Word, concept, perception and brain

Current and future work

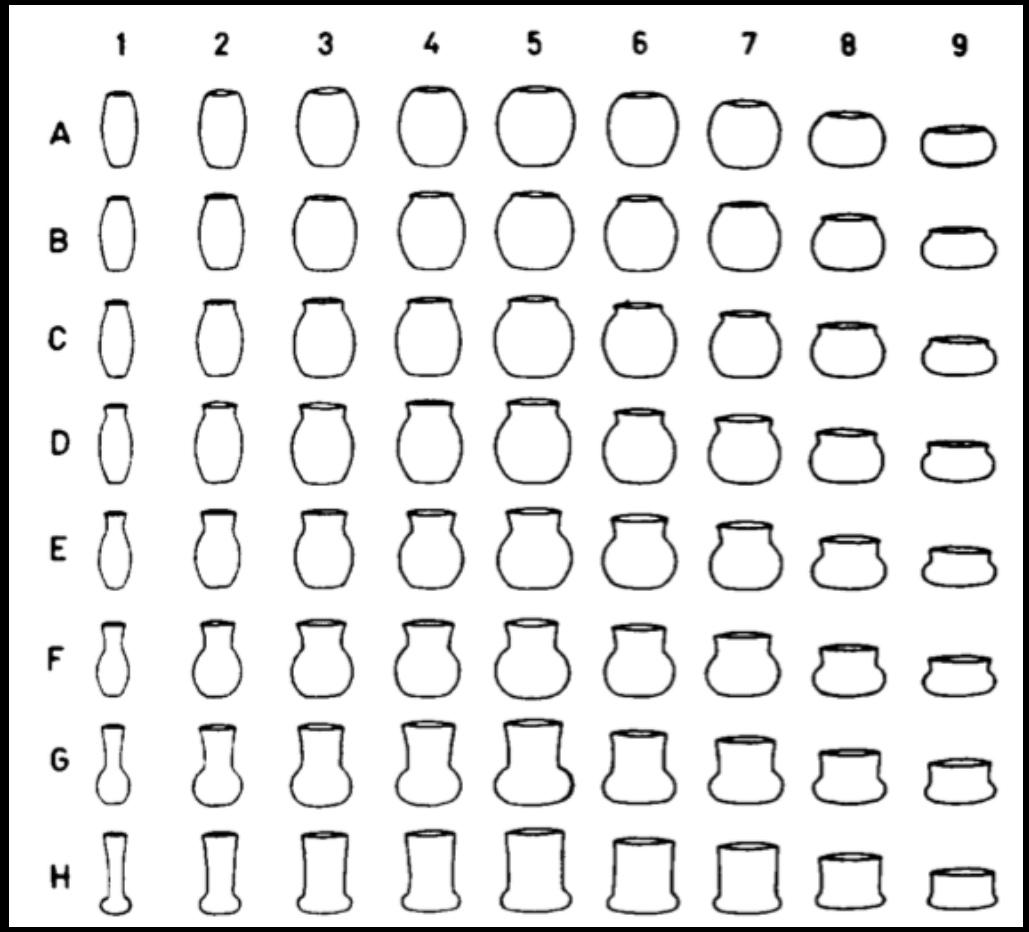
Alessandro Lopopolo Emiel van Miltenburg

