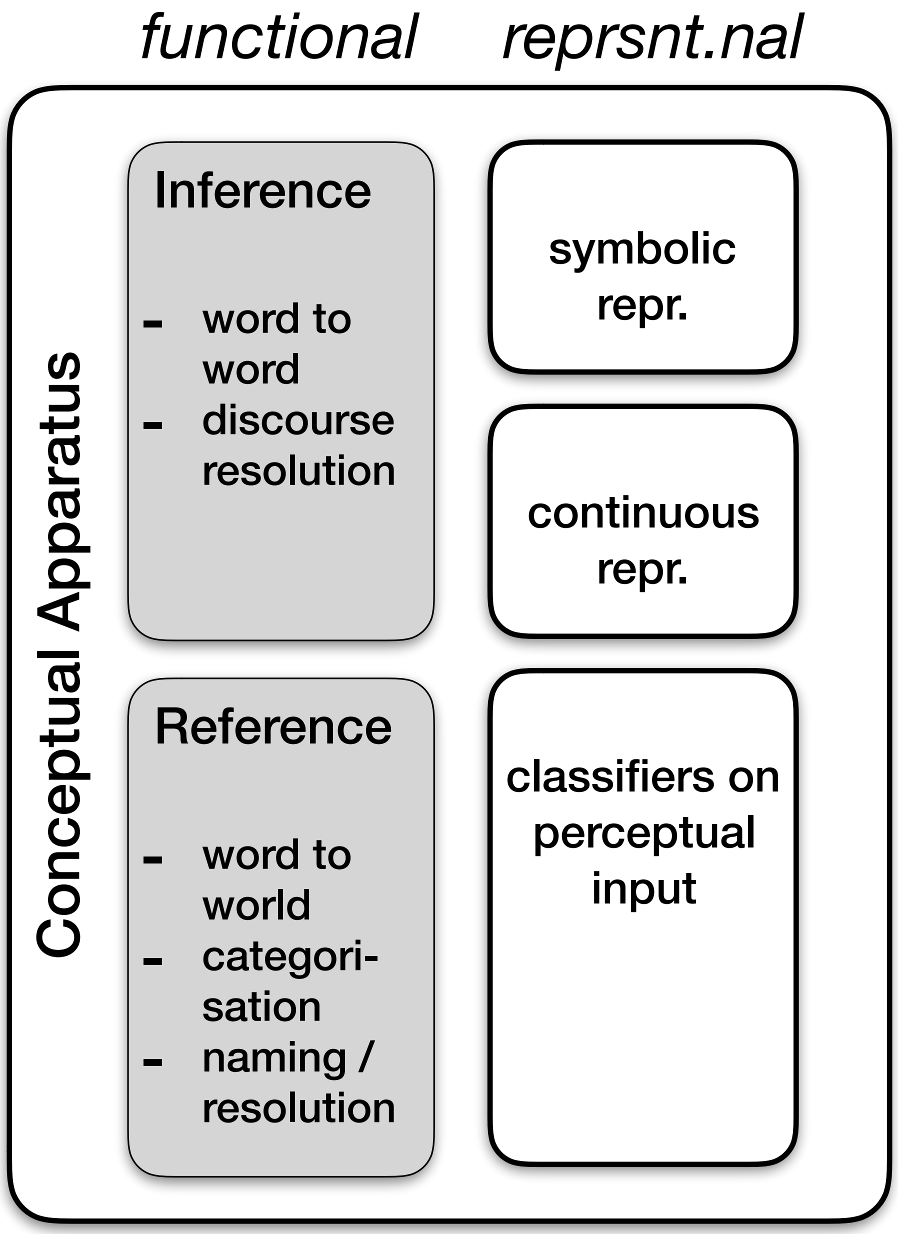
Look at the white dog!

We just saw a cute dog.

The cutest poodle ever!

Actually, that wasn't a poodle. It was too tall. It was a labradoodle.

- learning
  - incremental (within concept, within vocab)
  - fast
- implemented & tested on real data



functional *reprsnt.nal*

Conceptual Apparatus

Inference

- word to word
- discourse resolution

Reference

- word to world
- categori-sation
- naming / resolution

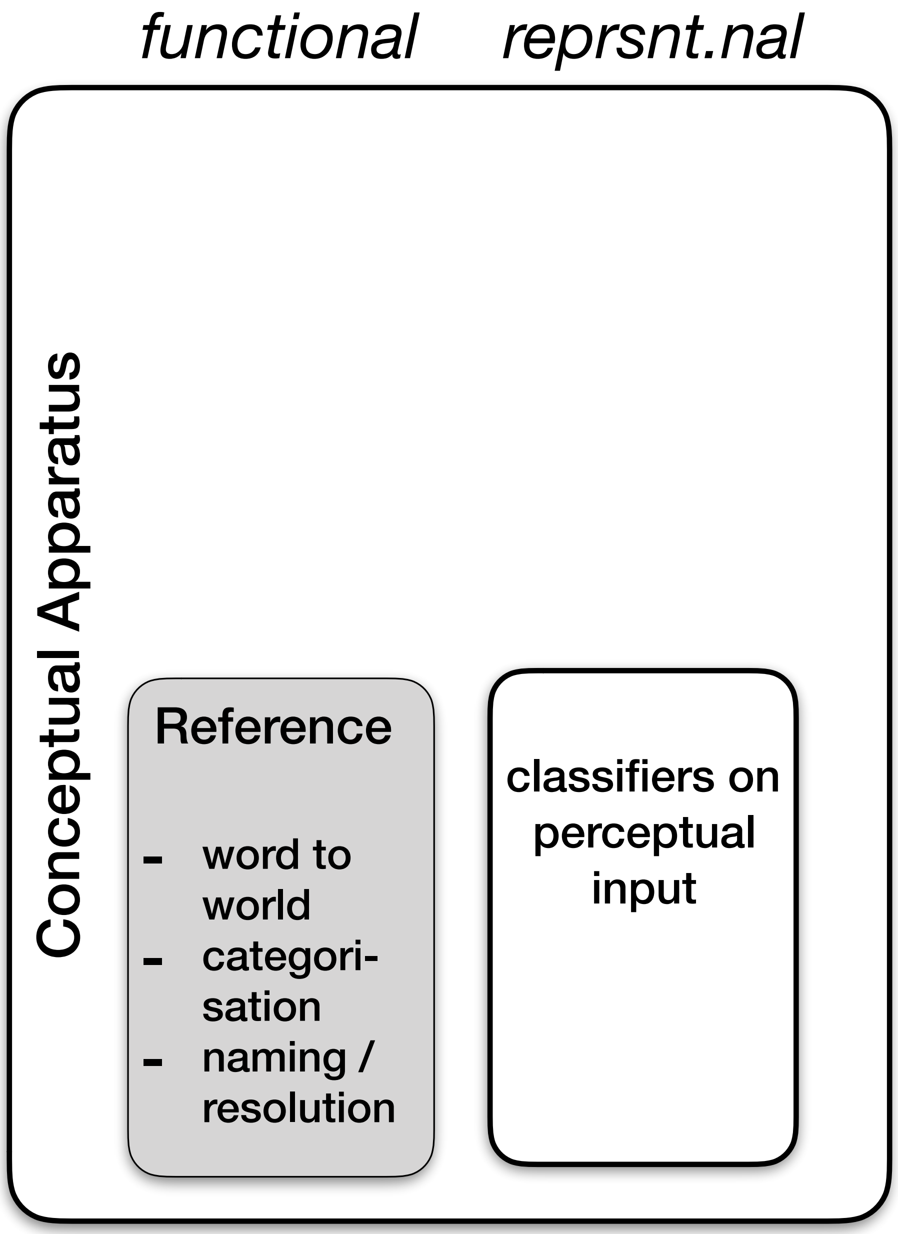
symbolic repr.

continuous repr.

classifiers on perceptual input

Look at the white dog!
We just saw a cute dog.
The cutest poodle ever!

- learning
  - incremental (within concept, within vocab)
  - fast
- implemented & tested on real data



functional *reprsnt.nal*

Conceptual Apparatus

Reference

- word to world
- categorisation
- naming / resolution

classifiers on perceptual input

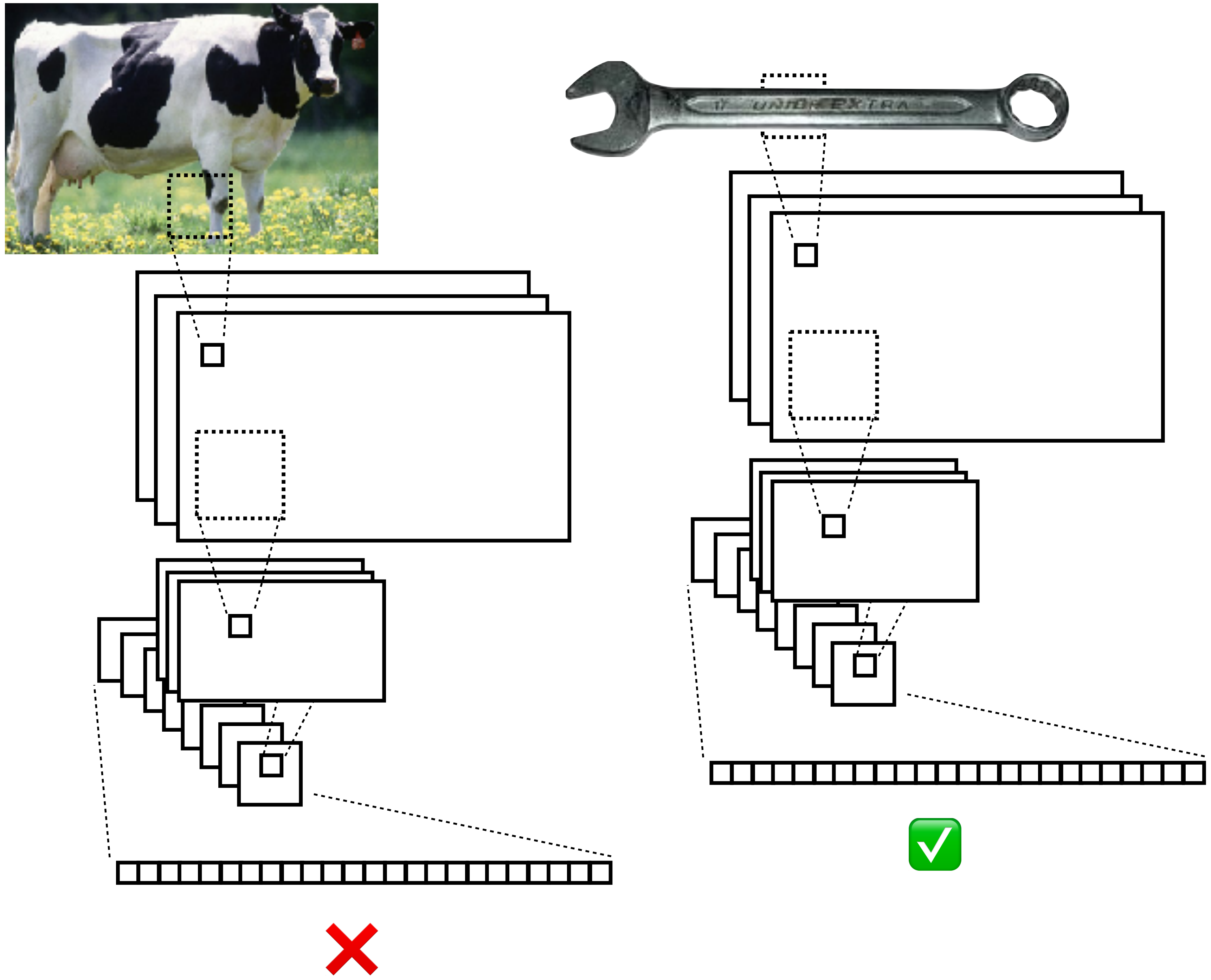
$$[[\mathrm{dog}]]^D = \{ (o, f_{dog}(o)) \}$$

what kind of function is this? (what is the range?)

where do we get this function?

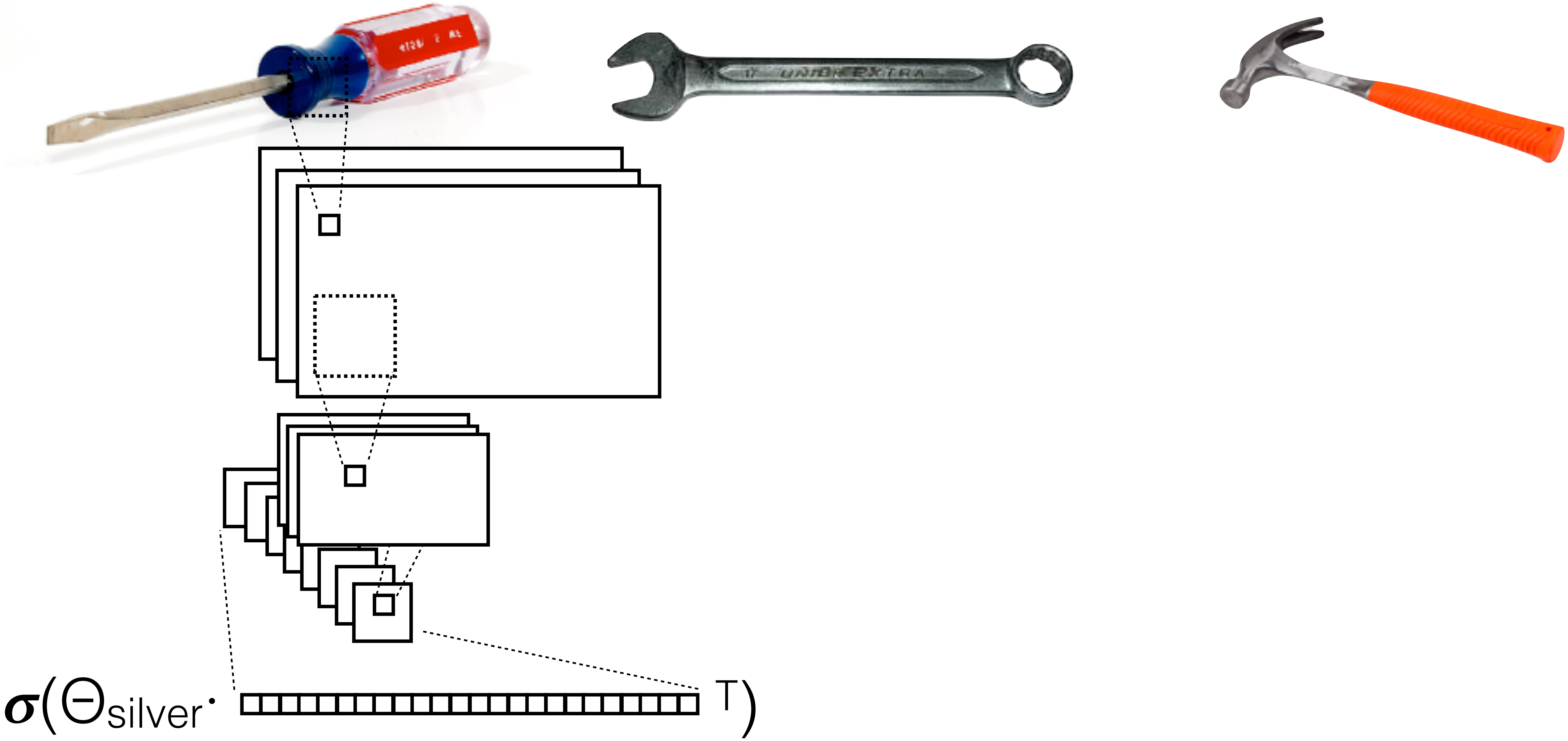
how do we present image object to function?

what kind of set is this?



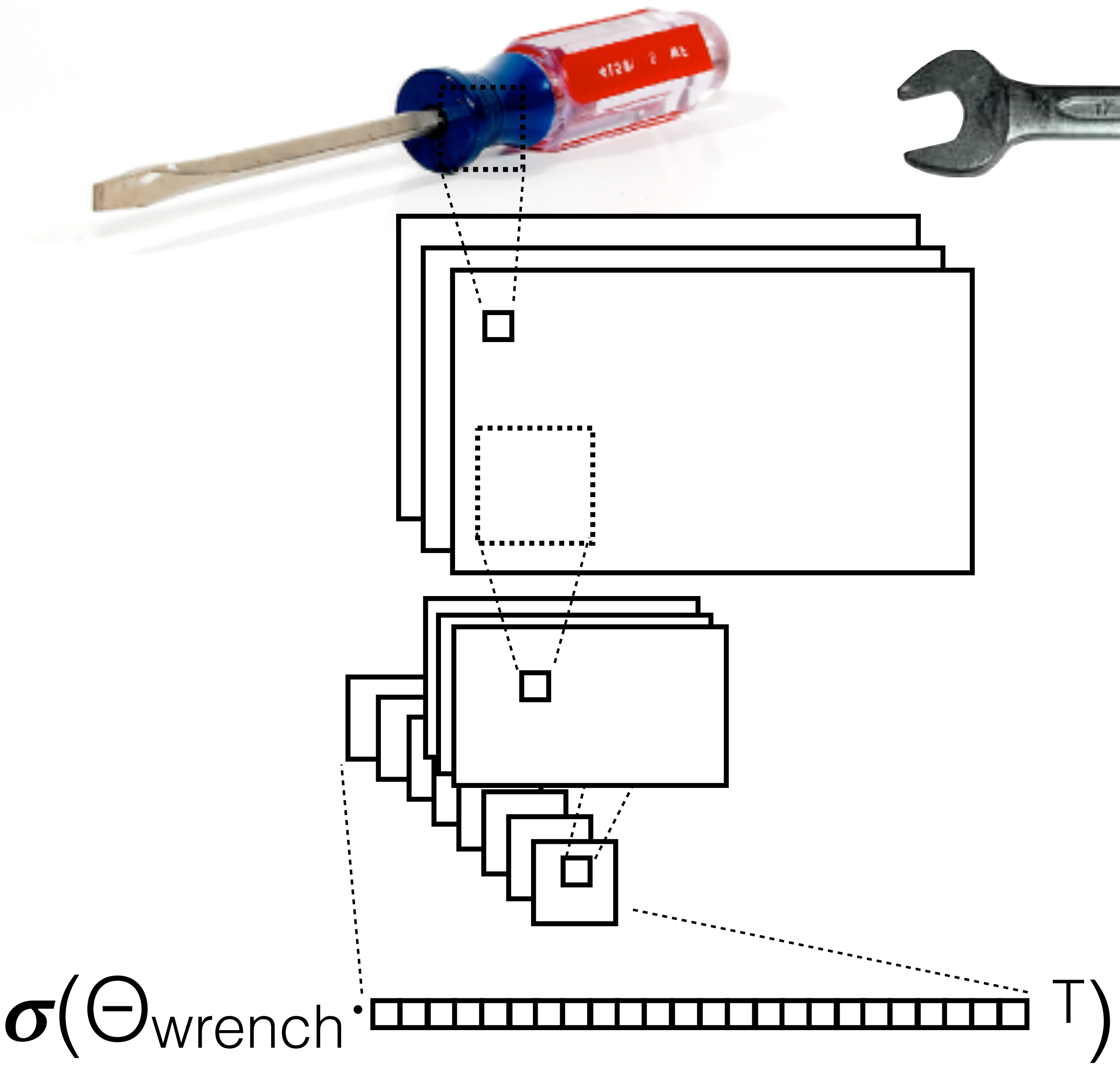
- one classifier per word / concept
- can always add new ones
- can always improve existing ones

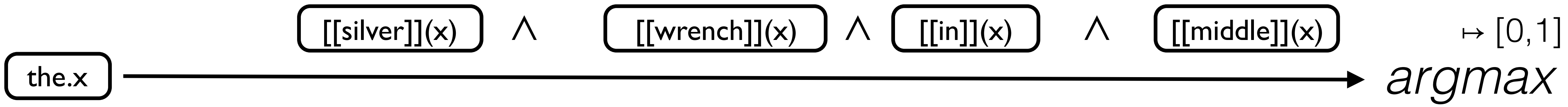
$$\boldsymbol{\sigma}(\Theta_{\text{silver}} \cdot \square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\square\ ^{T})$$



[[silver]](x) ∧ [[wrench]](x) ∧ [[in]](x) ∧ [[middle]](x) ↦ [0,1]

the.x → *argmax*





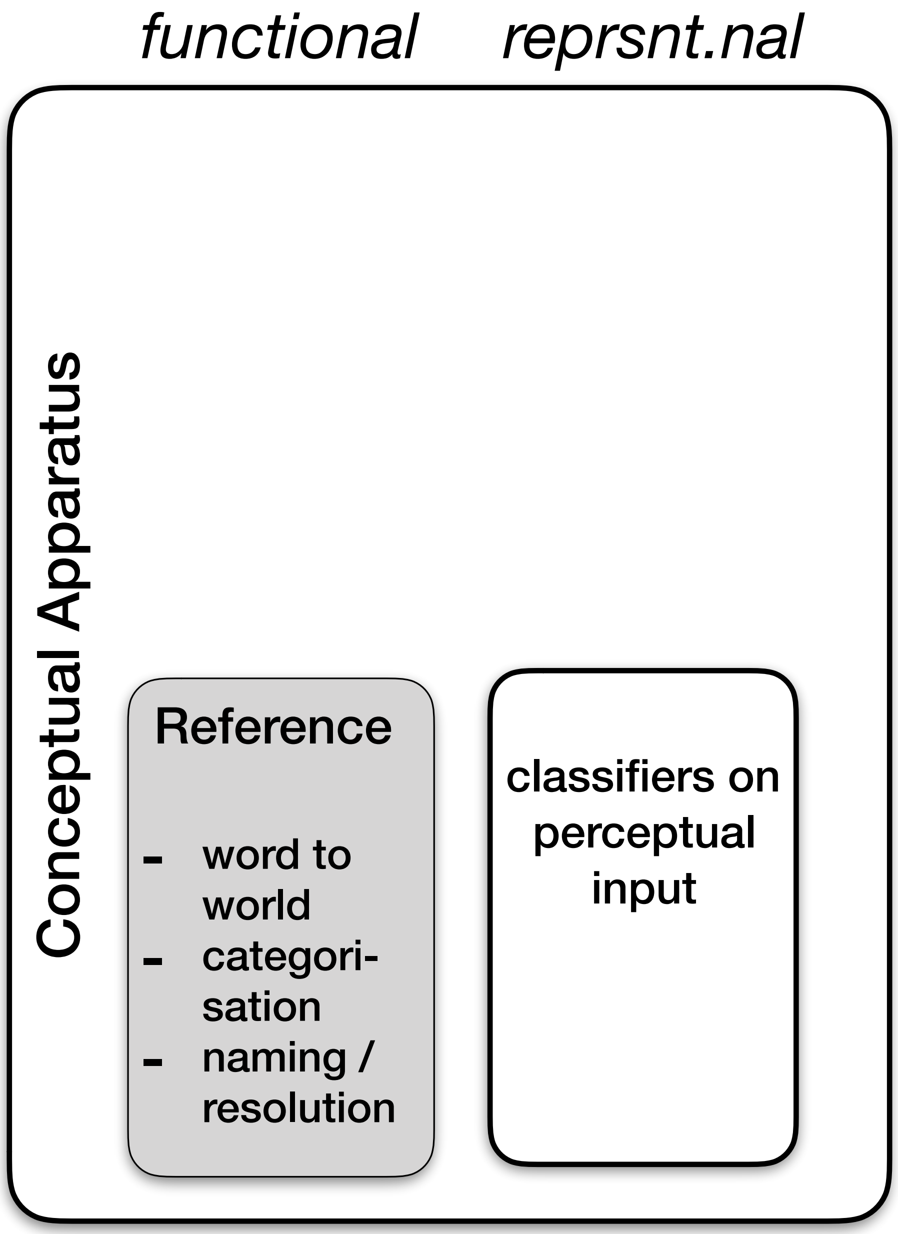
the.x

[[silver]](x) ∧ [[wrench]](x) ∧ [[in]](x) ∧ [[middle]](x) ↦ [0,1]

argmax

some notes

- "*bag of words*" application (for now)
- two things happening at same time:
    - learning of concept
    - name of concept
- theoretical claim (separation inference / reference; sources of information) + implementation (these types of classifiers, this performance)
    - e.g., on relevance of *learnability.* Would make claim that Goodman (1955)-type predicate "grue" (*X is grue if observed before 2019 and blue, or observed after 2019 and blue*) is harder to learn, as it involves two subsystems.



functional *reprsnt.nal*

Conceptual Apparatus

Reference

- word to world
- categori-sation
- naming / resolution

classifiers on perceptual input

$$[[\mathrm{dog}]]^D = \{ (o, f_{dog}(o)) \}$$

what kind of function is this? (what is the range?)

where do we get this function?

how do we present image object to function?

what kind of set is this?