Part I

Words and perception

Categorization

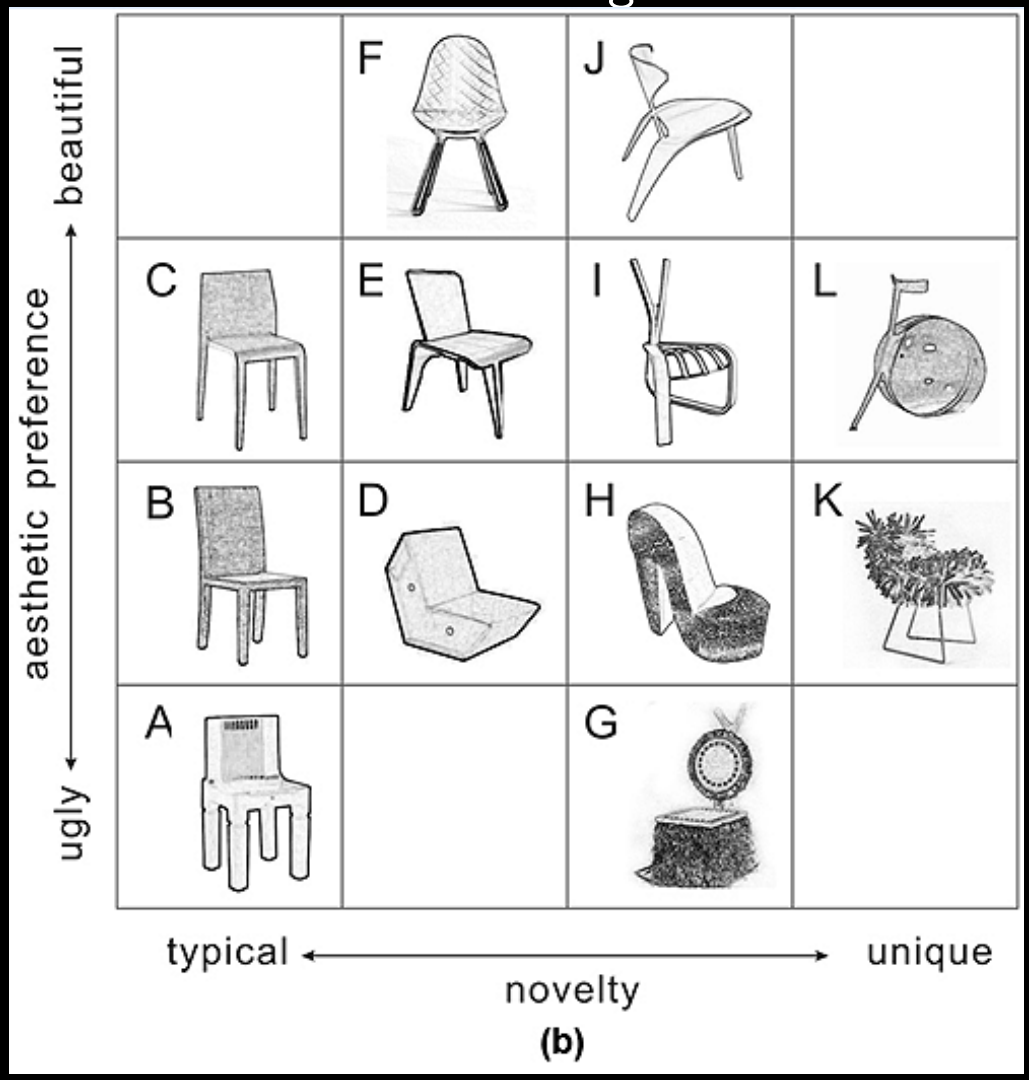
When is something a *pot*?



(Kempton 1984)

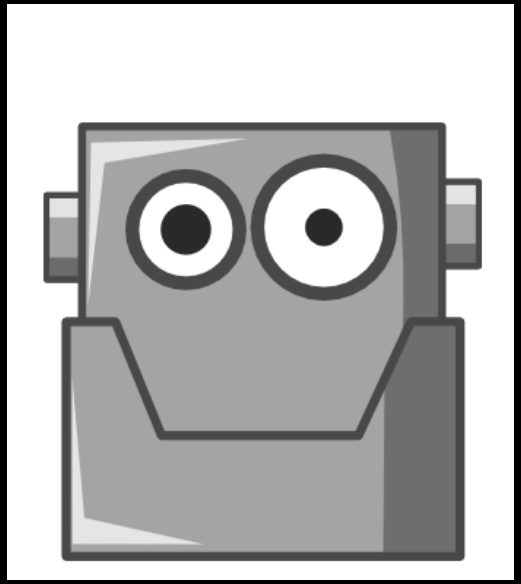
Categorization

When is something a *chair*?