What about the other direction? (Naming, Generation)

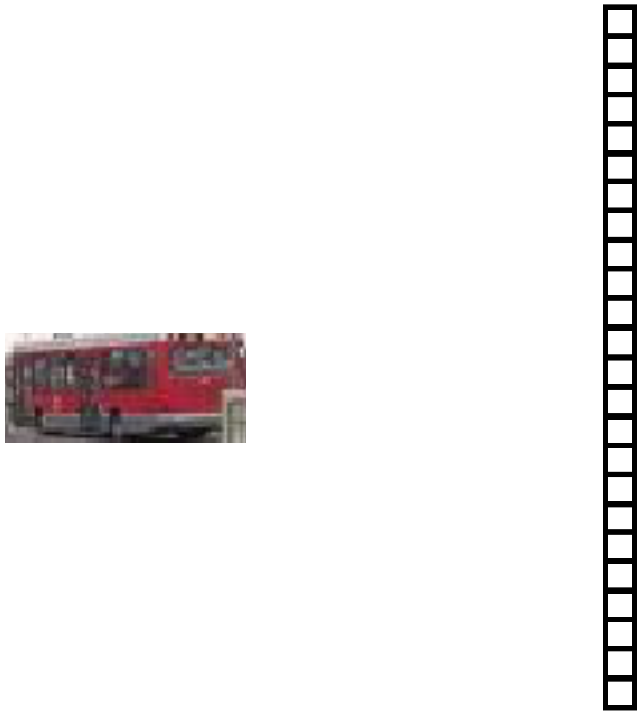
today

- the “words as classifiers” model of referential concepts
    - naming & factor graphs
    - naming colours
- learning word meaning representations from observations of linguistic contexts
    - different kinds of contexts
    - inference: referential compatiblity
    - predicting antonymy
- cross-over: zero-shot learning
- composition: syntax as interface
    - a side note: use visual denotations to induce structure?
    - composing continuous representations for inference: LSTMs, Transformers, Tree-LSTMs
    - composing references

object naming

- given an (image, or drawing of an) object, find its *name* (= noun for its category)
- important paradigm in cognitive psychology (e.g., Rosch 1978, Glaser 1992)
- but also of immediate practical relevance (remember Ann pointing out the dog to Bert)

object naming



run all available classifiers on object, pick the one that gives highest score

or run only subset, if you have a reason to do so?

object naming

in context

object naming

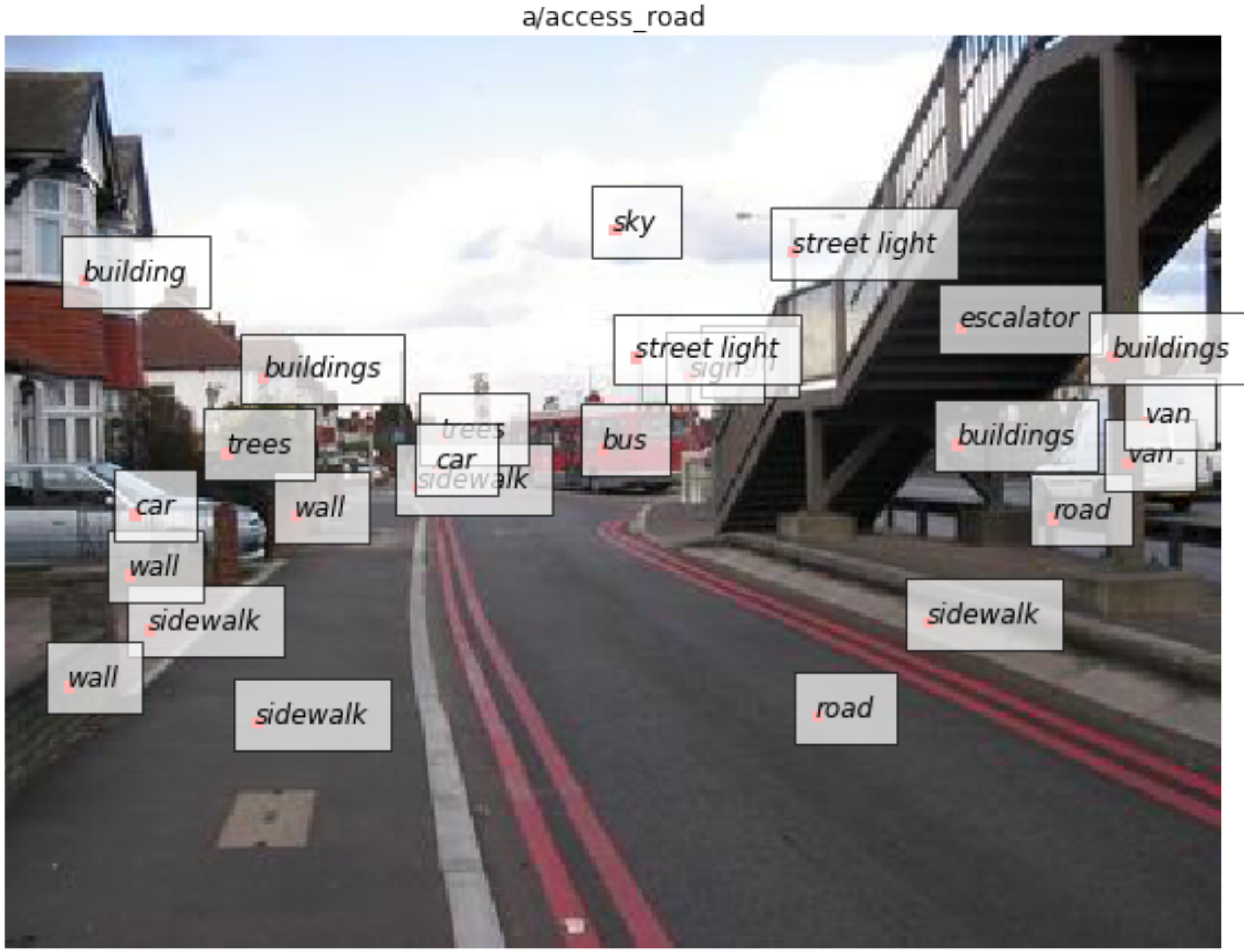
in context

a/access_road

562 WACs trained on ADE20k corpus, object labels

object naming

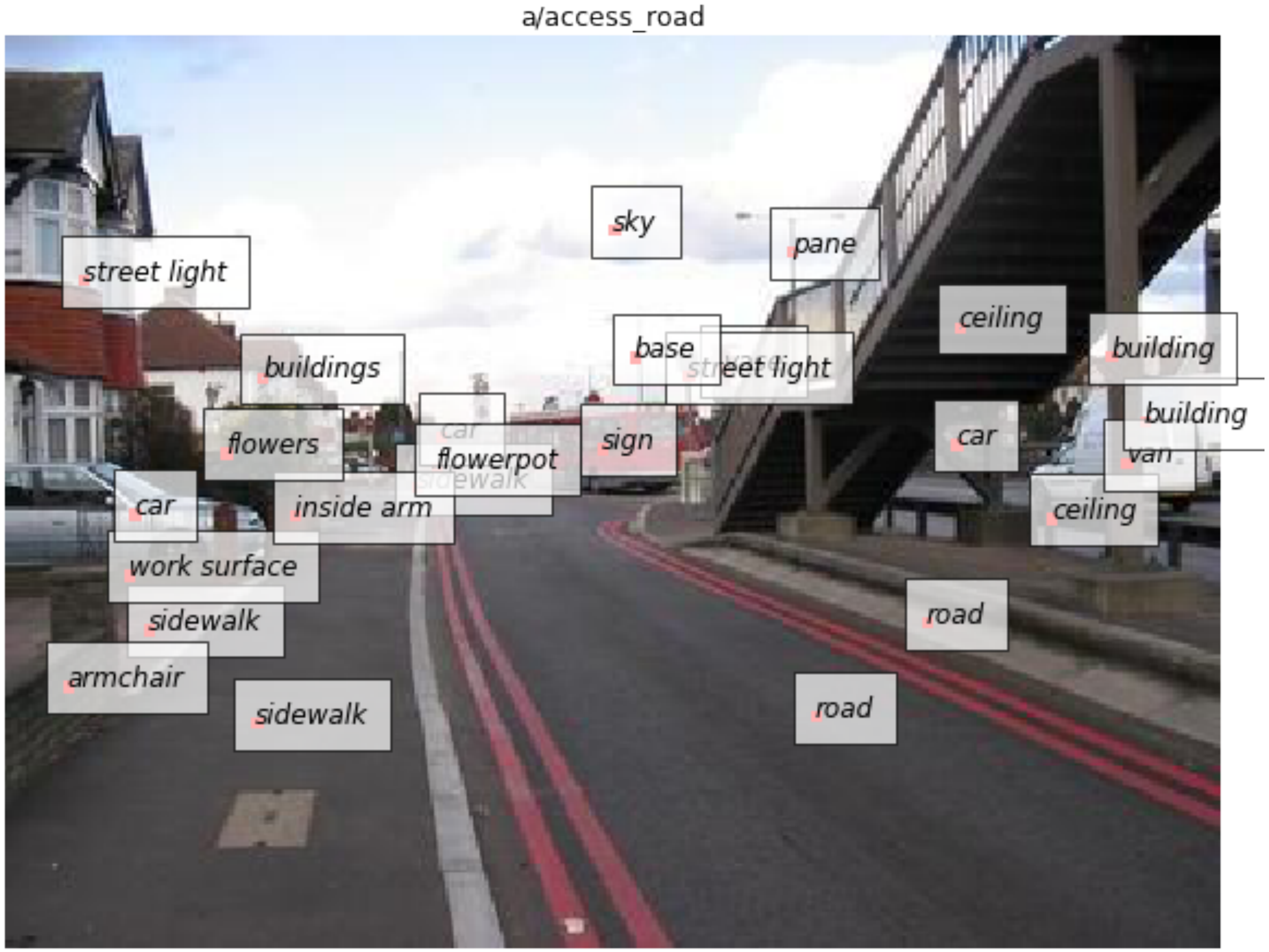
in context



ground truth

object naming

without context

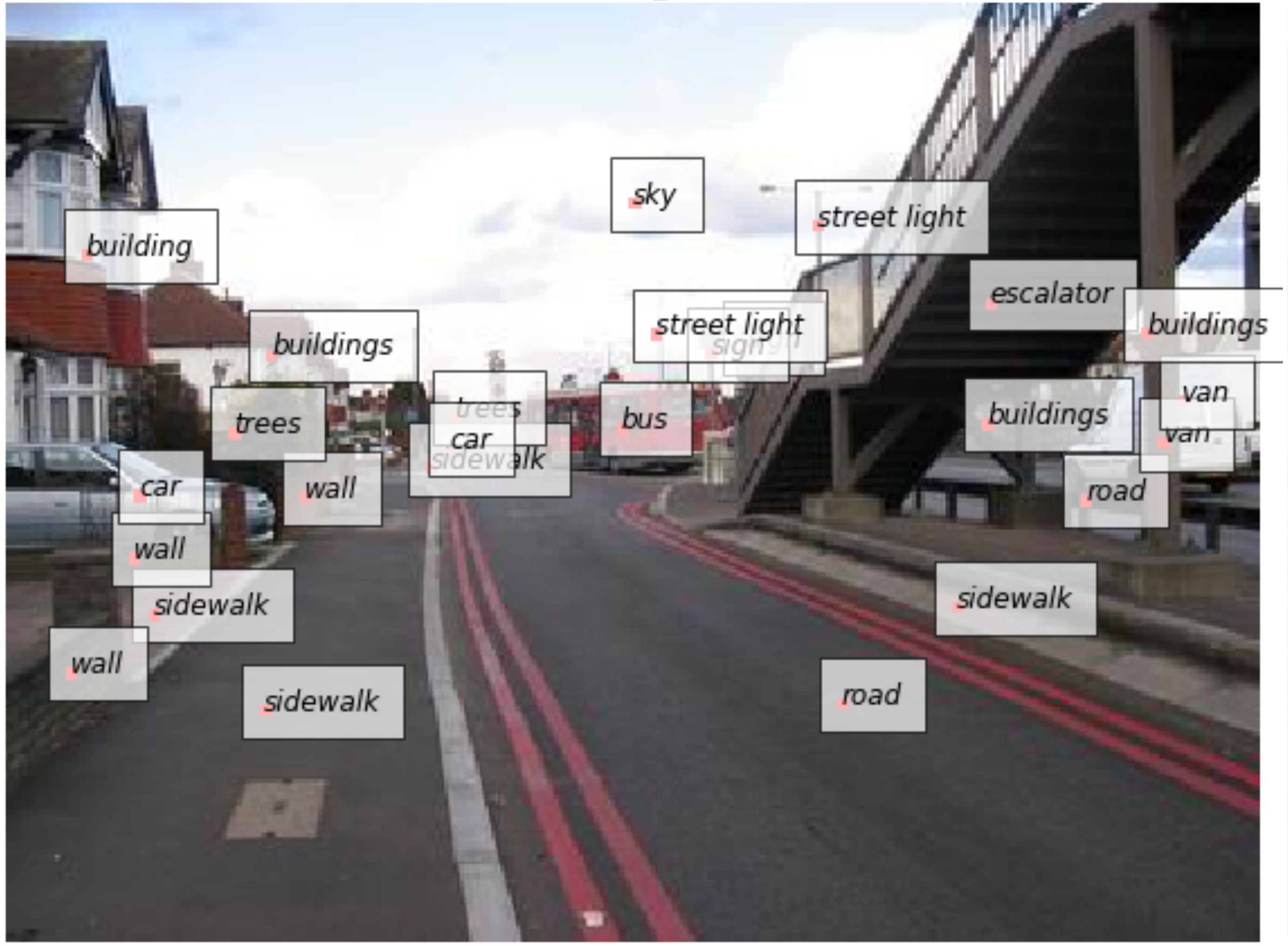


562 WACs trained on ADE20k corpus, object labels

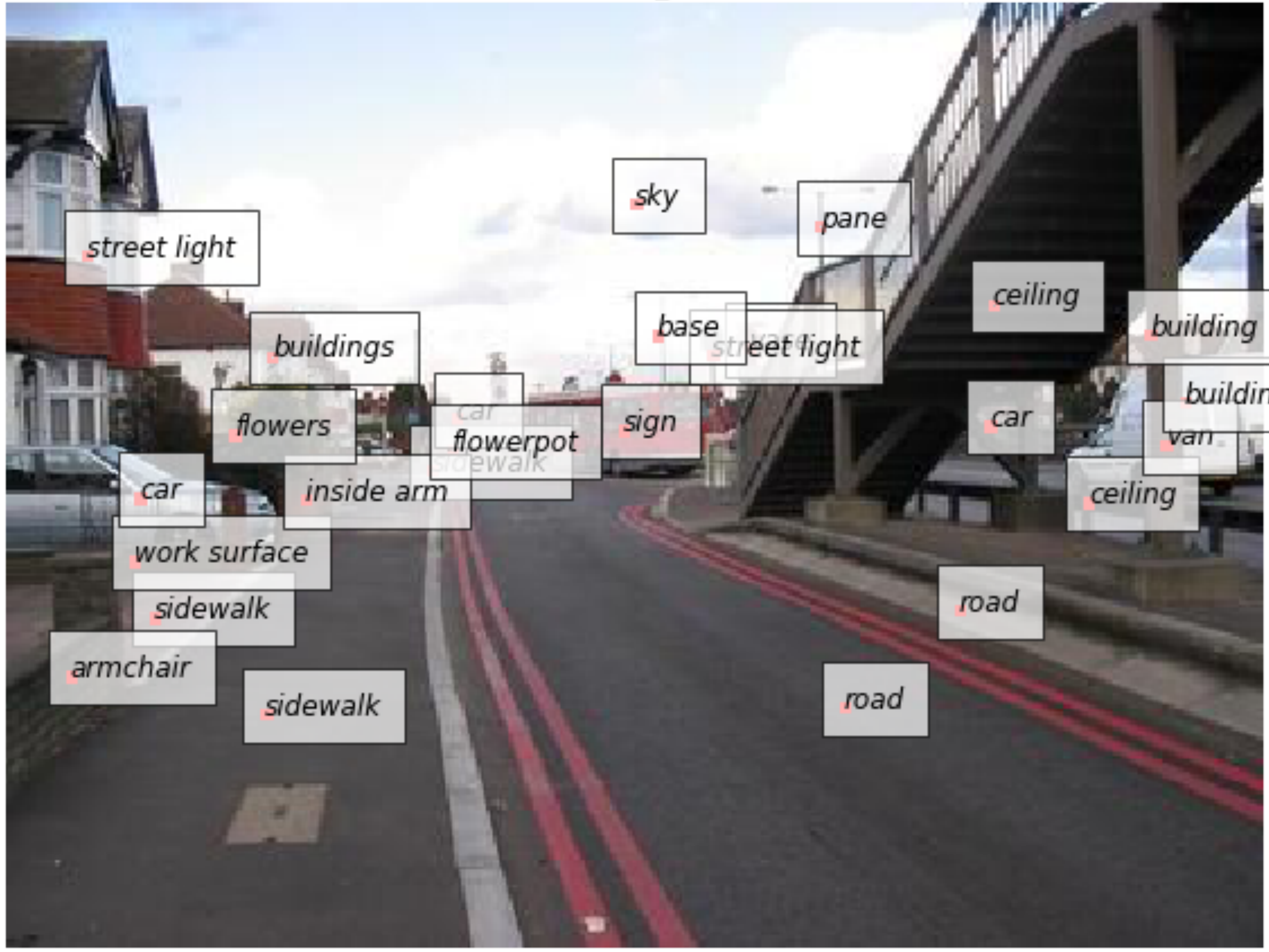
object naming

without context

a/access_road



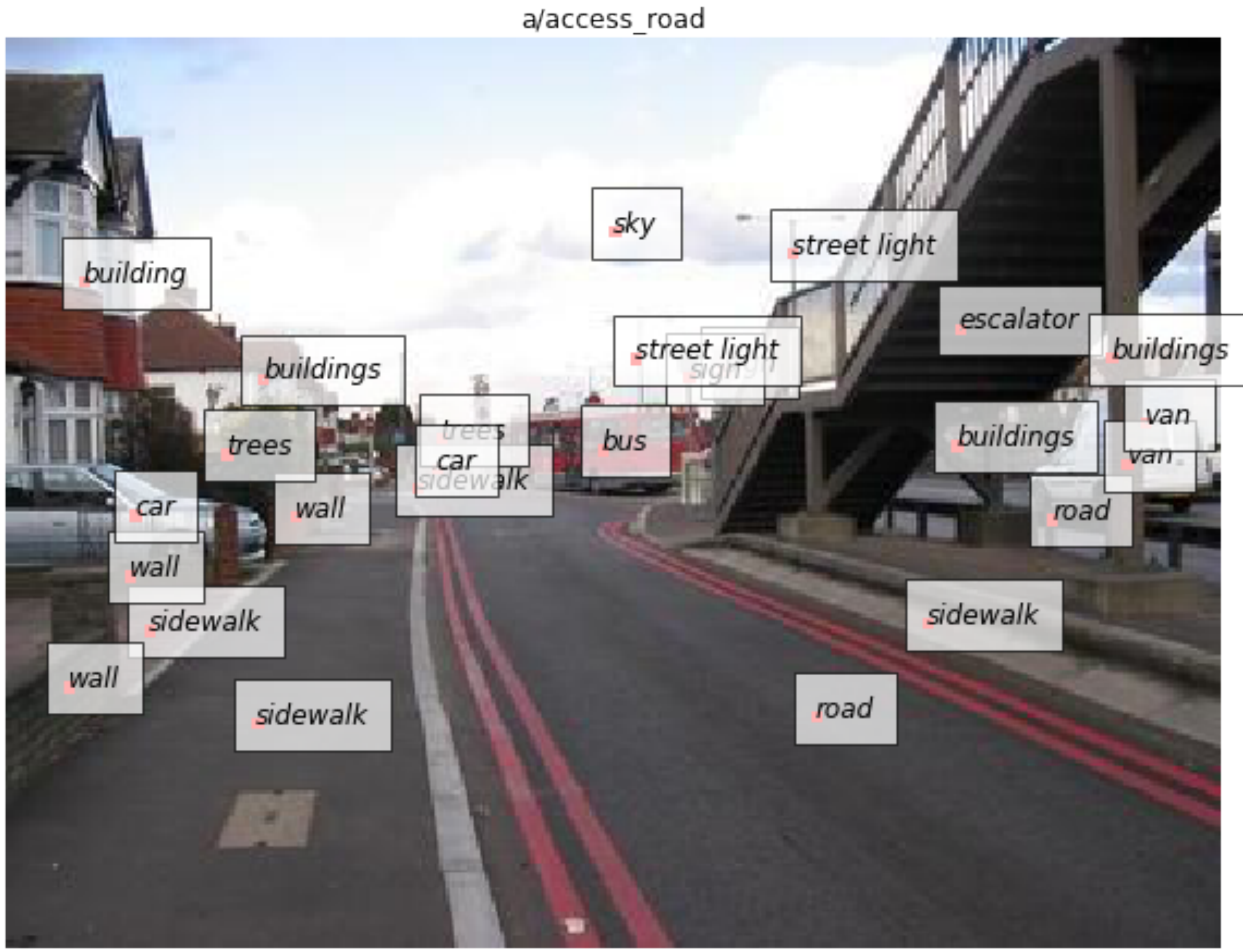
a/access_road



	R@1	R@5	MRR
w/o context	0.30	0.64	0.45
data distr	0.02	0.21	0.12
distr | type	0.12	0.41	0.26

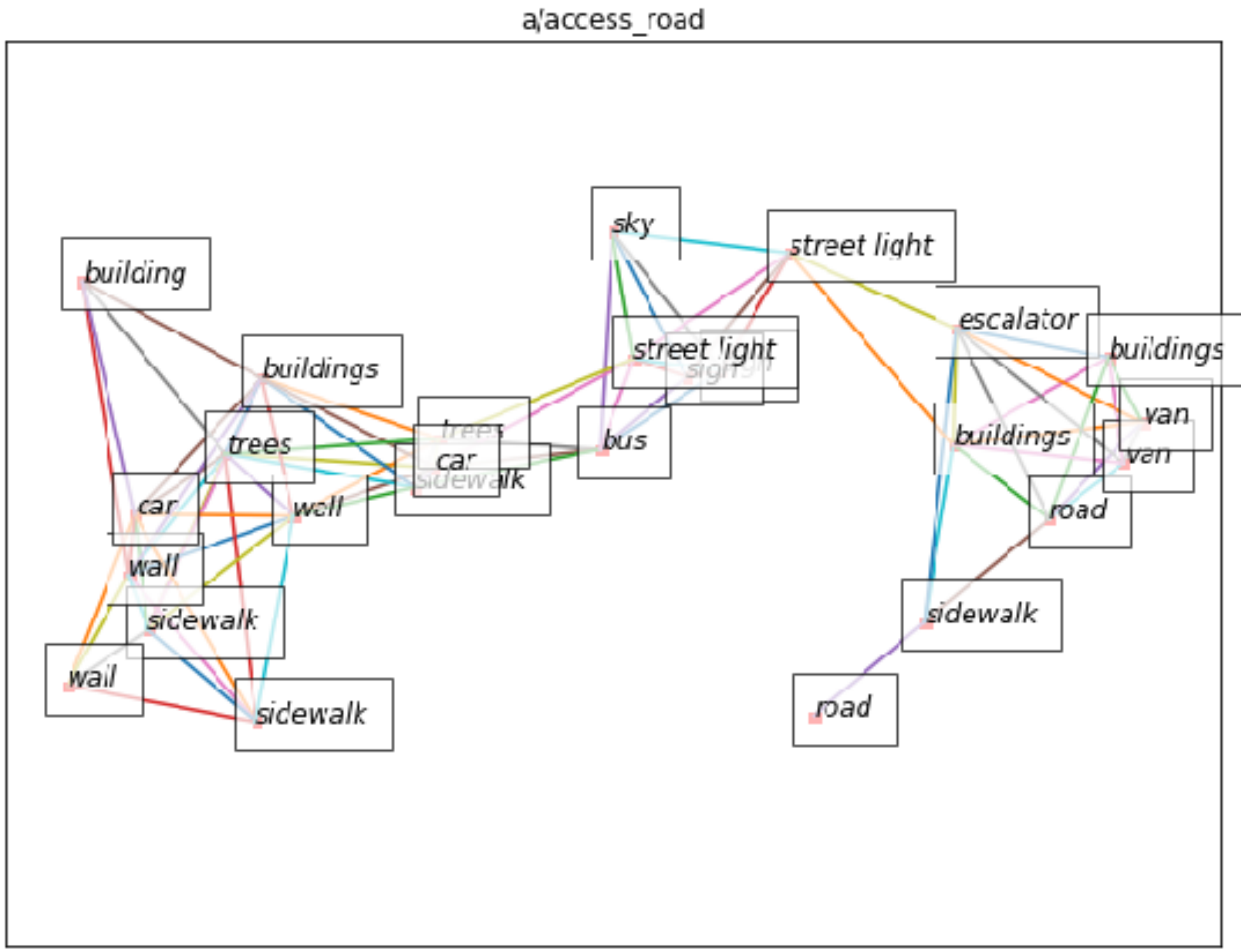
object naming

in context



object naming

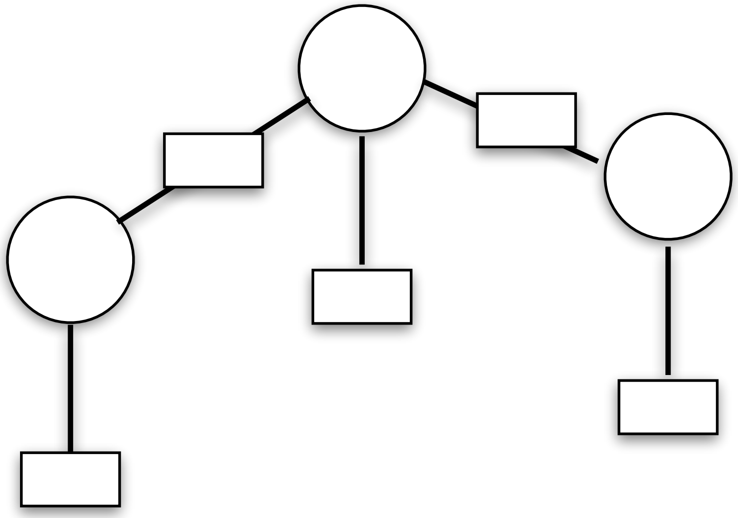
in context



create network of near objects
(within certain radius of each other)

object naming

in context



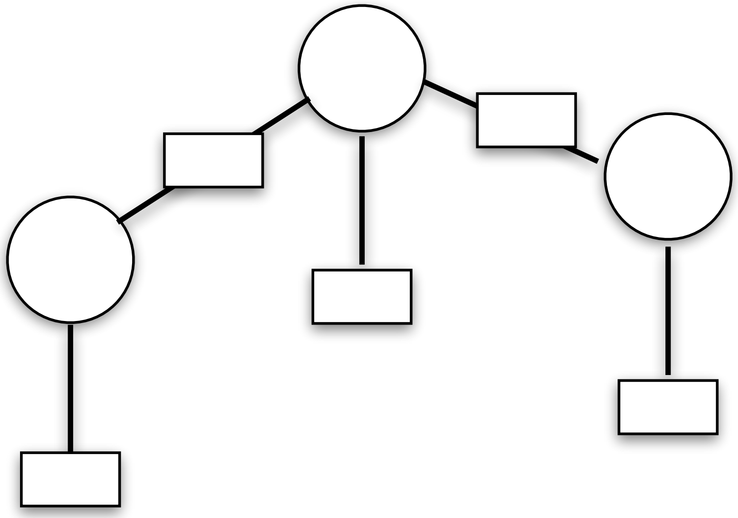
implemented using M Forbes' py-factorgraph package, which implements loopy belief propagation / max-sum inferences. (Forbes & Choi 2017)

https://github.com/mbforbes/py-factorgraph

use this to create factor graph, with the WACs being unary factors, and the co-occurence probabilites as binary factors

object naming

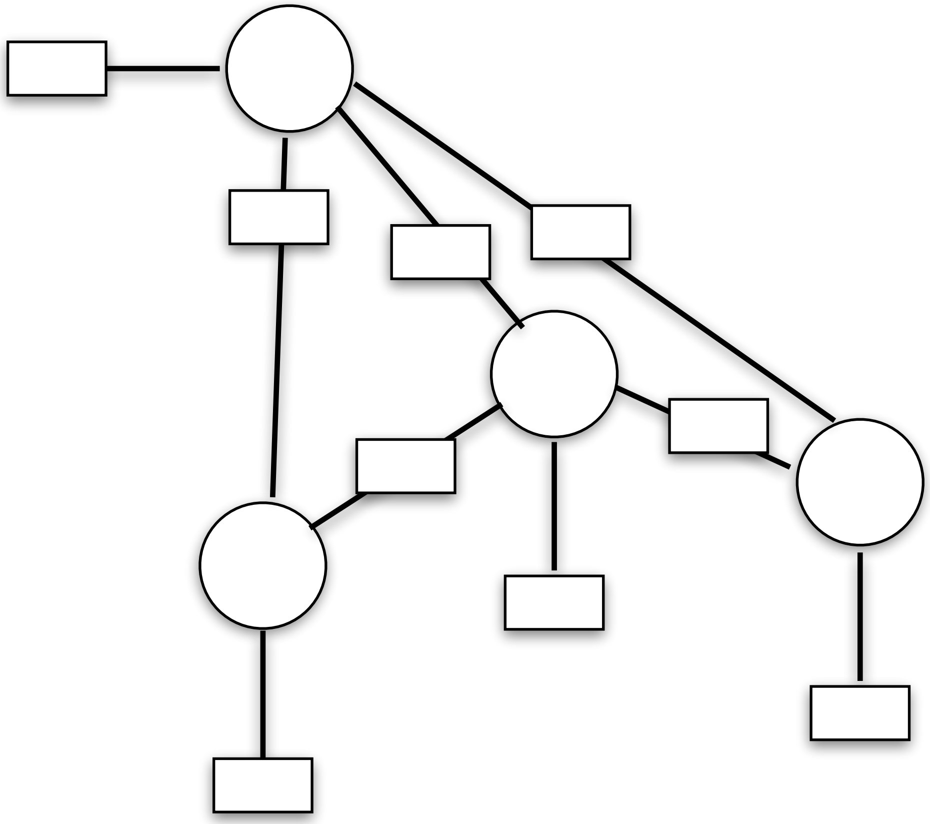
in context



	R@1	R@5	MRR
w/o context	0.30	0.64	0.45
w/ factors	0.41	0.70	0.54

object naming

in context

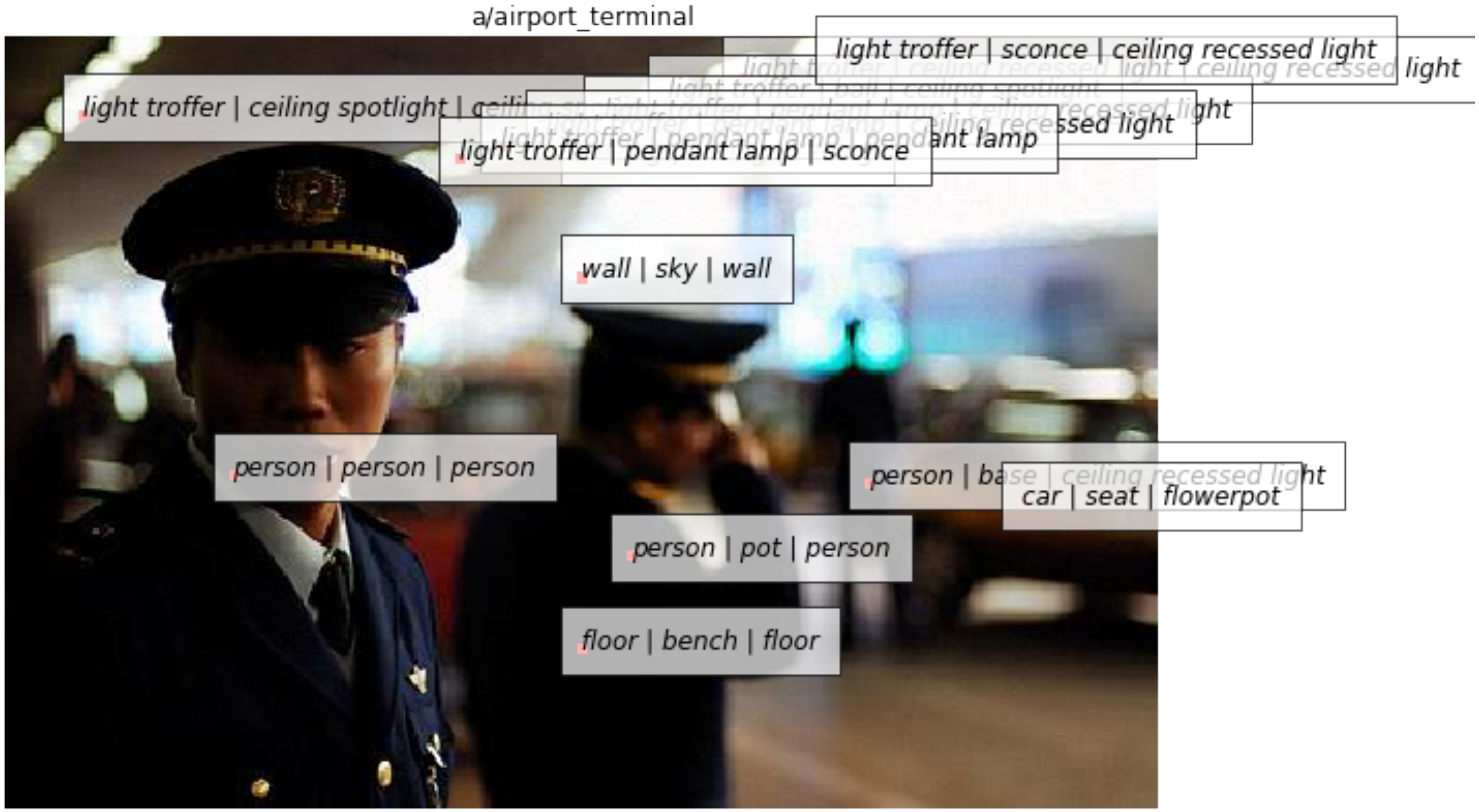


Adding in a scene type node, with a prior distribution as unary factor, and for each objects its conditional prior for this scene type.

	R@1	R@5	MRR
w/o context	0.30	0.64	0.45
w/ factors	0.41	0.70	0.54
w/ stf	0.42	0.71	0.55

object naming

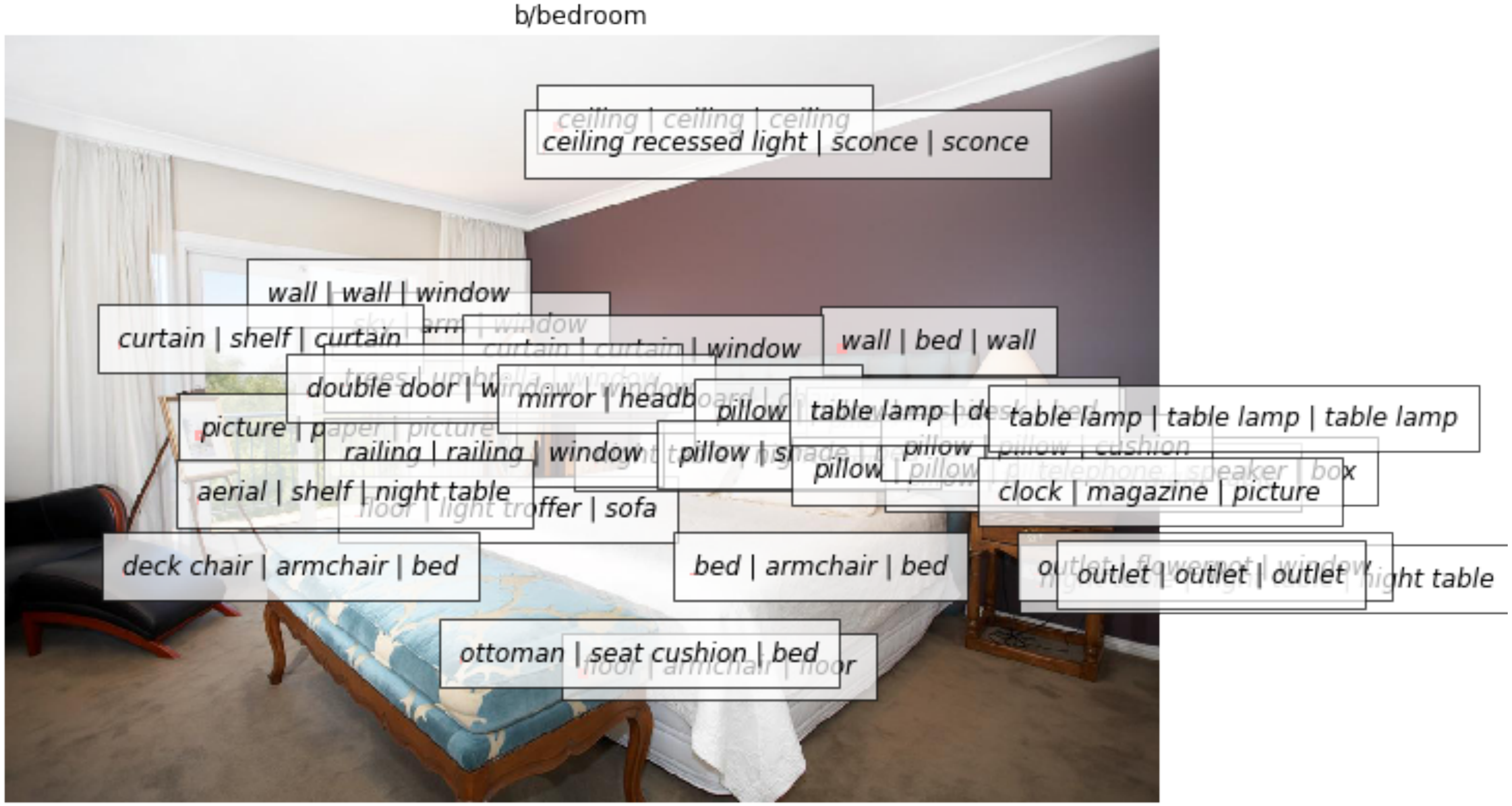
in context


a/airport_terminal
light troffer | sconce | ceiling recessed light
light troffer | ceiling recessed light | ceiling recessed light
light troffer | ball | ceiling spotlight
light troffer | ceiling spotlight | ceiling
light troffer | pendant lamp | ceiling recessed light
light troffer | pendant lamp | pendant lamp
light troffer | pendant lamp | sconce
wall | sky | wall
person | person | person
person | base | ceiling recessed light
car | seat | flowerpot
person | pot | person
floor | bench | floor

	R@1	R@5	MRR
w/o context	0.30	0.64	0.45
w/ factors	0.41	0.70	0.54
w/ stf	0.42	0.71	0.55

object naming

in context

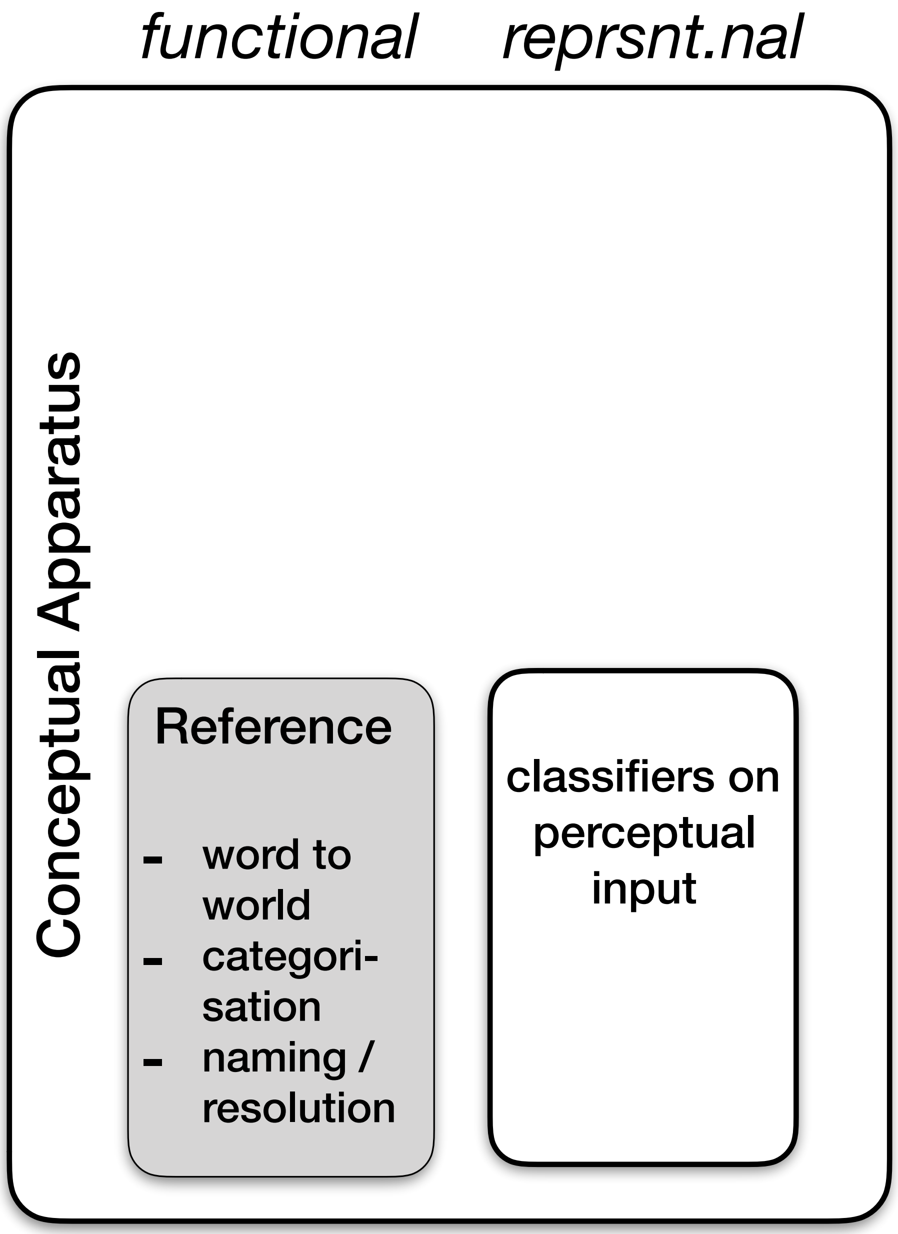

b/bedroom
ceiling | ceiling | ceiling
ceiling recessed light | sconce | sconce
wall | wall | window
sky | arm | window
curtain | shelf | curtain
curtain | curtain | window
wall | bed | wall
trees | umbrella | window
double door | window | window
mirror | headboard
pillow
table lamp | desk | bed
table lamp | table lamp | table lamp
picture | paper | picture
railing | railing | window
pillow | shade
pillow
pillow | cushion
pillow | pillow
telephone | speaker | box
aerial | shelf | night table
clock | magazine | picture
floor | light troffer | sofa
deck chair | armchair | bed
bed | armchair | bed
outlet | flowerpot | window
outlet | outlet | outlet
night table
ottoman | seat cushion | bed
floor | armchair | floor