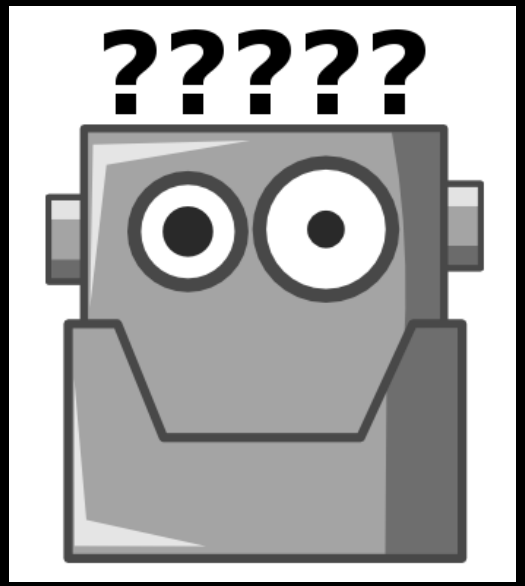


(http://www.ijdesign.org/ojs/index.php/IJDesign/article/view/1146/479)

“Could you get a chair for us, please?”



“Could you get a chair for us, please?”



Ambiguity

Link to ULM-1

- Not just an intralinguistic problem
- . . . but also a problem of reference
- Perceptual context is useful for disambiguation

Linking words to the world

How does the word *boot* points to an image of a boot?
How can a container object be descibed by the word *pot*?

Symbol Grounding

Relating words and external perceptual reality is a topic of interest for philosophers, linguists and researchers in artificial intelligence and cognitive neuroscience. (Harnad 1990)

ULM-2

Studying the relation between words, concepts and perception

Multidisciplinary:

- Linguistics
- Psychology/Anthropology
- Machine learning
- Neuroscience

Part II

Methods and theory

Methods

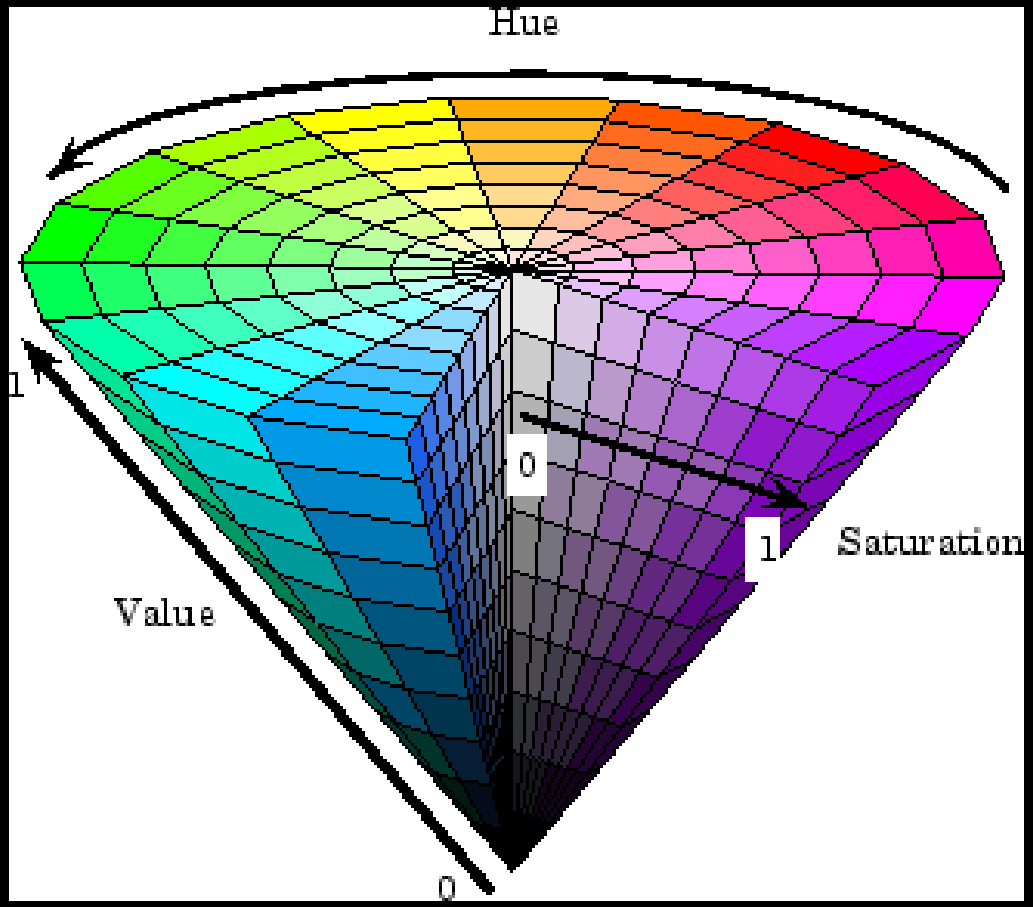
- Tools from NLP, image & sound processing
  $\rightarrow$ Putting current state of the art in perspective
- Behavioral data
- Brain imaging data