	R@1	R@5	MRR
w/o context	0.30	0.64	0.45
w/ factors	0.41	0.70	0.54
w/ stf	0.42	0.71	0.55

```
 ('a/art_gallery', 'c/casino/indoor'),
 ('a/art_school', 'l/living_room'),
 ('a/art_studio', 'l/living_room'),
 ('a/art_studio', 'l/living_room'),
('a/artists_loft', 'l/living_room'),
 ('a/attic', 'l/living_room'),
 ('a/attic', 'l/living_room'),
('a/auto_factory', 'l/living_room'),
 ('a/auto_mechanics/outdoor', 'a/
amphitheater'),
('b/bakery/shop', 'l/living_room'),
 ('b/bakery/shop', 'l/living_room'),
 ('b/bakery/shop', 'l/living_room'),
 ('b/balcony/interior', 'l/living_room'),
 ('b/balcony/interior', 'l/living_room'),
('b/bar', 'l/living_room'),
 ('b/bar', 'l/living_room'),
 ('b/bar', 'l/living_room'),
 ('b/barbershop', 'l/living_room'),
('b/bathhouse', 'l/living_room'),
('b/beach', 'b/bedroom'),
 ('b/beach', 's/street'),
 ('b/beach', 's/street'),
 ('b/beauty_salon', 'l/living_room'),
('b/bedroom', 'l/living_room'),
 ('b/bedroom', 'l/living_room'),
 ('b/bedroom', 'l/living_room'),
 ('b/bedroom', 'l/living_room'),
 ('b/bedroom', 'l/living_room'),
 ('b/bedroom', 'l/living_room'),
 ('b/botanical_garden', 'd/dining_room'),
 ('b/bow_window/indoor', 'l/living_room'),
 ('b/bow_window/outdoor', 'd/dining_room'),
 ('b/brewery/outdoor', 'l/living_room'),
 ('b/bridge', 'b/bedroom'),
 ('b/bridge', 'a/amphitheater'),
 ('b/building_facade', 'b/bedroom'),
 ('b/building_facade', 'b/bedroom'),
 ('b/building_facade', 'l/living_room'),
 ('b/building_facade', 'b/bedroom'),
 ('b/building_facade', 'l/living_room'),
 ('b/building_facade', 'l/living_room'),
 ('b/building_facade', 'l/living_room'),
 ('b/burial_chamber', 'l/living_room'),
 ('b/bus_interior', 'l/living_room'),
 ('b/bus_interior', 'a/amphitheater'),
 ('b/bus_station/outdoor', 'l/living_room'),
 ('c/cabin/outdoor', 'b/bedroom'),
 ('c/campus', 'b/bedroom'),
 ('c/canal/natural', 'b/bedroom'),
 ('c/candy_store', 'l/living_room'),
 ('c/canyon', 's/street'),
 ('c/car_interior/backseat', 'l/
living_room'),
 ('c/caravansary', 'l/living_room'),
 ('c/cargo_deck/airplane', 'c/casino/
indoor'),
 ('c/carport/outdoor', 'a/amphitheater'),
 ('c/casino/indoor', 'l/living_room'),
 ('c/casino/indoor', 'l/living_room'),
 ('c/castle', 'b/bedroom'),
 ('c/castle', 'b/bedroom'),
 ('c/catacomb', 'l/living_room'),
 ('c/cathedral/indoor', 'l/living_room'),
 ('c/catwalk', 'l/living_room'),
 ('c/cavern/outdoor', 'b/bedroom'),
 ('c/cemetery', 'a/amphitheater'),
 ('c/chapel', 'l/living_room'),
 ('c/cheese_factory', 'l/living_room'),
 ('c/chicken_coop/indoor', 'l/living_room'),
 ('d/dinette/vehicle', 'l/living_room'),
 ('d/dinette/vehicle', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dining_room', 'l/living_room'),
 ('d/dirt_track', 'l/living_room'),
 ('d/dock', 'b/bedroom'),
 ('d/donjon', 'b/bedroom'),
 ('d/doorway/indoor', 'l/living_room'),
 ('d/doorway/outdoor', 'l/living_room'),
 ('d/dorm_room', 'l/living_room'),
 ('d/dorm_room', 'l/living_room'),
 ('d/dorm_room', 'l/living_room'),
 ('d/drainage_ditch', 'b/bedroom'),
 ('d/drill_rig', 'l/living_room'),
 ('d/driveway', 'd/dining_room'),
 ('e/earth_fissure', 'b/bedroom'),
 ('e/elevator/door', 'l/living_room'),
 ('e/elevator/interior', 'l/living_room'),
 ('e/elevator_shaft', 'l/living_room'),
 ('e/engine_room', 'l/living_room'),
 ('e/estuary', 's/street'),
 ('f/factory/indoor', 'l/living_room'),
 ('f/field/cultivated', 'b/bedroom'),
 ('f/field/wild', 's/street'),
 ('f/field/wild', 's/street'),
 ('f/fire_escape', 'l/living_room'),
 ('f/firing_range/indoor', 'l/living_room'),
 ('f/fishmarket', 'l/living_room'),
 ('f/fitting_room/interior', 'l/
living_room'),
 ('f/flight_of_stairs/urban', 'b/bedroom'),
 ('f/florist_shop/indoor', 'c/casino/
indoor'),
 ('f/florist_shop/outdoor', 'l/living_room'),
 ('f/football_field', 'b/bedroom'),
 ('f/forest/broadleaf', 's/street'),
 ('f/forest/broadleaf', 'd/dining_room'),
 ('f/forest/broadleaf', 'b/bedroom'),
 ('f/forest/needleleaf', 'b/bedroom'),
 ('f/forest/needleleaf', 's/street'),
 ('f/forest_path', 'b/bedroom'),
 ('f/forest_road', 'b/bedroom'),
 ('f/fort', 'b/bedroom'),
 ('f/foundry/outdoor', 'b/bedroom'),
 ('f/freeway', 'a/amphitheater'),
 ('f/furnace_room', 'l/living_room'),
 ('g/galley', 'l/living_room'),
 ('g/game_room', 'l/living_room'),
 ('g/game_room', 'l/living_room'),
 ('g/game_room', 'l/living_room'),
 ('g/game_room', 'c/casino/indoor'),
 ('g/garage/indoor', 'l/living_room'),
 ('g/garage/indoor', 'l/living_room'),
 ('g/garage/indoor', 'l/living_room'),
 ('g/garage/outdoor', 'b/bedroom'),
 ('g/gatehouse', 'b/bedroom'),
 ('g/geodesic_dome/indoor', 'l/living_room'),
 ('g/ghost_town', 'l/living_room'),
 ('g/glacier', 's/street'),
 ('g/golf_course', 'd/dining_room'),
 ('g/gorge', 's/street'),
 ('g/great_hall', 'l/living_room'),
 ('i/ice_skating_rink/outdoor', 'l/
living_room'),
 ('i/igloo', 'b/bedroom'),
 ('i/industrial_area', 'l/living_room'),
 ('i/inn/indoor', 'l/living_room'),
 ('i/inn/outdoor', 'b/bedroom'),
 ('i/islet', 'l/lobby'),
 ('j/jacuzzi/indoor', 'l/living_room'),
 ('j/jacuzzi/indoor', 'l/living_room'),
 ('j/jail/indoor', 'l/living_room'),
 ('j/jail_cell', 'l/living_room'),
 ('j/japanese_garden', 'a/amphitheater'),
 ('j/joss_house', 'l/living_room'),
 ('j/junkyard', 'b/bedroom'),
 ('k/kennel/outdoor', 'a/amphitheater'),
 ('k/kindergarden_classroom', 'l/
living_room'),
 ('k/kiosk/outdoor', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('k/kitchen', 'l/living_room'),
 ('l/lab_classroom', 'c/casino/indoor'),
 ('l/labyrinth/outdoor', 'b/bedroom'),
 ('l/lake/natural', 'a/amphitheater'),
 ('l/lake/natural', 'b/bedroom'),
 ('l/landing_deck', 'l/living_room'),
 ('l/lawn', 'b/bedroom'),
 ('l/lecture_room', 'l/living_room'),
 ('l/levee', 'a/amphitheater'),
 ('l/lido_deck/outdoor', 'l/living_room'),
 ('l/lighthouse', 'a/amphitheater'),
 ('l/lighthouse', 'a/amphitheater'),
 ('l/liquor_store/outdoor', 'b/bedroom'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/living_room', 'l/living_room'),
 ('l/lobby', 'l/living_room'),
 ('l/lobby', 'l/living_room'),
 ('l/locker_room', 'l/living_room'),
 ('l/lookout_station/outdoor', 'b/bedroom'),
 ('m/mansion', 'b/bedroom'),
 ('m/manufactured_home', 'b/bedroom'),
 ('m/marsh', 'l/lobby'),
 ('m/martial_arts_gym', 'l/living_room'),
 ('m/medina', 'b/bedroom'),
 ('m/military_hut', 'l/living_room'),
 ('m/mission', 'a/amphitheater'),
 ('p/planetarium/outdoor', 'b/bedroom'),
 ('p/playground', 'a/amphitheater'),
 ('p/playroom', 'l/living_room'),
 ('p/playroom', 'l/living_room'),
 ('p/plaza', 'b/bedroom'),
 ('p/podium/outdoor', 'l/living_room'),
 ('p/poolroom/establishment', 'l/
living_room'),
 ('p/poolroom/home', 'l/living_room'),
 ('p/poolroom/home', 'l/living_room'),
 ('p/poolroom/home', 'l/living_room'),
 ('p/poolroom/home', 'l/living_room'),
 ('p/promenade_deck', 'l/living_room'),
 ('p/pulpit', 'l/living_room'),
 ('q/quadrangle', 'l/living_room'),
 ('r/ramp', 'b/bedroom'),
 ('r/ranch_house', 'd/dining_room'),
 ('r/reception', 'l/living_room'),
 ('r/reception', 'l/living_room'),
 ('r/resort', 'b/bedroom'),
 ('r/restaurant_patio', 'c/casino/indoor'),
 ('r/revolving_door', 'l/living_room'),
 ('r/river', 'a/amphitheater'),
 ('r/river', 's/street'),
 ('r/rock_arch', 's/street'),
 ('r/root_cellar', 'l/living_room'),
 ('r/roundabout', 'a/amphitheater'),
 ('r/roundabout', 'a/amphitheater'),
 ('r/ruin', 's/street'),
 ('r/runway', 'b/bedroom'),
 ('s/sand_trap', 'b/bedroom'),
 ('s/sandbox', 'l/living_room'),
 ('s/sauna', 'l/living_room'),
 ('s/schoolhouse', 'b/bedroom'),
 ('s/seawall', 'l/living_room'),
 ('s/shipping_room', 'l/living_room'),
 ('s/shoe_shop', 'l/living_room'),
 ('s/shower', 'l/living_room'),
 ('s/shrine', 'l/living_room'),
 ('s/ski_lodge', 'a/amphitheater'),
 ('s/ski_slope', 'l/living_room'),
 ('s/skyscraper', 'b/bedroom'),
 ('s/skyscraper', 'b/bedroom'),
 ('s/skyscraper', 'a/amphitheater'),
 ('s/skyscraper', 'l/living_room'),
 ('s/skyscraper', 'a/amphitheater'),
 ('s/skyscraper', 'b/bedroom'),
 ('s/skyscraper', 'b/bedroom'),
 ('s/skyscraper', 'l/living_room'),
 ('s/skyscraper', 'b/bedroom'),
 ('s/skyscraper', 'l/living_room'),
 ('s/skyscraper', 'b/bedroom'),
 ('s/skyscraper', 's/street'),
 ('s/skyscraper', 's/street'),
 ('s/slum', 'l/living_room'),
 ('s/snowfield', 'b/bedroom'),
 ('s/spillway', 's/street'),
 ('s/stadium/baseball', 'l/living_room'),
 ('s/stage/indoor', 'l/living_room'),
 ('s/staircase', 'l/living_room'),
 ('s/staircase', 'l/living_room'),
 ('s/staircase', 'l/living_room'),
 ('s/steel_mill/outdoor', 'b/bedroom'),
 ('s/street', 'a/amphitheater'),
 ('s/street', 'a/amphitheater'),
 ('s/street', 'b/bedroom'),
 ('s/street', 'a/amphitheater'),
 ('s/street', 'a/amphitheater'),
 ('s/street', 'a/amphitheater'),
 ('s/street', 'a/amphitheater'),
```

open questions

- what kind of knowledge is this?
- knowledge about the world, about which types of things tend to occur together?
- or is this part of the knowledge of the *word* “bathroom”?



functional *reprsnt.nal*

Conceptual Apparatus

Reference

- word to world
- categorisation
- naming / resolution

classifiers on perceptual input

$$[[\mathrm{dog}]]^{D} = \{ (o, f_{dog}(o)) \}$$

what kind of function is this? (what is the range?)

where do we get this function?

how do we present image object to function?

what kind of set is this?

What about the other direction? (Naming, Generation)

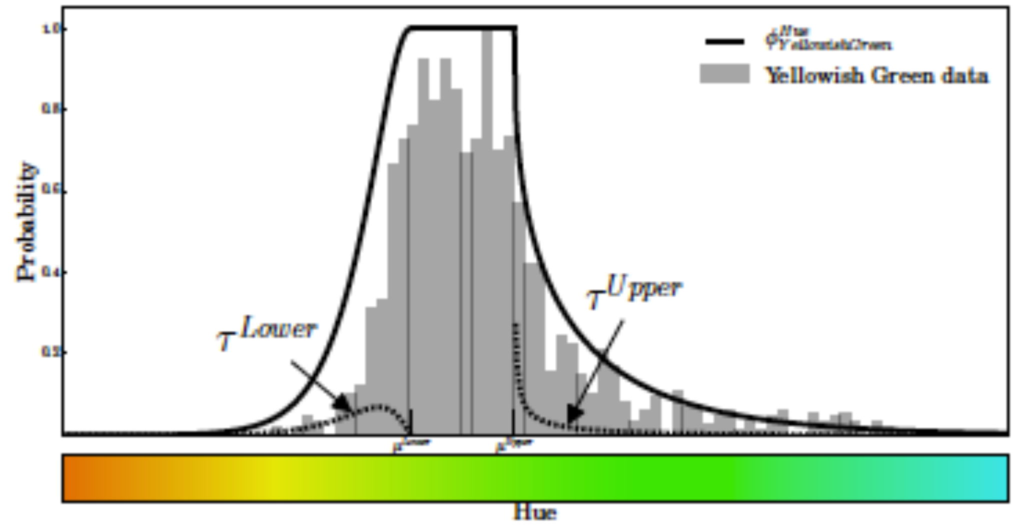
Kind of works, if top-down constraints are factored in.

another example: colour terms

- (Zarrieß & Schlangen, INLG 2016), “Towards Generating Colour Terms for Referents in Photographs: Prefer the Expected or the Unexpected?”

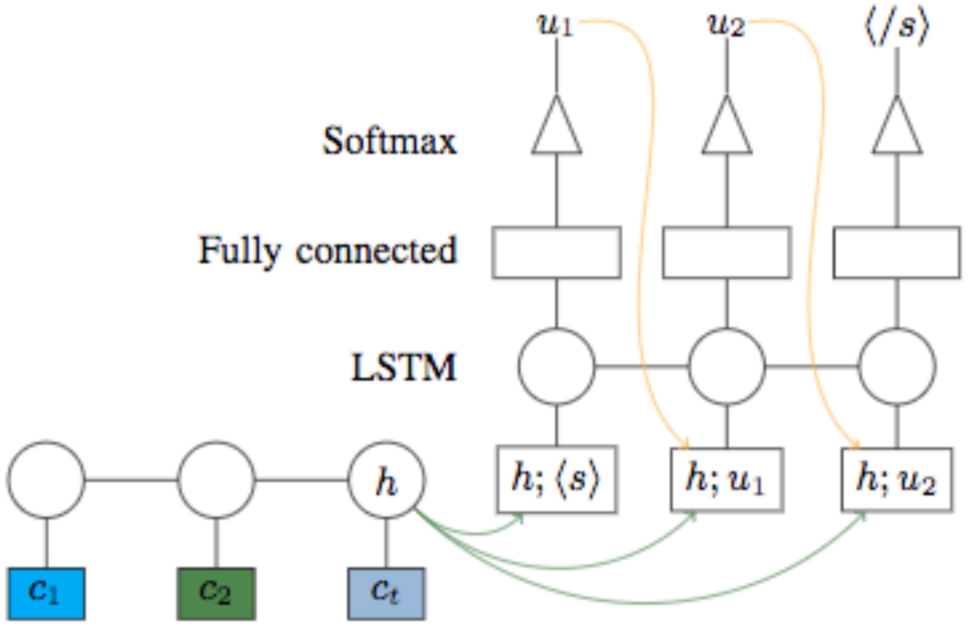
(Now JProf @ Jena)

Example: Models for Colour Term Meaning



yellowish green

(MacMahan and Stone, TACL 2015)



(Monroe et al, TACL 2017)

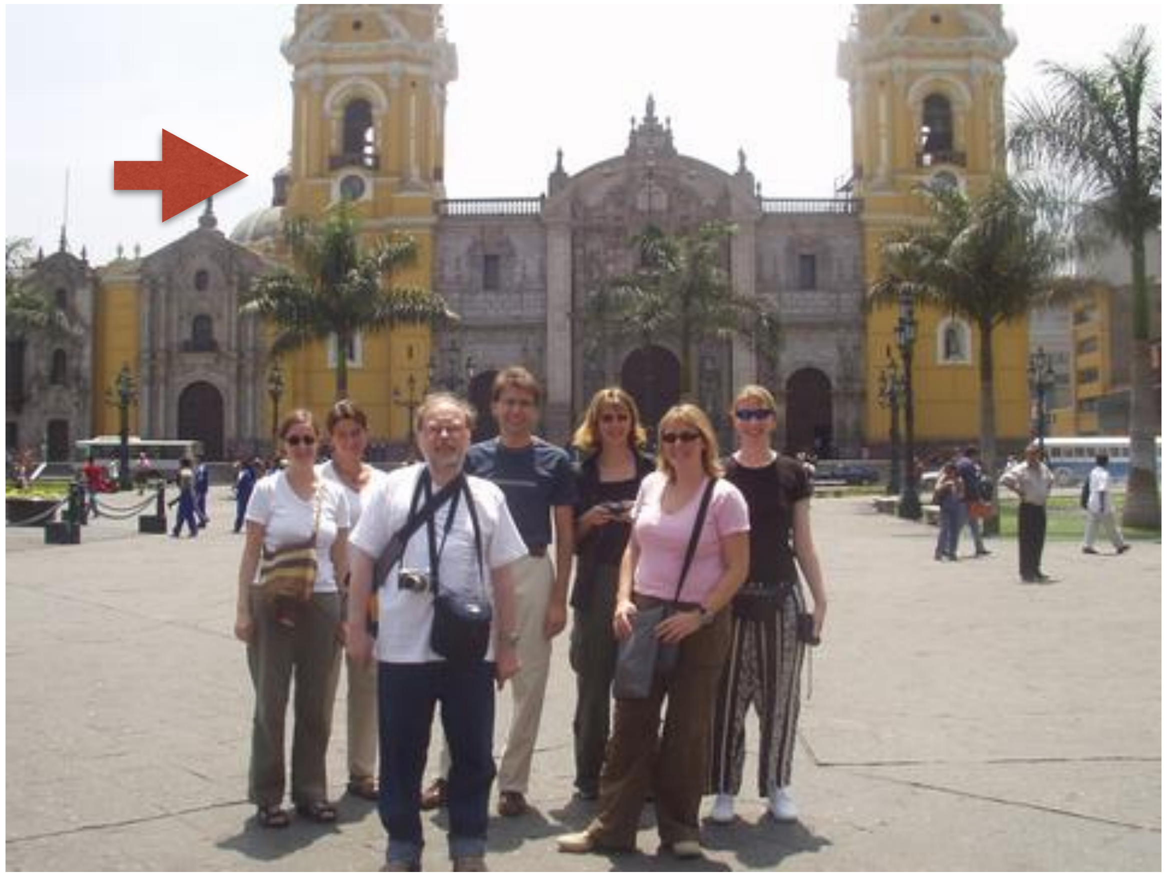
The Task

- predict colour terms for objects in photographs

- data: referring expressions paired with image regions

- content decision is known, focus on REs with colour terms (12K pairs)

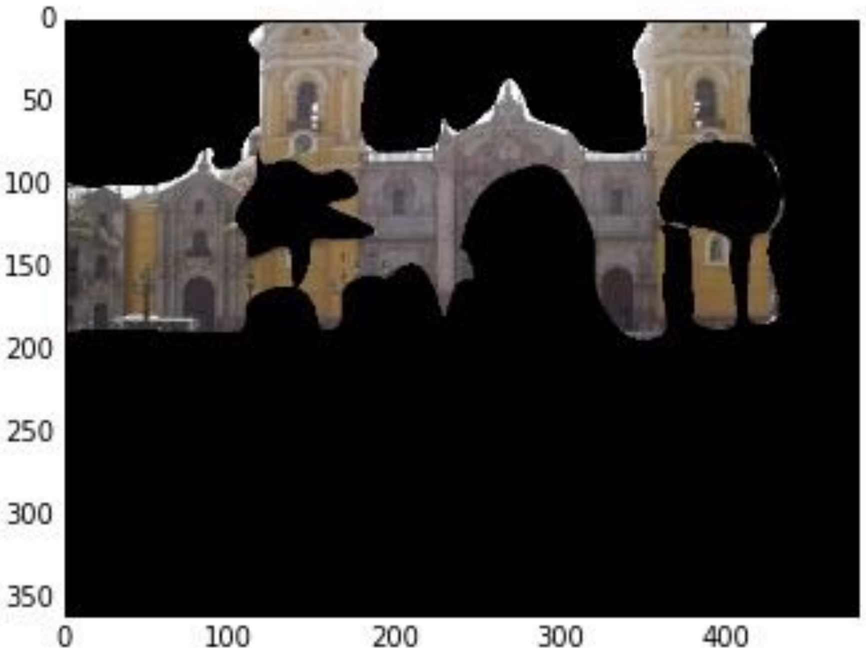
- (objects are labelled)



*the **yellow** building*

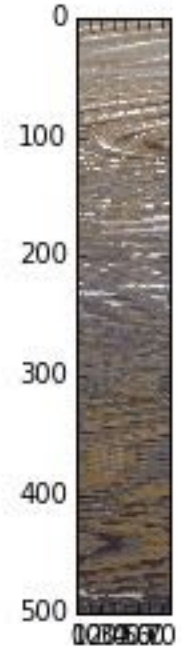
ReferIt Corpus (Kazemzadeh et al., 2014)

From Colour Patches to Realistic Inputs

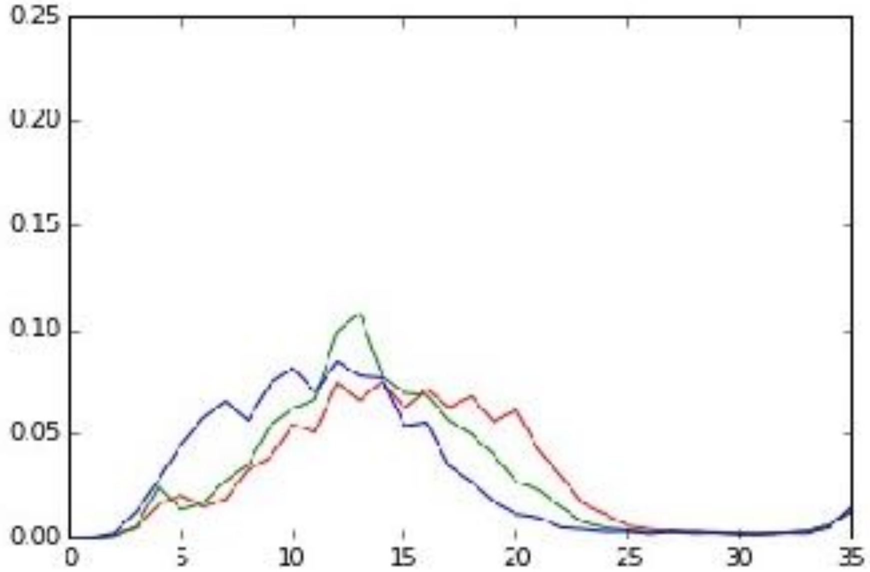


region,
high-quality
segmentation!

„bag-of-pixels“



RGB histogram

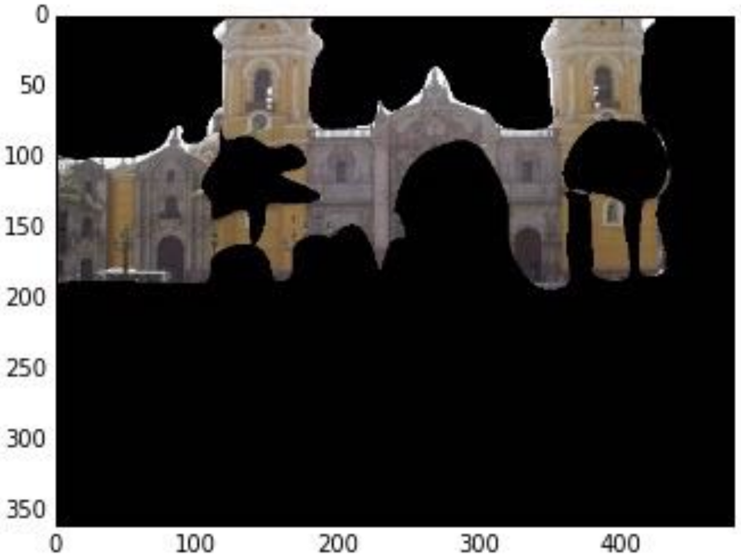


yellow?

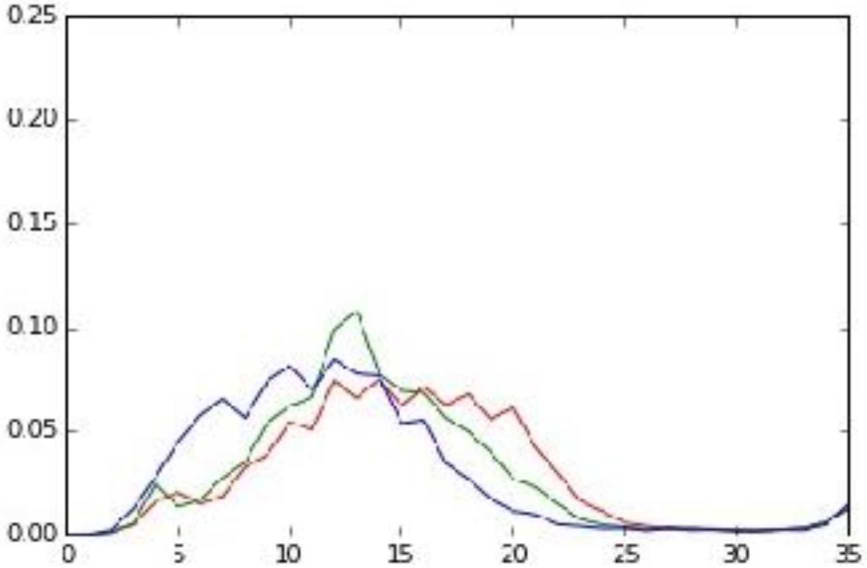
gray?

brown?

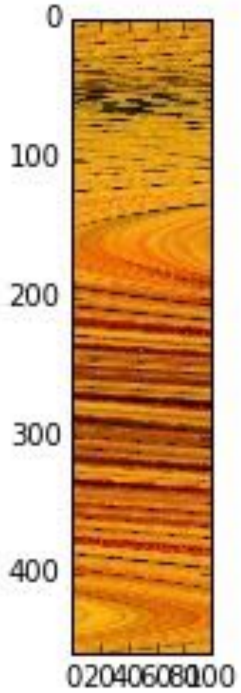
white?

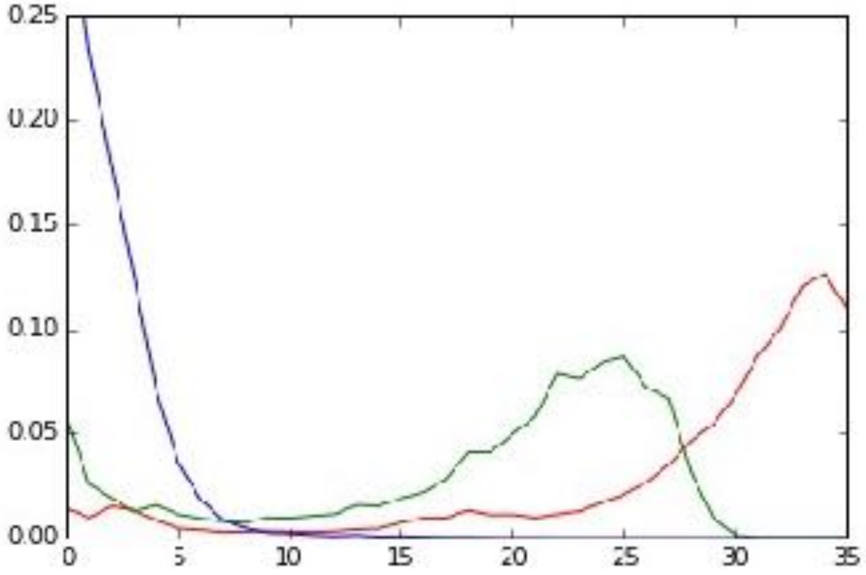


yellow building



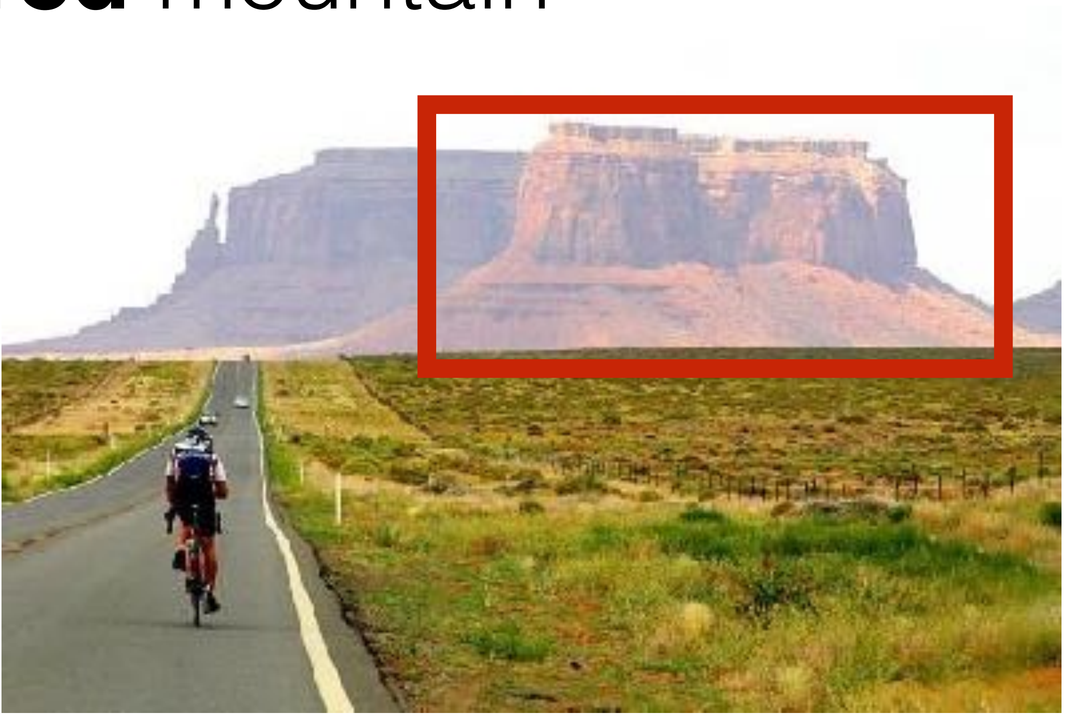
yellow flower

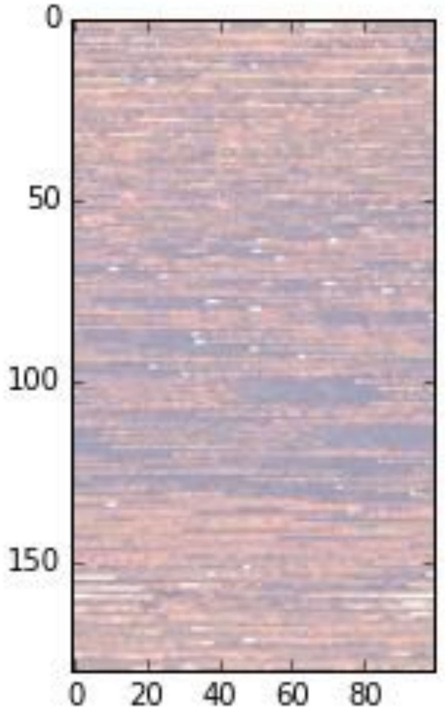




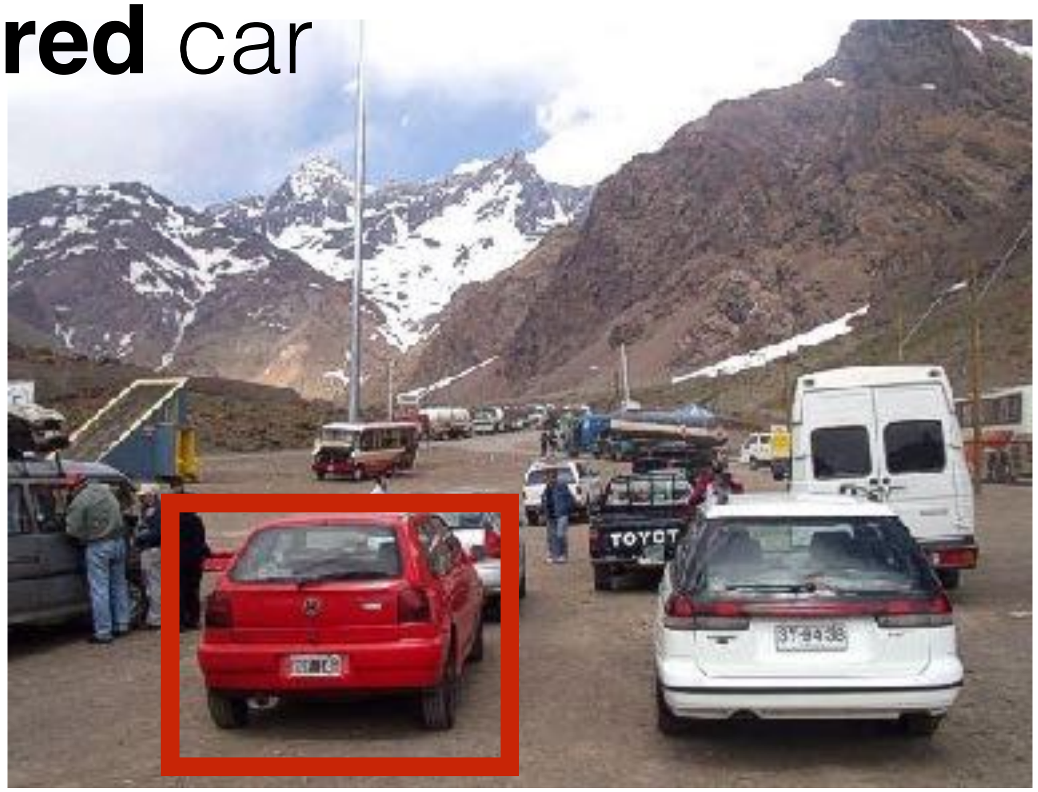
Variation Among Objects

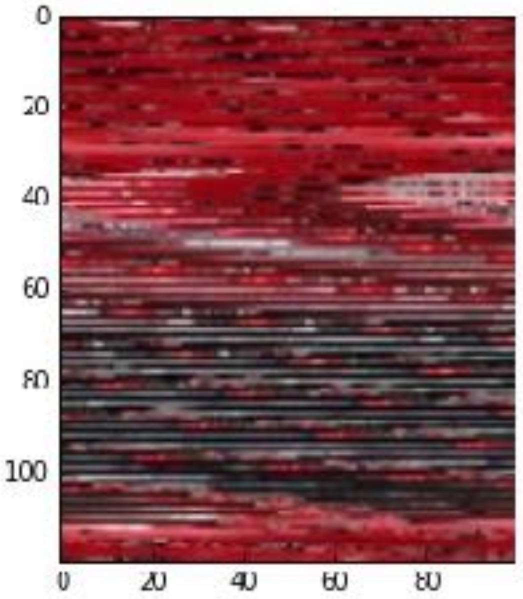
red mountain





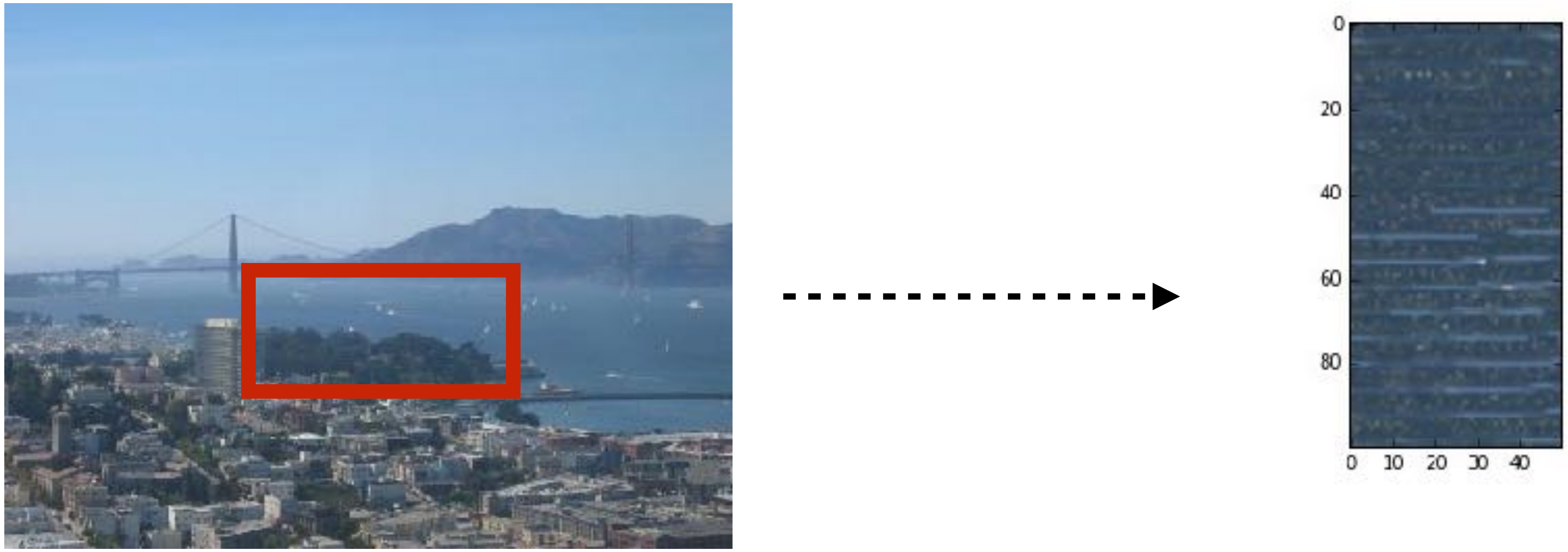
red car



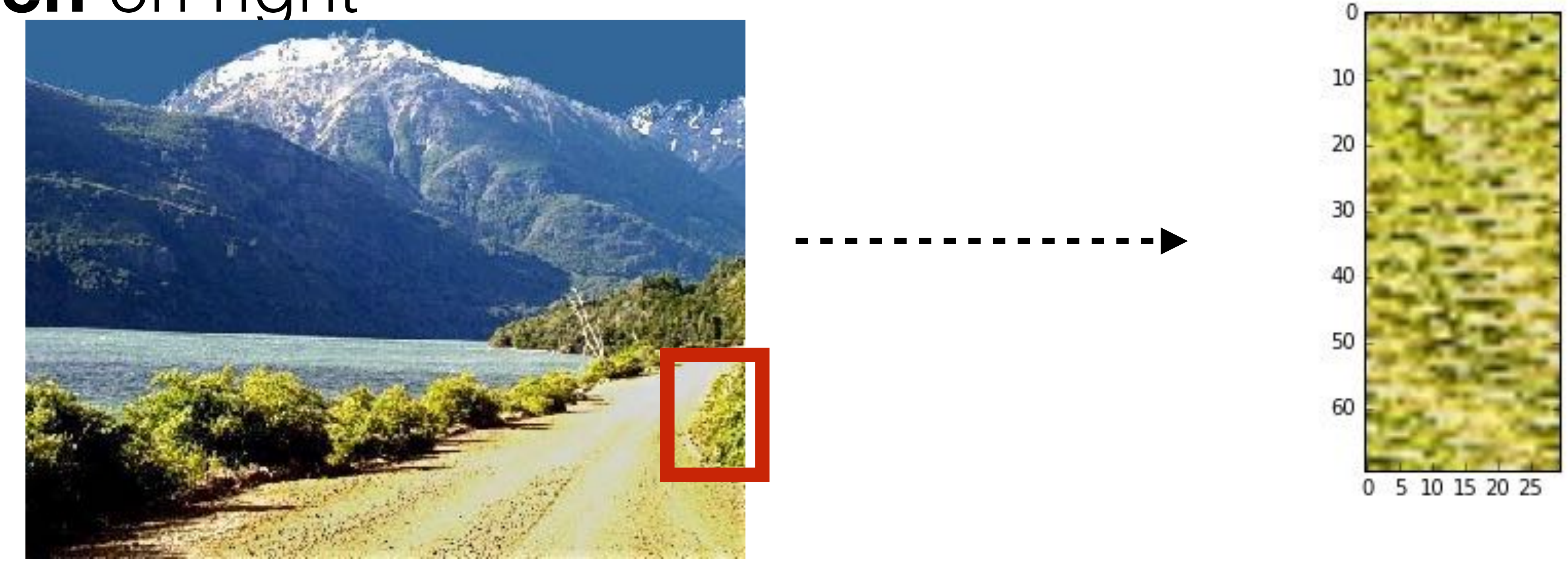


Variable Context

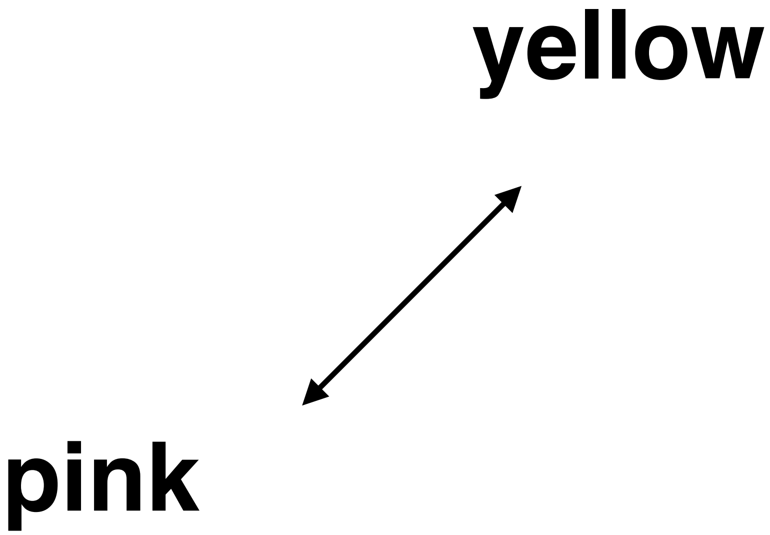
green trees in the middle

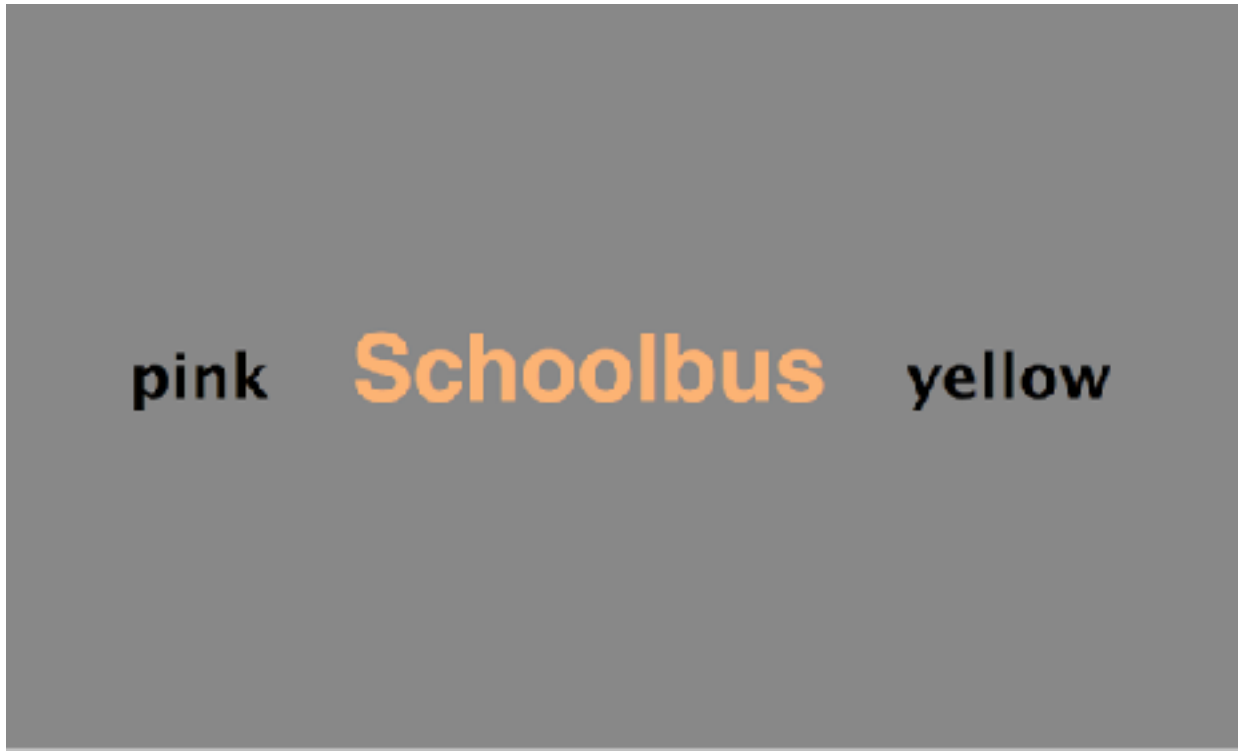


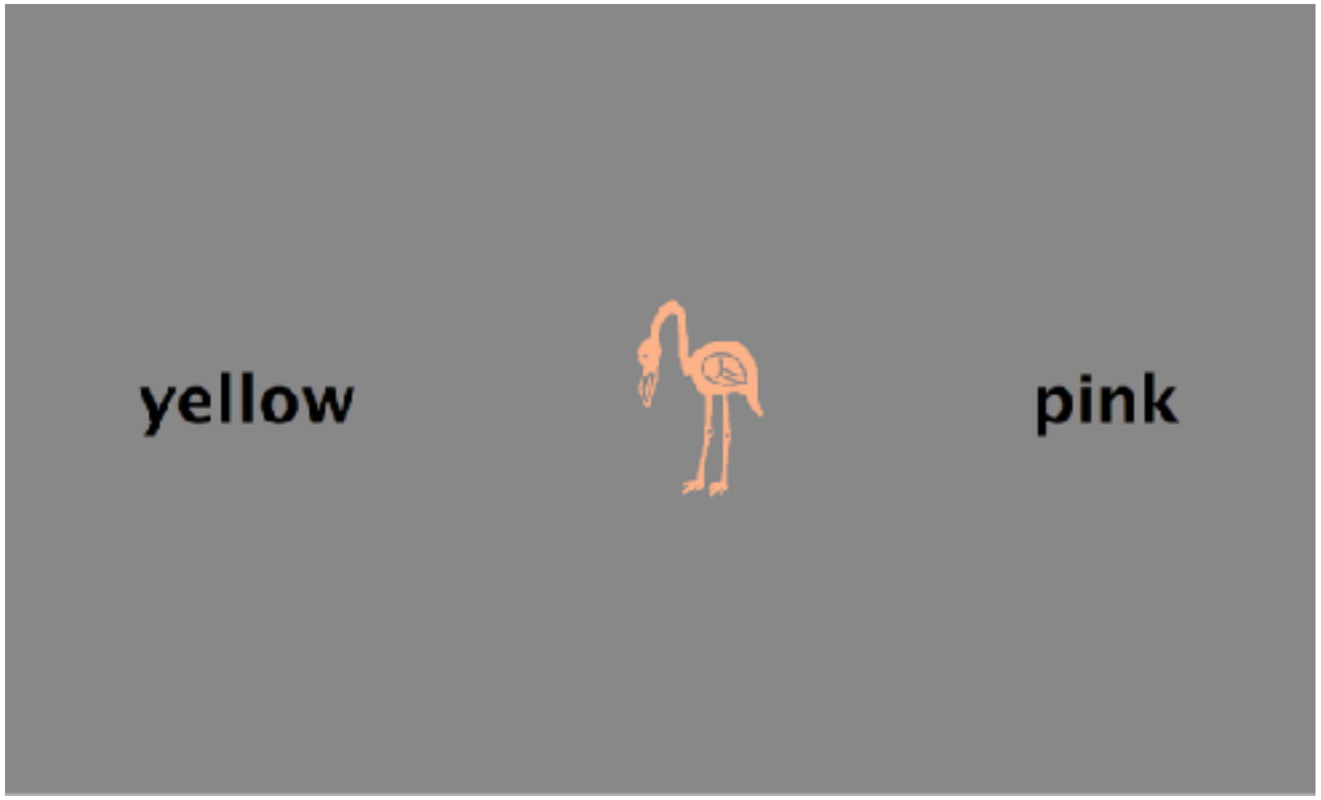
green on right



Top-down: Colours and World Knowledge




pink Schoolbus yellow


yellow pink

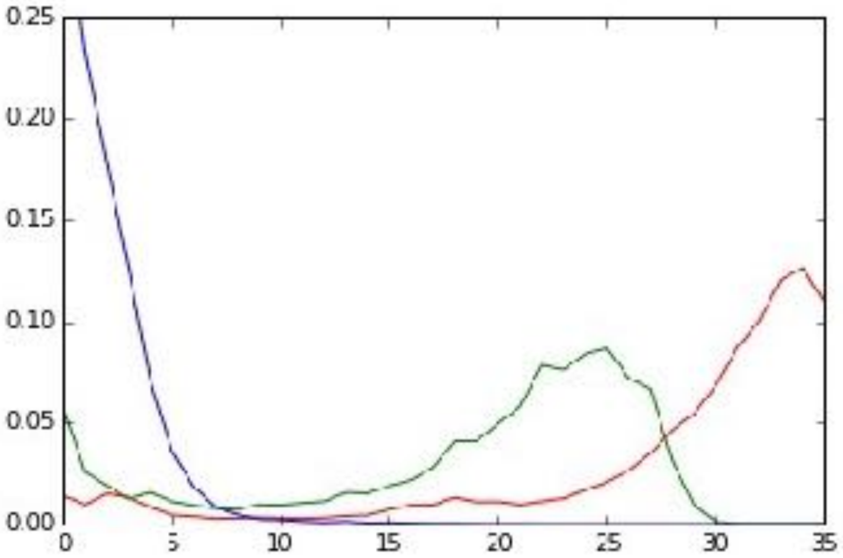
(Kubat, Mirman and Roy, 2009)

people calibrate ambiguous hues to **expected** colours

recalibration accounts for the colour constancy problem (Mitterer and de Ruiter, 2008)

Bottom-up: Perceptual Classifiers

- Multi-layer Perceptron: 512 input nodes (RGB hist), 11 output nodes (basic colour terms), 2 hidden layers
- ConvNet features don't seem to work for colour!



yellow 0.3
red 0.25
black 0.15
green 0.14
...

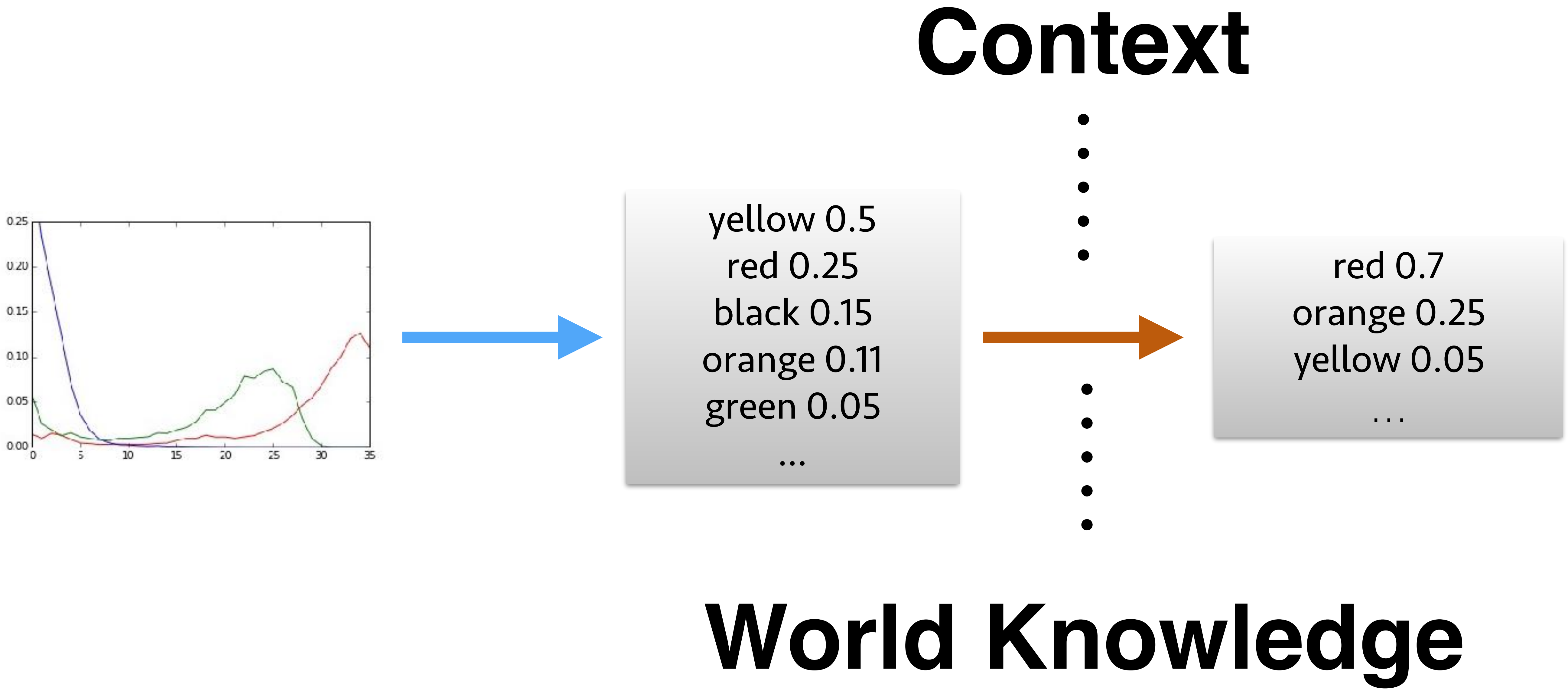
1 general classifier
or separate classifiers for each object?

Results: Perceptual Classifiers

- train and test on all objects, 1 classifier
- **63.7% accuracy**

- train and test separately on 52 object classes
- **45.1% accuracy**
- more sensitive to ``noise'', often recalibrate when general classifier is confident

Top-down: Recalibration

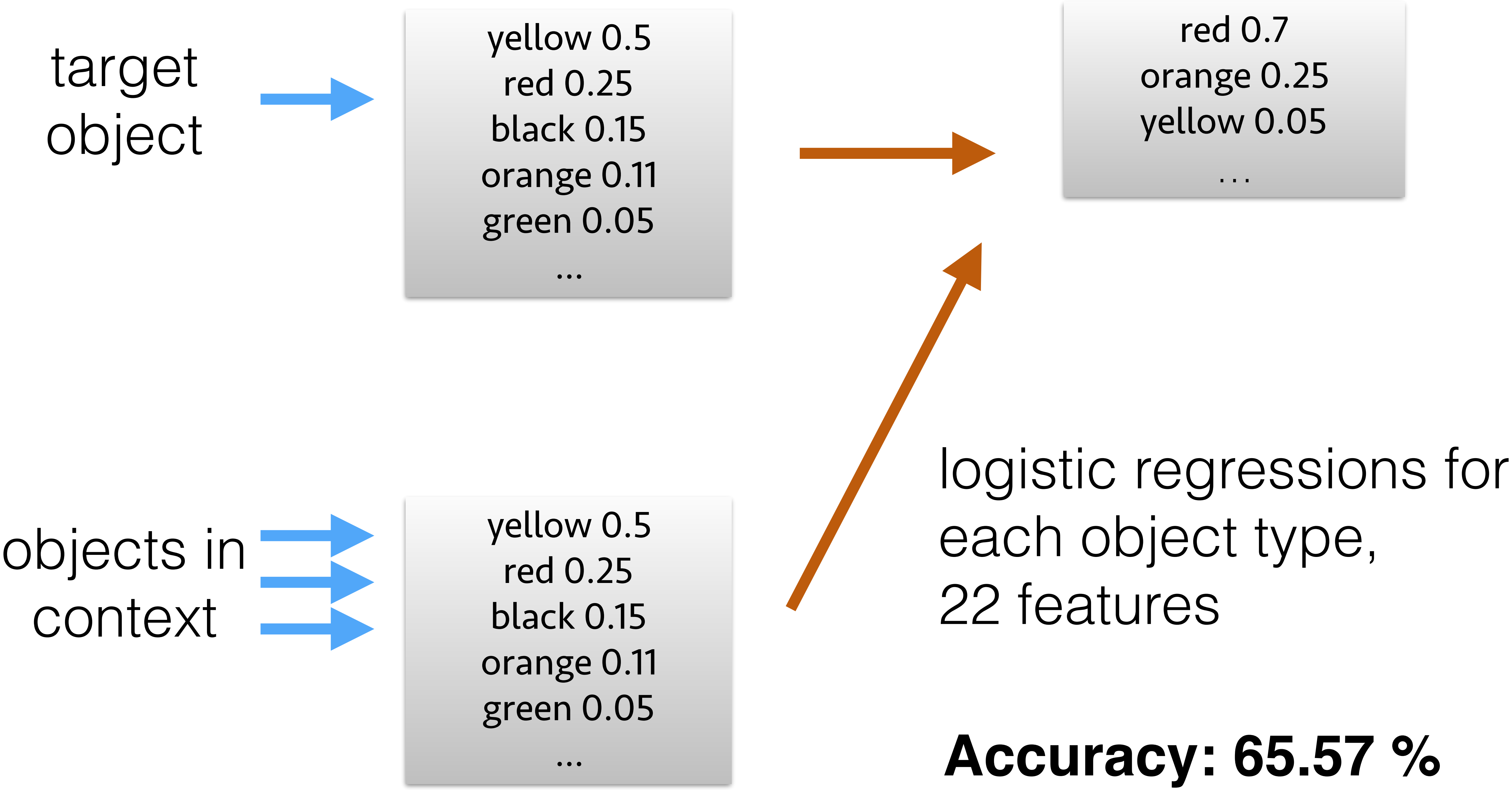


Context

yellow 0.5
red 0.25
black 0.15
orange 0.11
green 0.05
...

red 0.7
orange 0.25
yellow 0.05
. . .

World Knowledge



target object

yellow 0.5
red 0.25
black 0.15
orange 0.11
green 0.05
...

red 0.7
orange 0.25
yellow 0.05
. . .

objects in context

yellow 0.5
red 0.25
black 0.15
orange 0.11
green 0.05
...

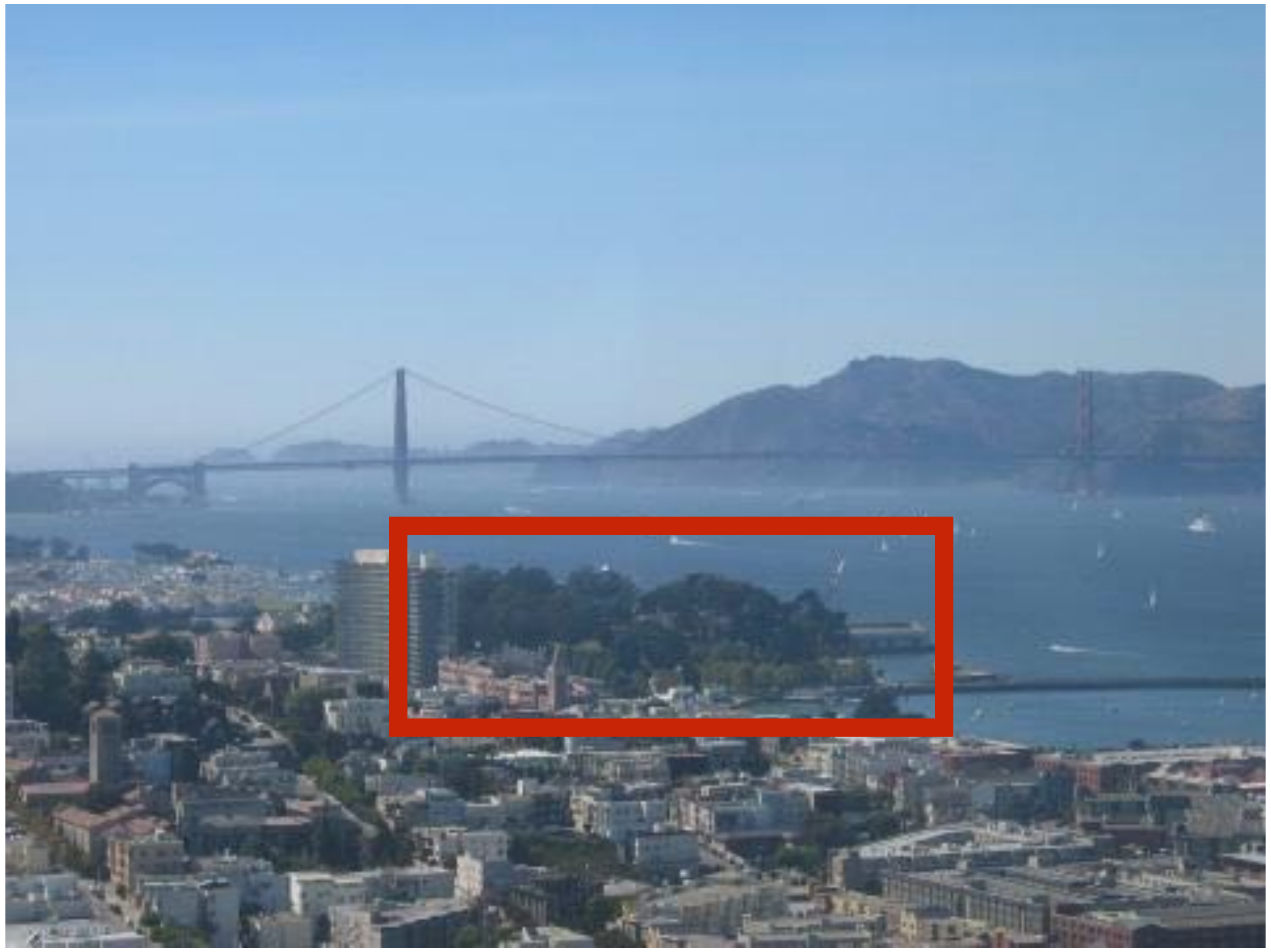
logistic regressions for each object type, 22 features

- cross-validation

Accuracy: 65.57 %

Perceptual: 63.7%

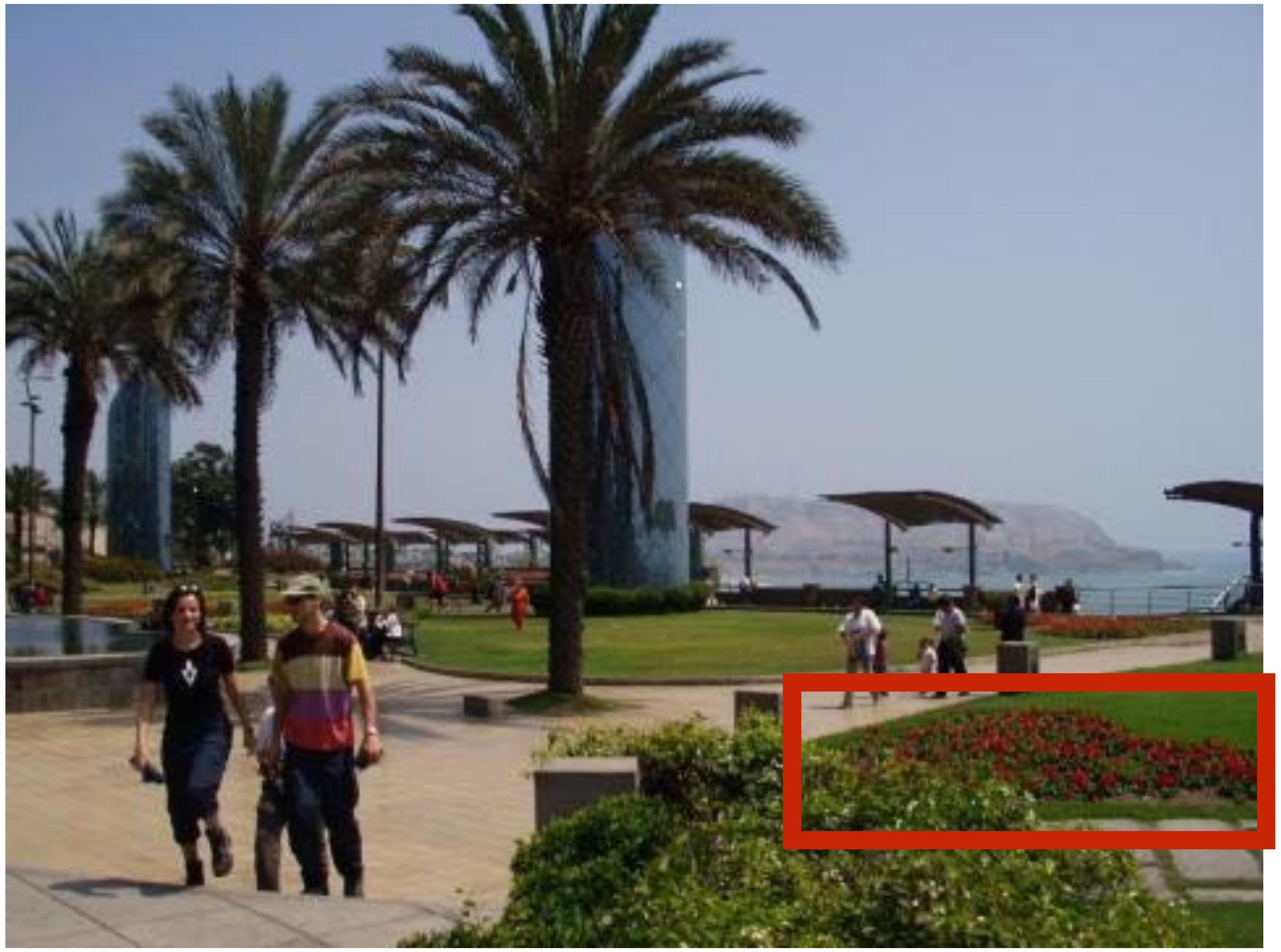
Object-specific: 45.1%



corpus: **green**
perceptual: blue
recalibrated: **green**

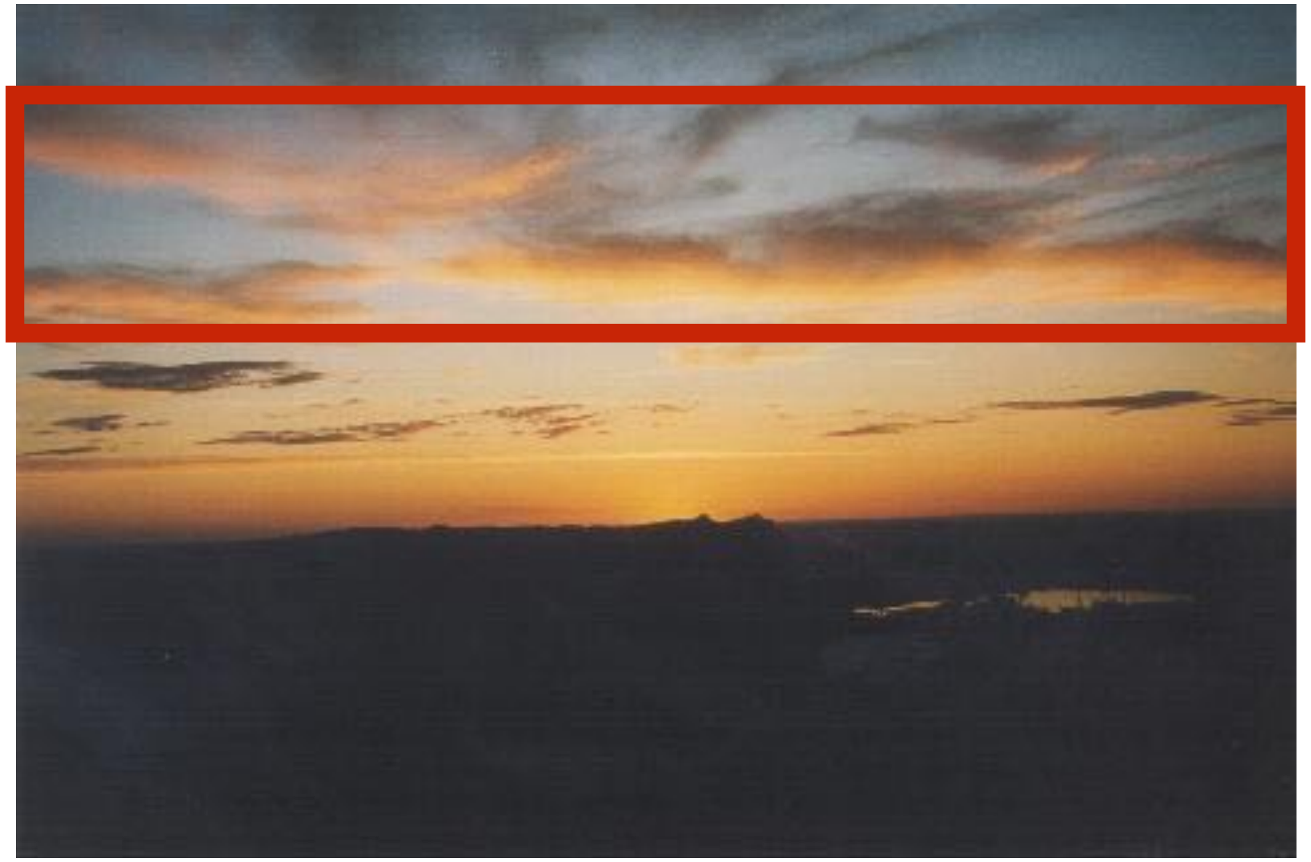
→ *expected colour*

Successful Recalibration



corpus: **red**
perceptual: green
recalibrated: **red**

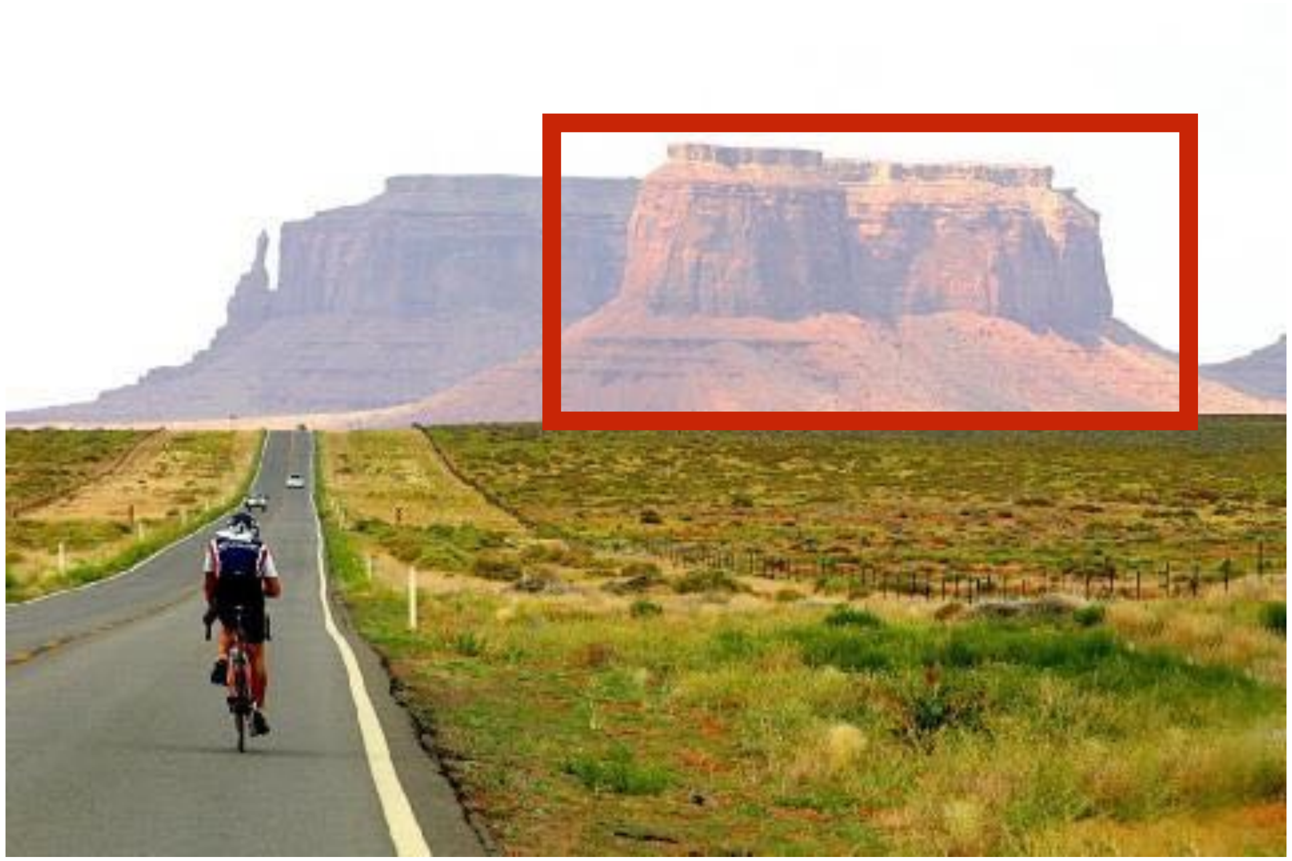
→ *salient colour*



corpus: **yellow**
perceptual: yellow
recalibrated: **white**

→ *expected colour*

Unsuccessful Recalibration



corpus: **red**
perceptual: pink
recalibrated: **white**

→ *expected colour*

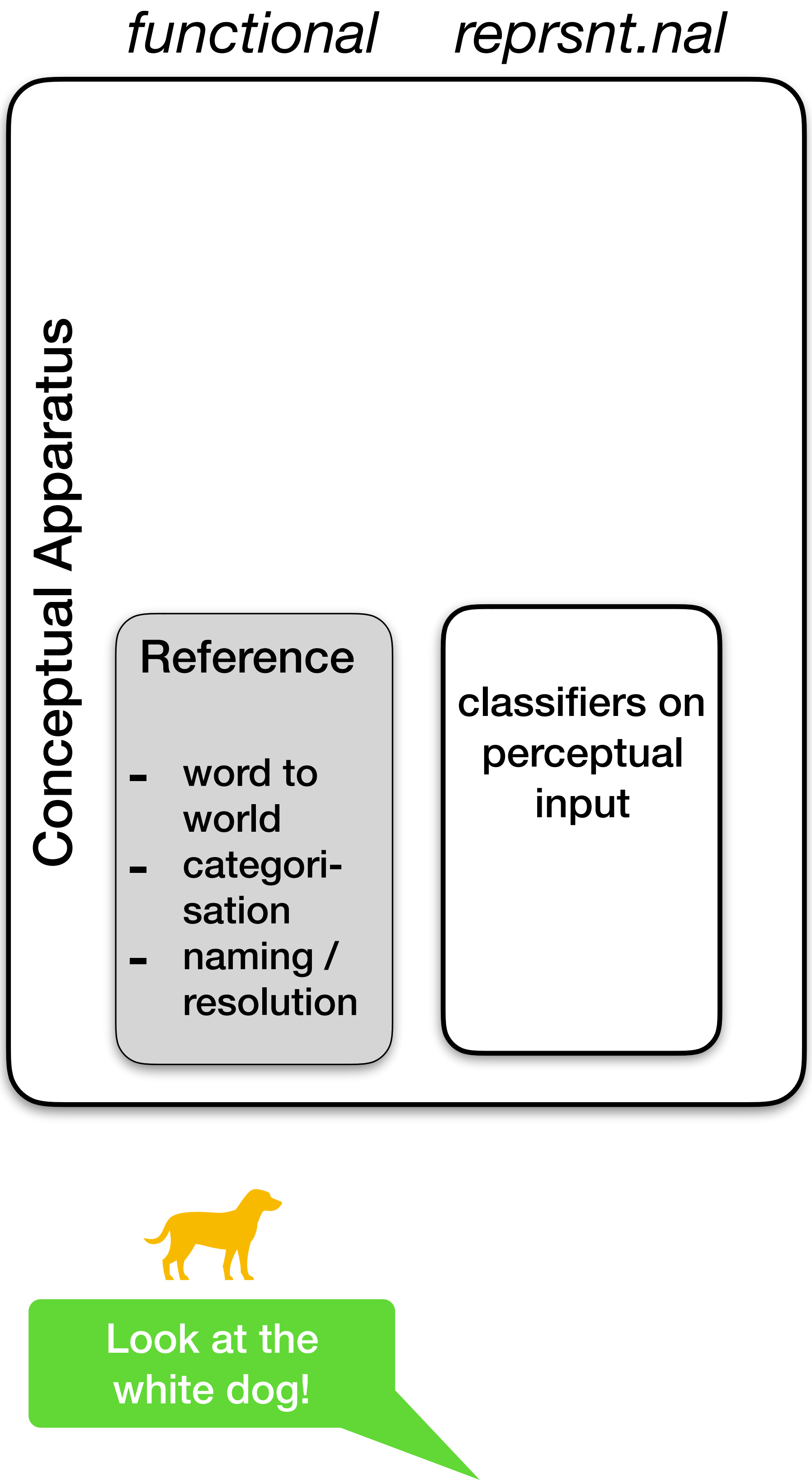
Grounding and Calibrating Colour Terms

- Bottom-up perceptual classifiers:
  visual input → distribution over colour terms

- Top-down recalibration:
  world knowledge: object types, typicality
  context: saliency, atypicality

to sum up

- when deciding on (name, attribute) categorisation of objects in a scene, just looking at the one single object that is to be categorised is not enough
  - need to look at the whole scene (visual context)
  - need to know something about expectations about connections
- other questions to pursue: how to recognise base level categories (category names), how to select appropriate level



functional *reprsnt.nal*

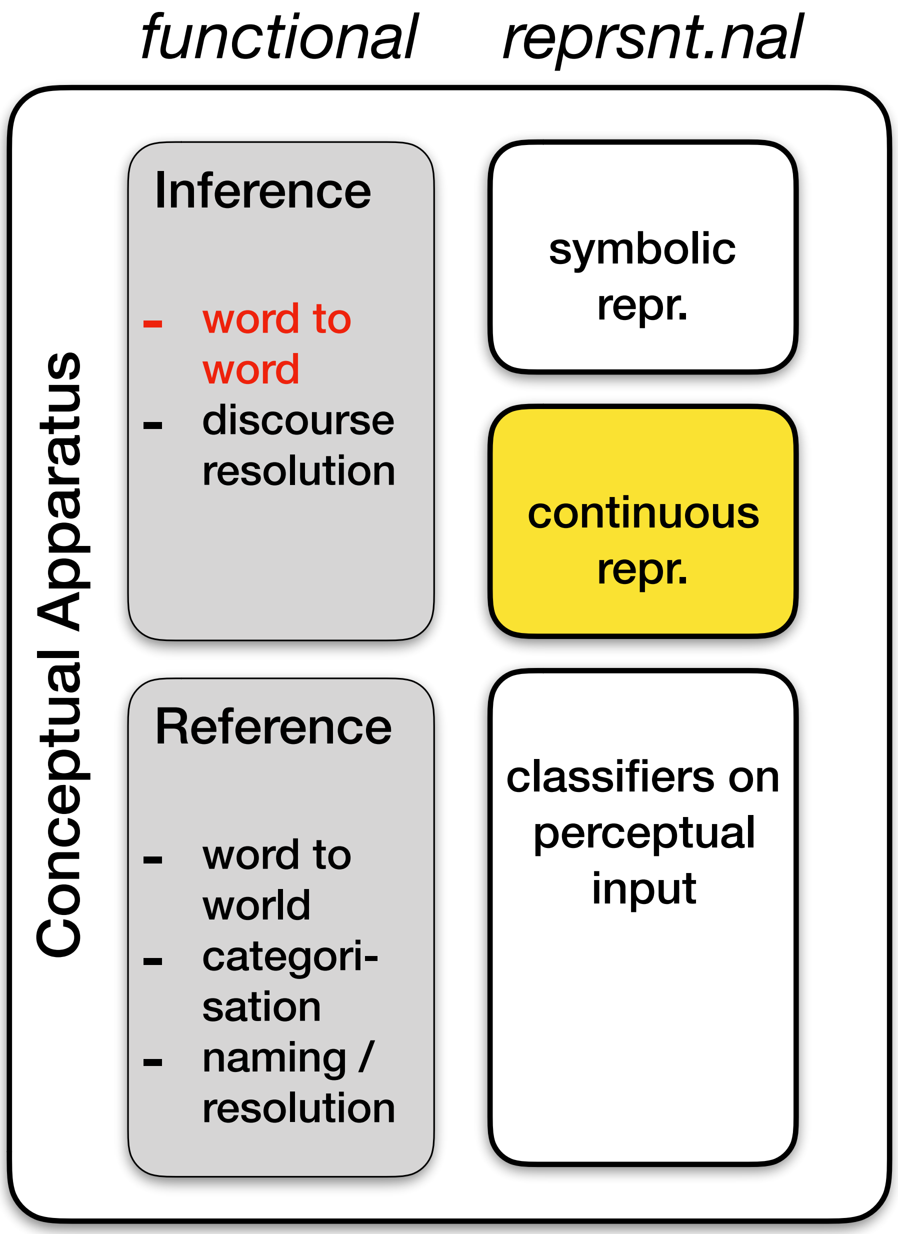
Conceptual Apparatus

Reference

- word to world
- categorisation
- naming / resolution

classifiers on perceptual input

Look at the white dog!



functional *reprsnt.nal*

Conceptual Apparatus

Inference

- word to word
- discourse resolution

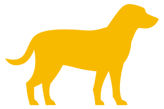
symbolic repr.

continuous repr.

Reference

- word to world
- categorisation
- naming / resolution

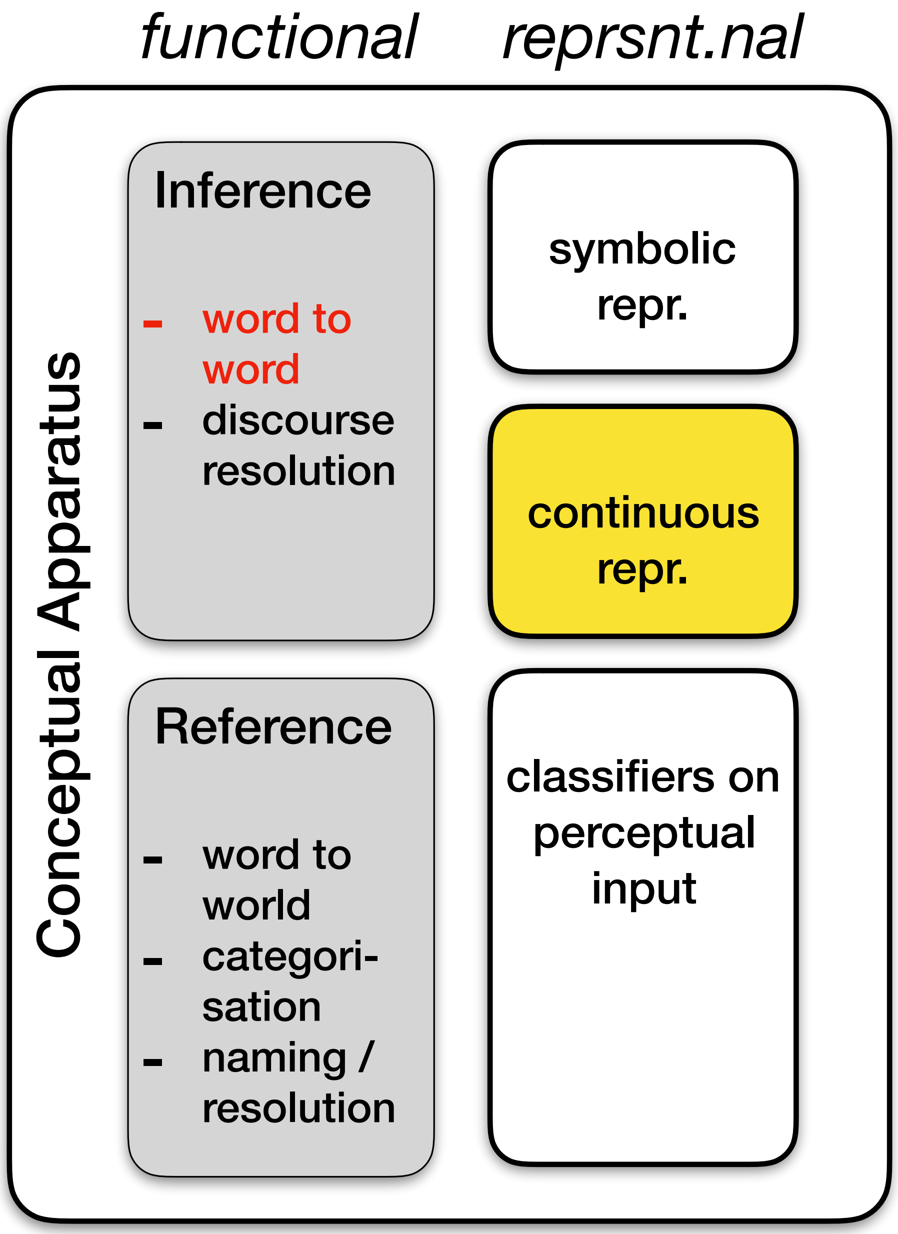
classifiers on perceptual input



Look at the white dog!

We just saw a cute dog.

The cutest poodle ever!



functional *reprsnt.nal*

Conceptual Apparatus

Inference

- word to word
- discourse resolution

Reference

- word to world
- categorisation
- naming / resolution

symbolic repr.

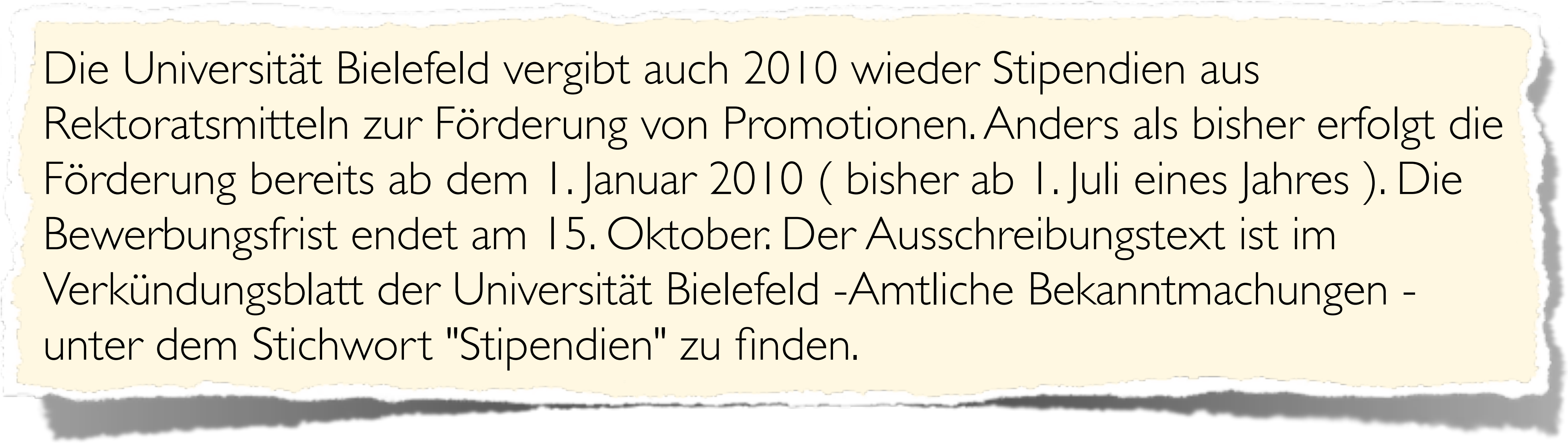
continuous repr.

classifiers on perceptual input

Harris (1954): “If A and B have almost identical environments we say that they are synonyms.”

Firth (1957): “You shall know a word by the company it keeps!”

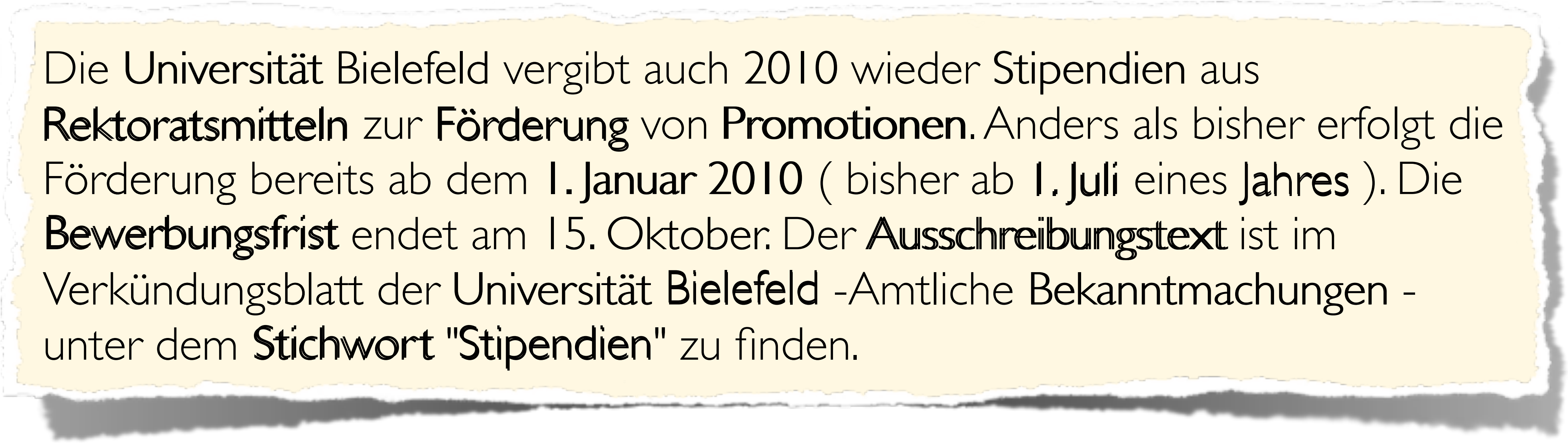
What is an “environment”, and what is “company” for a word?



Die Universität Bielefeld vergibt auch 2010 wieder Stipendien aus Rektoratsmitteln zur Förderung von Promotionen. Anders als bisher erfolgt die Förderung bereits ab dem 1. Januar 2010 (bisher ab 1. Juli eines Jahres). Die Bewerbungsfrist endet am 15. Oktober. Der Ausschreibungstext ist im Verkündungsblatt der Universität Bielefeld -Amtliche Bekanntmachungen - unter dem Stichwort "Stipendien" zu finden.

http://ekvv.uni-bielefeld.de/blog/uniaktuell/

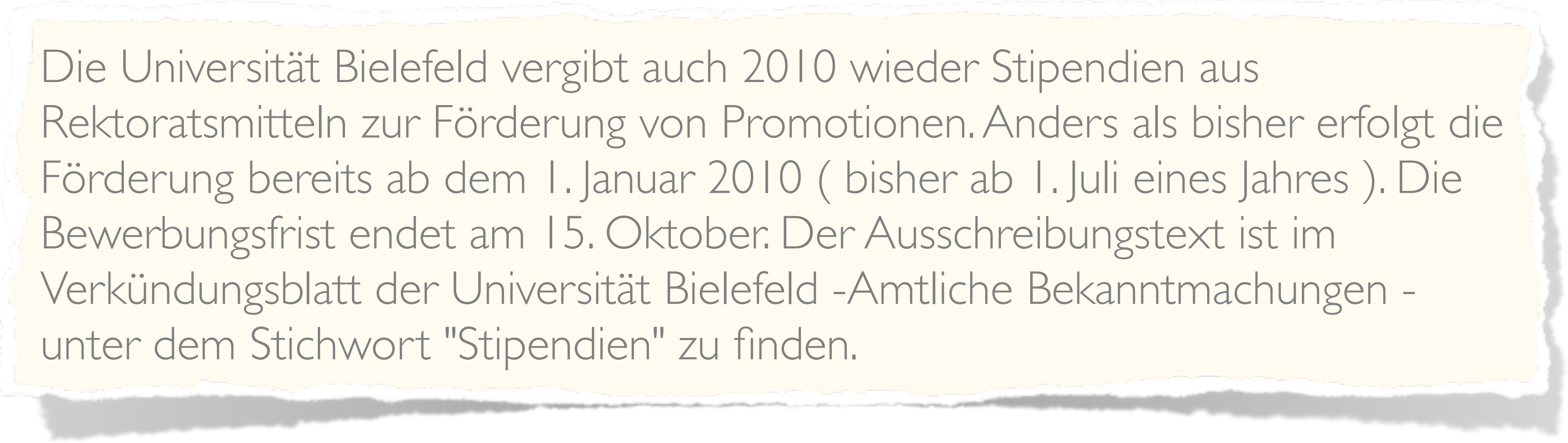
http://ekvv.uni-bielefeld.de/blog/uniaktuell/



Die Universität Bielefeld vergibt auch 2010 wieder Stipendien aus Rektoratsmitteln zur Förderung von Promotionen. Anders als bisher erfolgt die Förderung bereits ab dem 1. Januar 2010 (bisher ab 1. Juli eines Jahres). Die Bewerbungsfrist endet am 15. Oktober. Der Ausschreibungstext ist im Verkündungsblatt der Universität Bielefeld -Amtliche Bekanntmachungen - unter dem Stichwort "Stipendien" zu finden.

bag of words representation

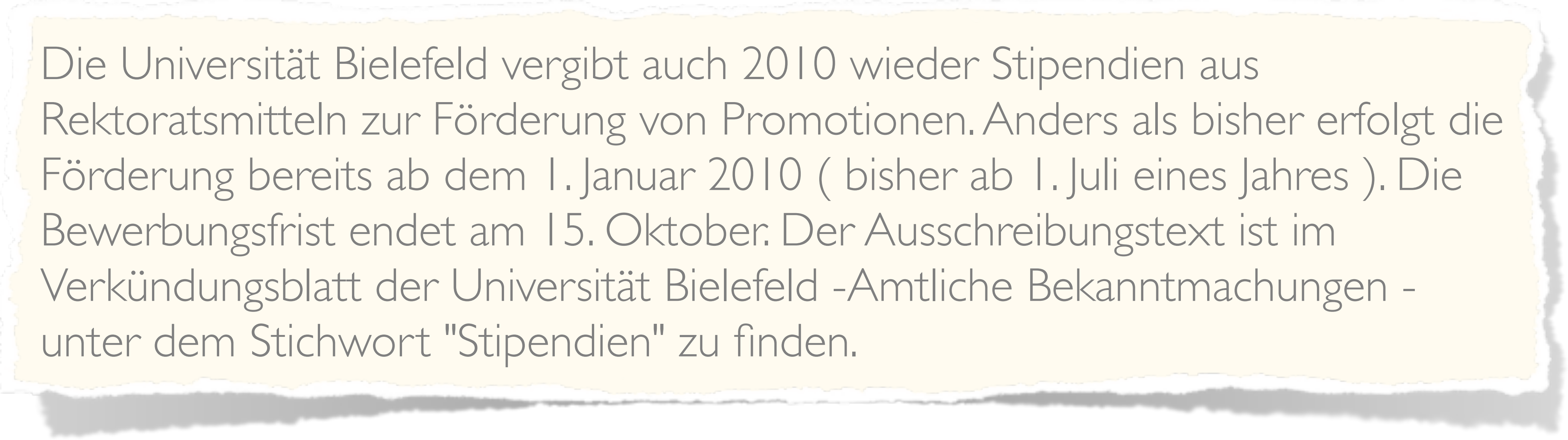
http://ekvv.uni-bielefeld.de/blog/uniaktuell/



Die Universität Bielefeld vergibt auch 2010 wieder Stipendien aus Rektoratsmitteln zur Förderung von Promotionen. Anders als bisher erfolgt die Förderung bereits ab dem 1. Januar 2010 (bisher ab 1. Juli eines Jahres). Die Bewerbungsfrist endet am 15. Oktober. Der Ausschreibungstext ist im Verkündungsblatt der Universität Bielefeld -Amtliche Bekanntmachungen - unter dem Stichwort "Stipendien" zu finden.

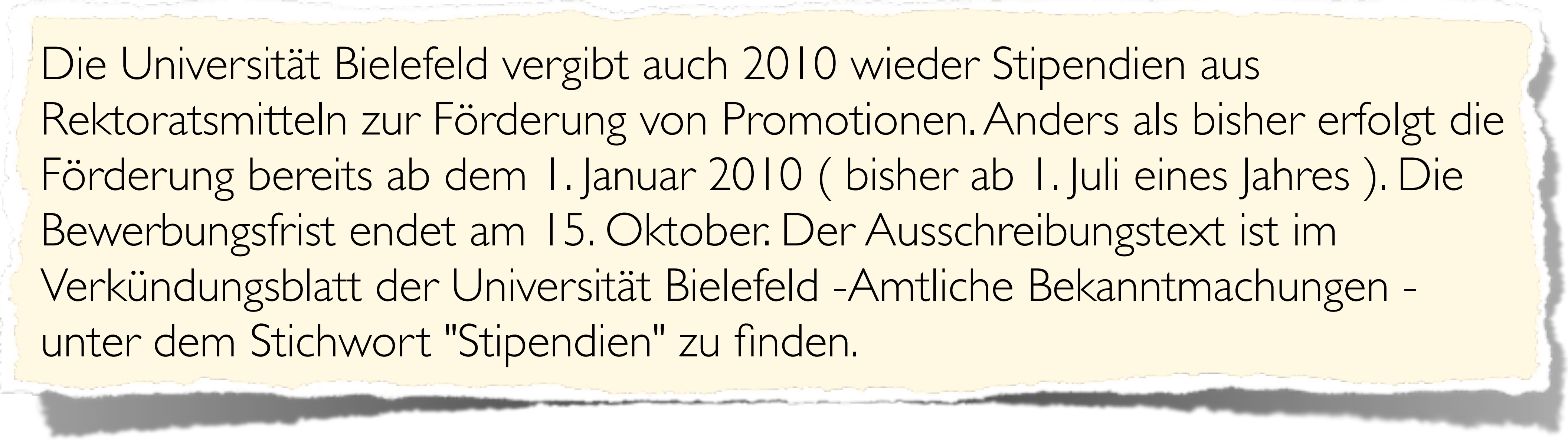
Ausschreibungstext	1
Bewerbungsfrist	1
Bekanntmachungen	1
Bielefeld	2
Förderung	2
Jahres	1
Promotionen	1
Rektoratsmitteln	1
Stichwort	1
Stipendien	2
Universität	2
Verkündungsblatt	1

http://ekvv.uni-bielefeld.de/blog/uniaktuell/



Die Universität Bielefeld vergibt auch 2010 wieder Stipendien aus Rektoratsmitteln zur Förderung von Promotionen. Anders als bisher erfolgt die Förderung bereits ab dem 1. Januar 2010 (bisher ab 1. Juli eines Jahres). Die Bewerbungsfrist endet am 15. Oktober. Der Ausschreibungstext ist im Verkündungsblatt der Universität Bielefeld -Amtliche Bekanntmachungen - unter dem Stichwort "Stipendien" zu finden.

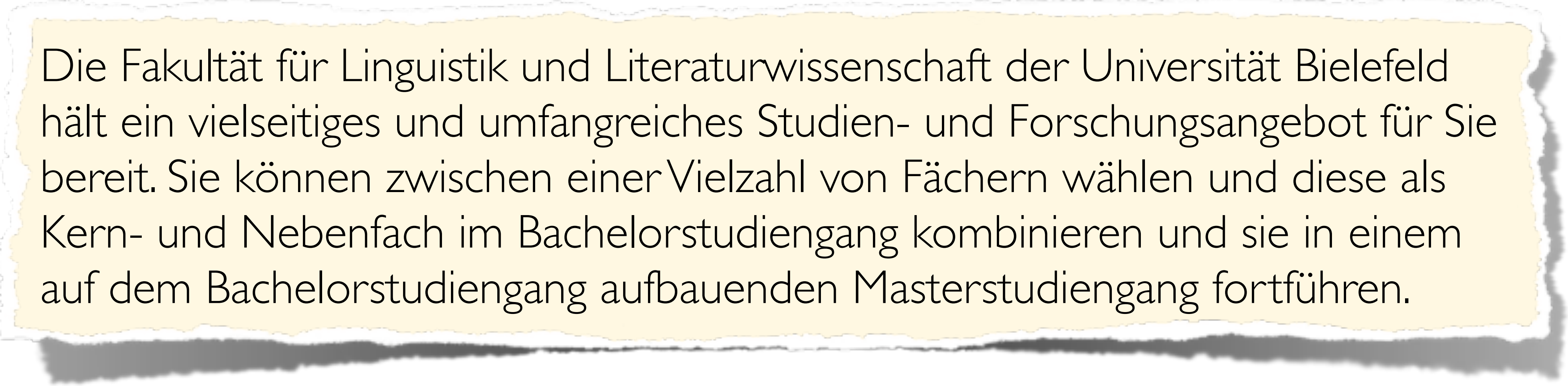
	Doc1
Ausschreibungstext	1
Bewerbungsfrist	1
Bekanntmachungen	1
Bielefeld	2
Förderung	2
Jahres	1
Promotionen	1
Rektoratsmitteln	1
Stichwort	1
Stipendien	2
Universität	2
Verkündungsblatt	1



Die Universität Bielefeld vergibt auch 2010 wieder Stipendien aus Rektoratsmitteln zur Förderung von Promotionen. Anders als bisher erfolgt die Förderung bereits ab dem 1. Januar 2010 (bisher ab 1. Juli eines Jahres). Die Bewerbungsfrist endet am 15. Oktober. Der Ausschreibungstext ist im Verkündungsblatt der Universität Bielefeld -Amtliche Bekanntmachungen - unter dem Stichwort "Stipendien" zu finden.

http://ekvv.uni-bielefeld.de/blog/uniaktuell/

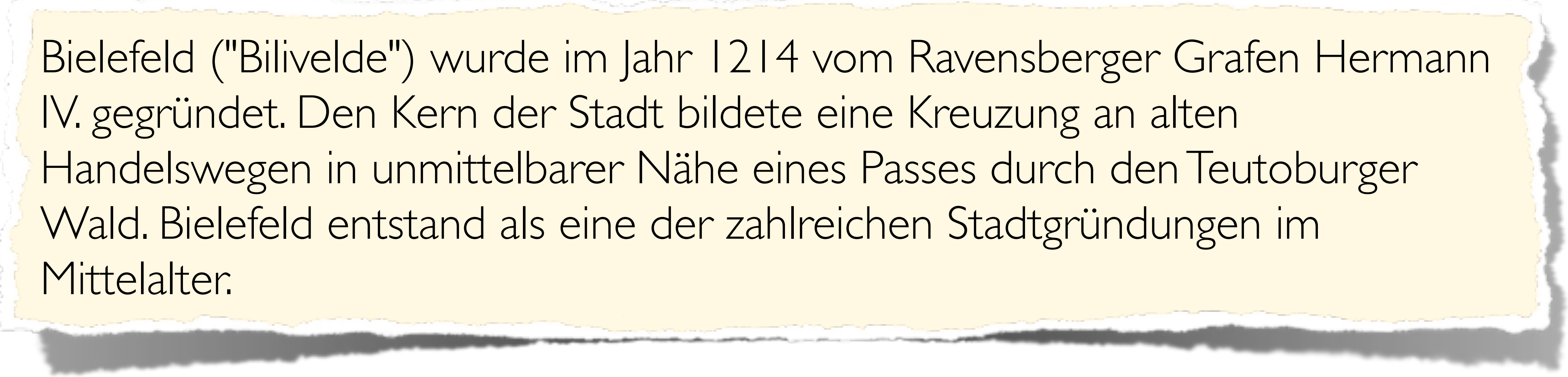
Uni BF



Die Fakultät für Linguistik und Literaturwissenschaft der Universität Bielefeld hält ein vielseitiges und umfangreiches Studien- und Forschungsangebot für Sie bereit. Sie können zwischen einer Vielzahl von Fächern wählen und diese als Kern- und Nebenfach im Bachelorstudiengang kombinieren und sie in einem auf dem Bachelorstudiengang aufbauenden Masterstudiengang fortführen.

http://www.uni-bielefeld.de/lili/startseite/

LiLi
Uni BF



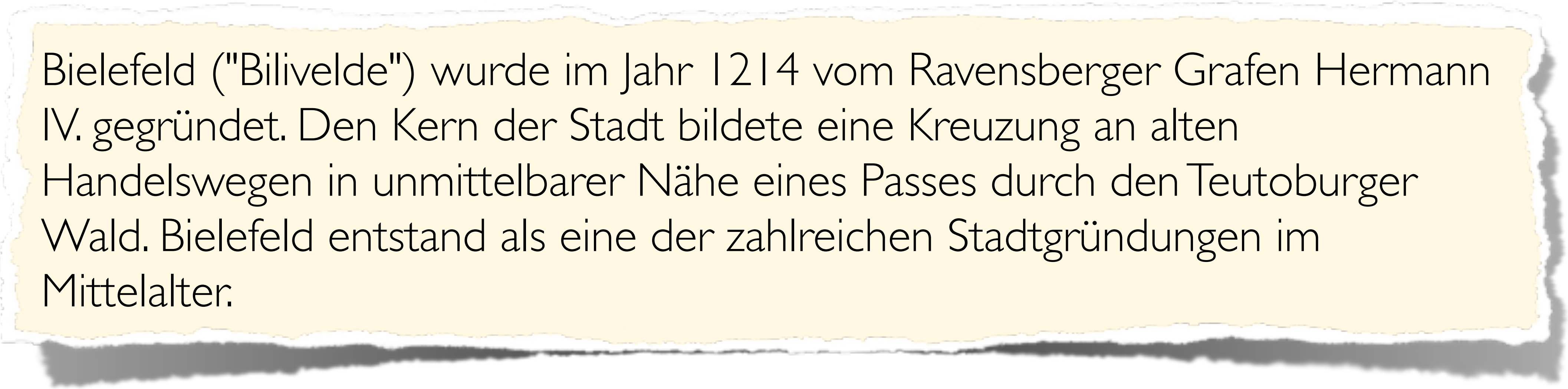
Bielefeld ("Bilivelde") wurde im Jahr 1214 vom Ravensberger Grafen Hermann IV. gegründet. Den Kern der Stadt bildete eine Kreuzung an alten Handelswegen in unmittelbarer Nähe eines Passes durch den Teutoburger Wald. Bielefeld entstand als eine der zahlreichen Stadtgründungen im Mittelalter.

http://www.bielefeld.de/de/si/geschichte/

Stadt BF

Bielefeld ("Bilivelde") wurde im Jahr 1214 vom Ravensberger Grafen Hermann IV. gegründet. Den Kern der Stadt bildete eine Kreuzung an alten Handelswegen in unmittelbarer Nähe eines Passes durch den Teutoburger Wald. Bielefeld entstand als eine der zahlreichen Stadtgründungen im Mittelalter.

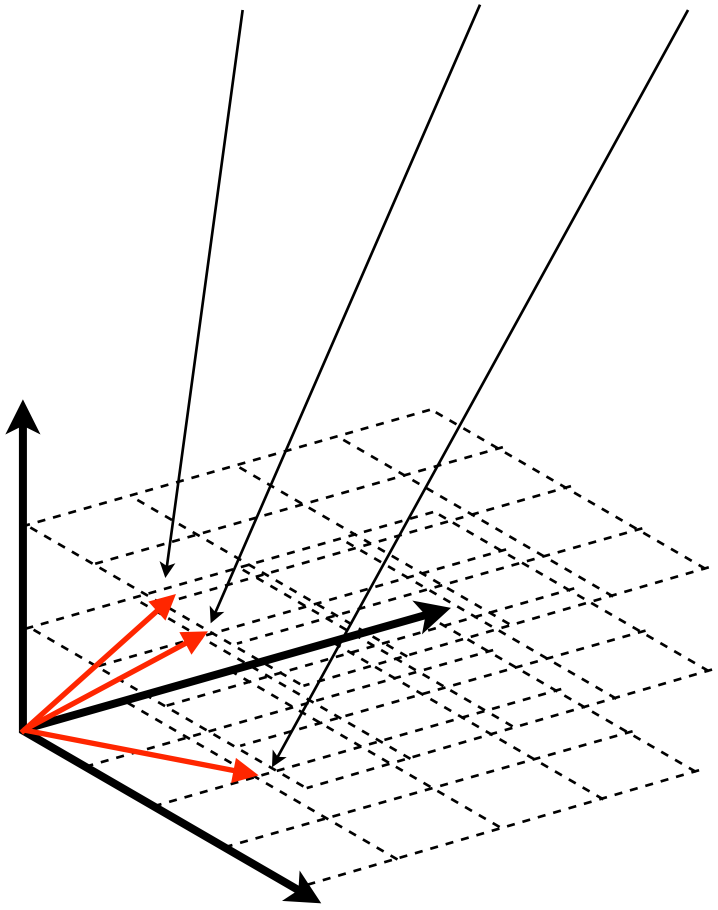
	Doc1	Doc2	Doc3
Ausschreibungstext	1	0	0
Bachelorstudiengang	0	1	0
Bewerbungsfrist	1	0	0
Bekanntmachungen	1	0	0
Bielefeld	2	1	2
Bilivelde	0	0	1
Fächern	0	1	0
Fakultät	0	1	0
Forschungsangebot	0	1	0
Förderung	2	0	0
Grafen	0	0	1
Handelswegen	0	0	1
Jahres	1	0	0
Kern-	0	1	1
Kreuzung	0	0	1
Linguistik	0	1	0
Literaturwissenschaft	0	1	0
Masterstudiengang	0	1	0
Mittelalter	0	0	1
Nebenfach	0	1	0
Promotionen	1	0	0
Rektoratsmitteln	1	0	0
Stadt	0	0	1
Stichwort	1	0	0
Stipendien	2	0	0
Studien-	0	1	0
Universität	2	1	0
Verkündungsblatt	1	0	0
Vielzahl	0	1	0



Bielefeld ("Bilivelde") wurde im Jahr 1214 vom Ravensberger Grafen Hermann IV. gegründet. Den Kern der Stadt bildete eine Kreuzung an alten Handelswegen in unmittelbarer Nähe eines Passes durch den Teutoburger Wald. Bielefeld entstand als eine der zahlreichen Stadtgründungen im Mittelalter.

	Doc1	Doc2	Doc3
Ausschreibungstext	1	0	0
Bachelorstudiengang	0	1	0
Bewerbungsfrist	1	0	0
Bekanntmachungen	1	0	0
Bielefeld	2	1	2
Bilivelde	0	0	1
Fächern	0	1	0
Fakultät	0	1	0
Forschungsangebot	0	1	0
Förderung	2	0	0
Grafen	0	0	1
Handelswegen	0	0	1
Jahres	1	0	0
Kern-	0	1	1
Kreuzung	0	0	1
Linguistik	0	1	0
Literaturwissenschaft	0	1	0
Masterstudiengang	0	1	0
Mittelalter	0	0	1
Nebenfach	0	1	0
Promotionen	1	0	0
Rektoratsmitteln	1	0	0
Stadt	0	0	1
Stichwort	1	0	0
Stipendien	2	0	0
Studien-	0	1	0
Universität	2	1	0
Verkündungsblatt	1	0	0
Vielzahl	0	1	0

	Doc1	Doc2	Doc3
Bielefeld	2	1	2
Kern	0	1	1
Universität	2	1	0

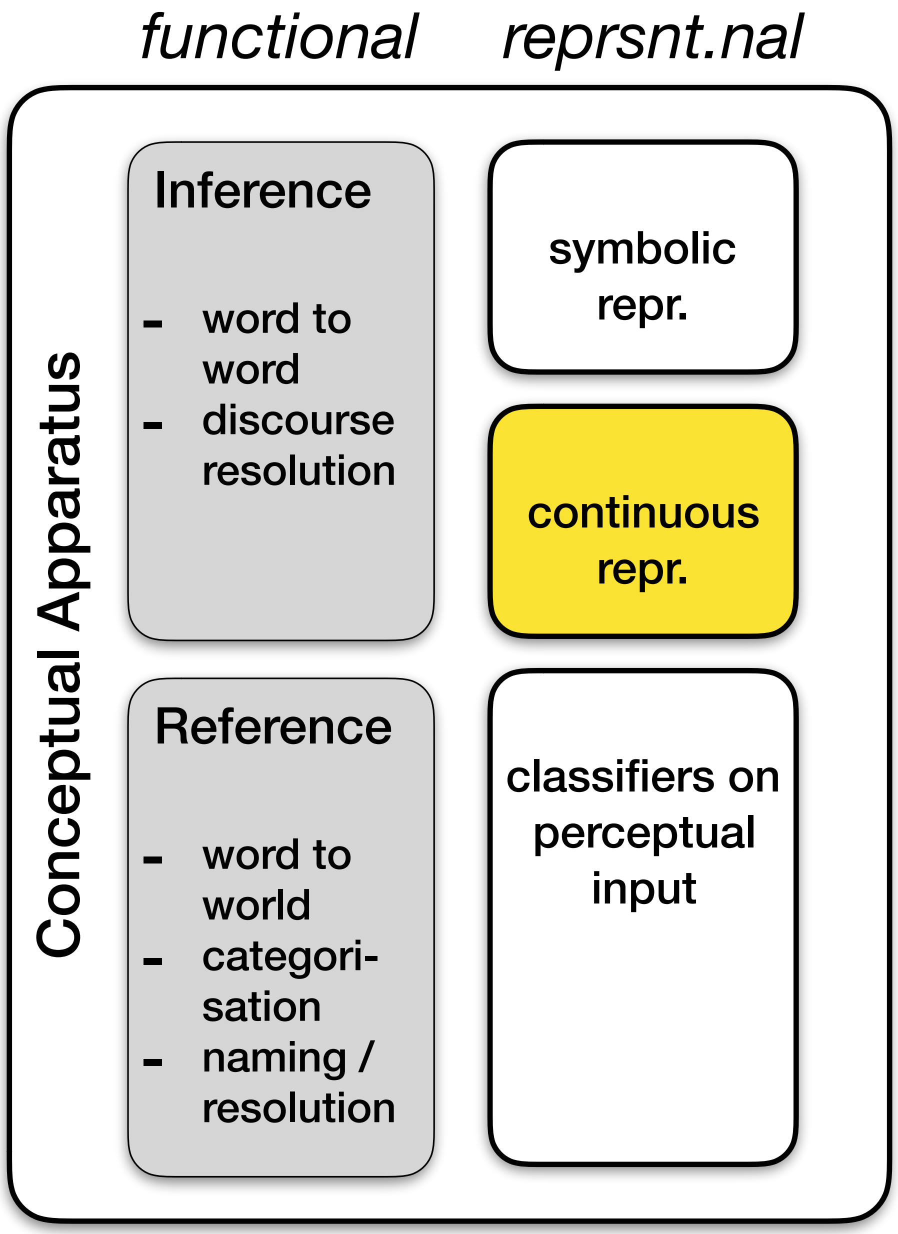


computing continuous word meaning representations

- this was the count-based way (which you would follow with a matrix factorisation step to reduce dimensionality & make representations less sparse)

- can also do this based on prediction: train NN to predict context words, and let it learn its input representations itself. To be good at this task, representations need to capture information about contexts.

- roughly equivalent (Goldberg & Levy 2014)



functional *reprsnt.nal*

Conceptual Apparatus

Inference

- word to word
- discourse resolution

Reference

- word to world
- categorisation
- naming / resolution

symbolic repr.

continuous repr.

classifiers on perceptual input

Harris (1954): “If A and B have almost identical environments we say that they are synonyms.”

Firth (1957): “You shall know a word by the company it keeps!”

What is an “environment”, and what is “company” for a word?

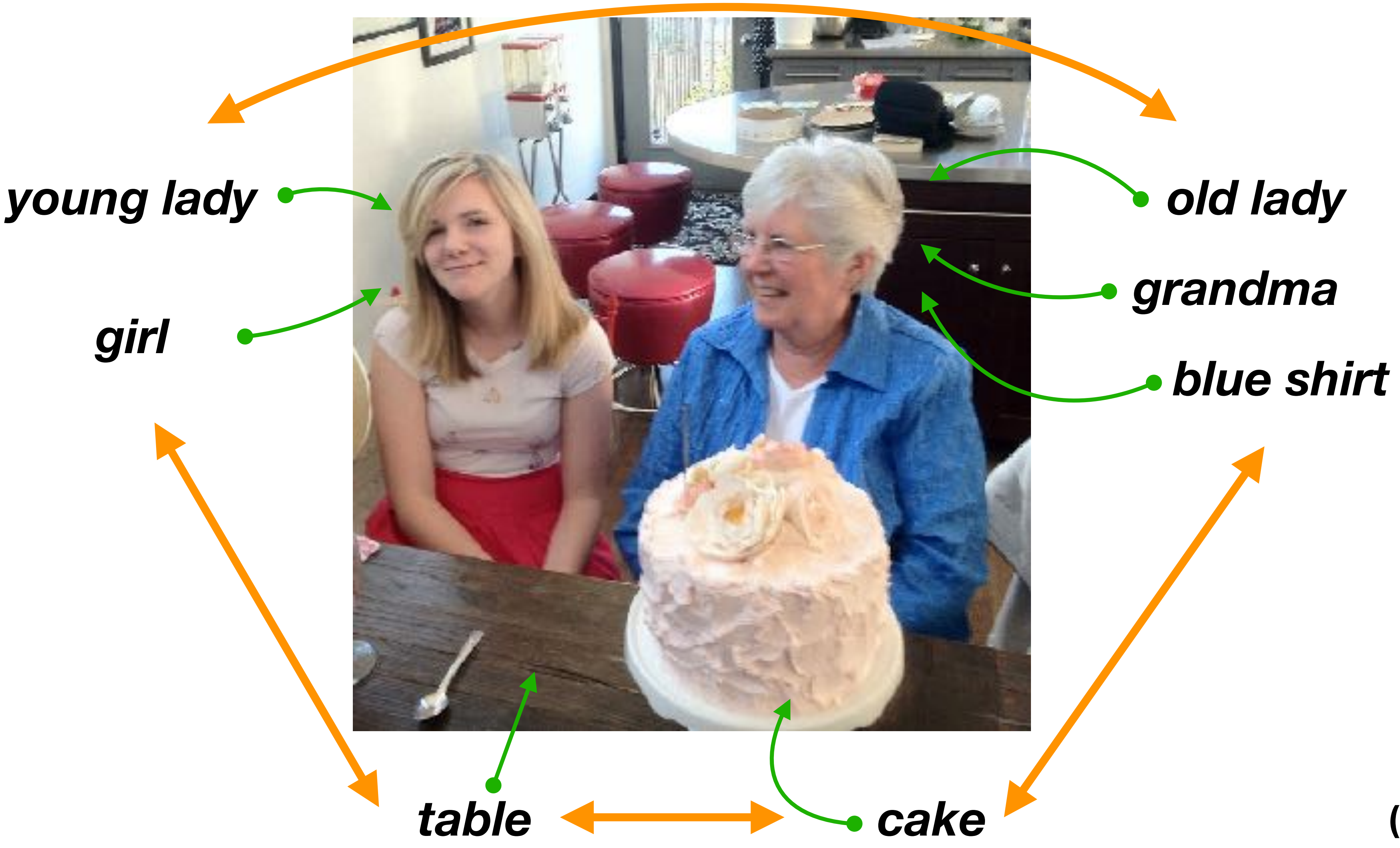
An *interpreted* observation of e: $o = \langle e, i, \delta, R_\delta, \omega, R_\omega \rangle$

$[[w]]_{ref} = f(O_w^A)$

$[[w]]_{inf} = g_w(f({}_iO^A))$ … needs only *uninterpreted* observations?

Observe *that* it is used, but not what it is used *for*?

structured contexts



young lady

girl

table

cake

old lady

grandma

blue shirt

continuous repr.

embeddings, from different kinds of context:

- ref.exp. as sentence, whole corpus, v_{txt}
- co-referential exp. as context, v_{ref}
- situation as context, v_{sit}

(Zarrieß & Schlangen, EMNLP 2017)

summary

- today has been about making connections:
  - between the objects in the scene, to yield a plausible interpretation
  - between objects and attributes, to name colours in an understandable way
  - between words / concepts, based on the contexts in which they were used [to be continued]
  - between representations, to transfer knowledge
  - between the parts out of which a phrase is built

Thank you.

Thanks also to my Bielefeld PhD students, Postdocs, and collaborators: Julian Hough, Sina Zarrieß, Casey Kennington, Nikolai Ilinykh, Soledad Lopez, Ting Han, Nazia Attari, Spyros Kousidis.

Funding received from CITEC, DFG.

References

References to our own work can be resolved via http://clp.ling.uni-potsdam.de/publications/ (where also the PDFs are available).
(First authors: Han, Kennington, Kousidis, Lopez, Schlangen, Zarrieß.)

- Bruni, E., Boleda, G., & Baroni, M. (2012). Distributional Semantics in Technicolor. In ACL 2012 (pp. 136–145).
- Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), Mechanisms of Language Acquisition. Hilsdale, New Jersey, USA: Lawrence Erlbaum Associates.
- Elman, J. L. (1990). Finding Structure in Time. Cognitive Science, 14, 179–211.
- Firth, J. R. (1957). A Synopsis of Linguistic Theory, 1930-1955. In Studies in linguistic analysis.
- Goldberg, Y., & Levy, O. (2014). word2vec Explained: Deriving Mikolov et al.'s Negative-Sampling Word-Embedding Method. arXiv preprint arXiv:1402.3722.
- Goodman, N. (1955). Fact, Fiction, & Forecast, Cambridge Massachusetts: Harvard University Press
- Harris, Z. S. (1954). Distributional Structure. Word, 10(2–3), 146–162.
- Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1–32.
- Kruszewski, G., & Baroni, M. (2015). So similar and yet incompatible : Toward automated identification of semantically compatible words. In The 2015 Annual Conference of the North American Chapter of the ACL (pp. 964–969).
- McRae, K., Cree, G. S., Seidenberg, M. S., & McNorgan, C. (2005). Semantic feature production norms for a large set of living and nonliving things. Behavior Research Methods, Instruments & Computers, 37(4), 547--559.
- Silberer, C., & Lapata, M. (2014). Learning Grounded Meaning Representations with Autoencoders. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), 721–732.
- Socher, R., Lin, C. C., Ng, A. Y., & Manning, C. D. (2011). Parsing Natural Scenes and Natural Language. In Proc. 28th International Conference on Machine Learning.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … Polosukhin, I. (2017). Attention Is All You Need. In NIPS.