Central metaphor

- Conceptual spaces (Gärdenfors 2000, 2004)
- Words map to points or regions in conceptual space

Conceptual spaces

Color



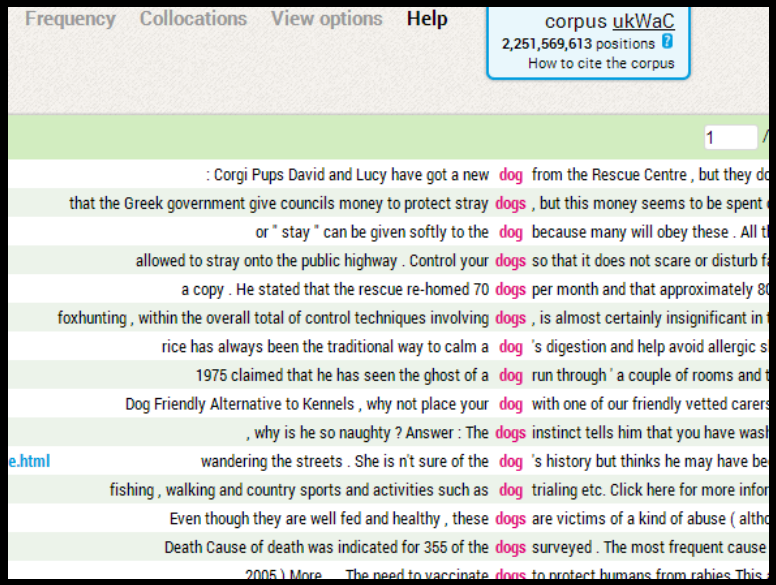
Conceptual spaces

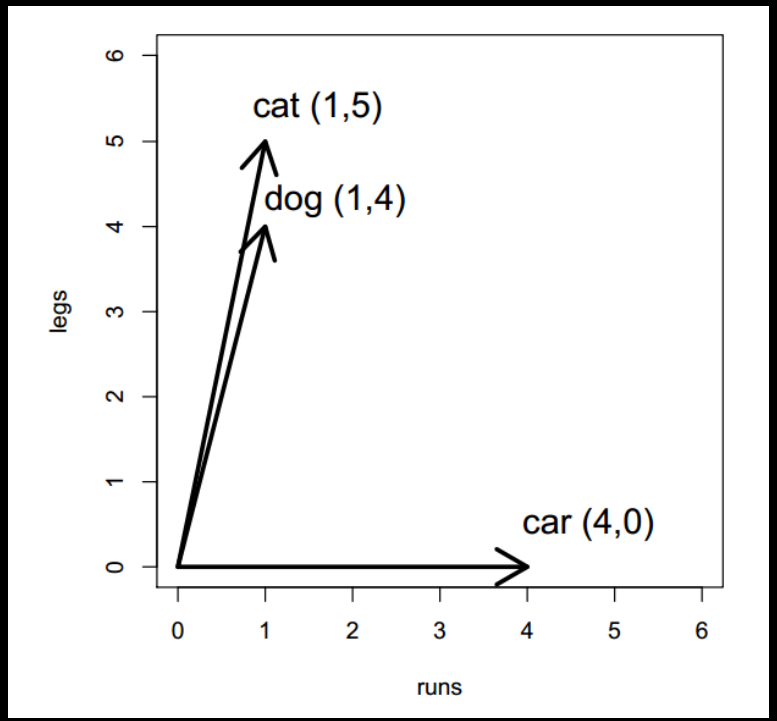
Part III

From theory to practice

Text-based space

You shall know a word by the company it keeps!

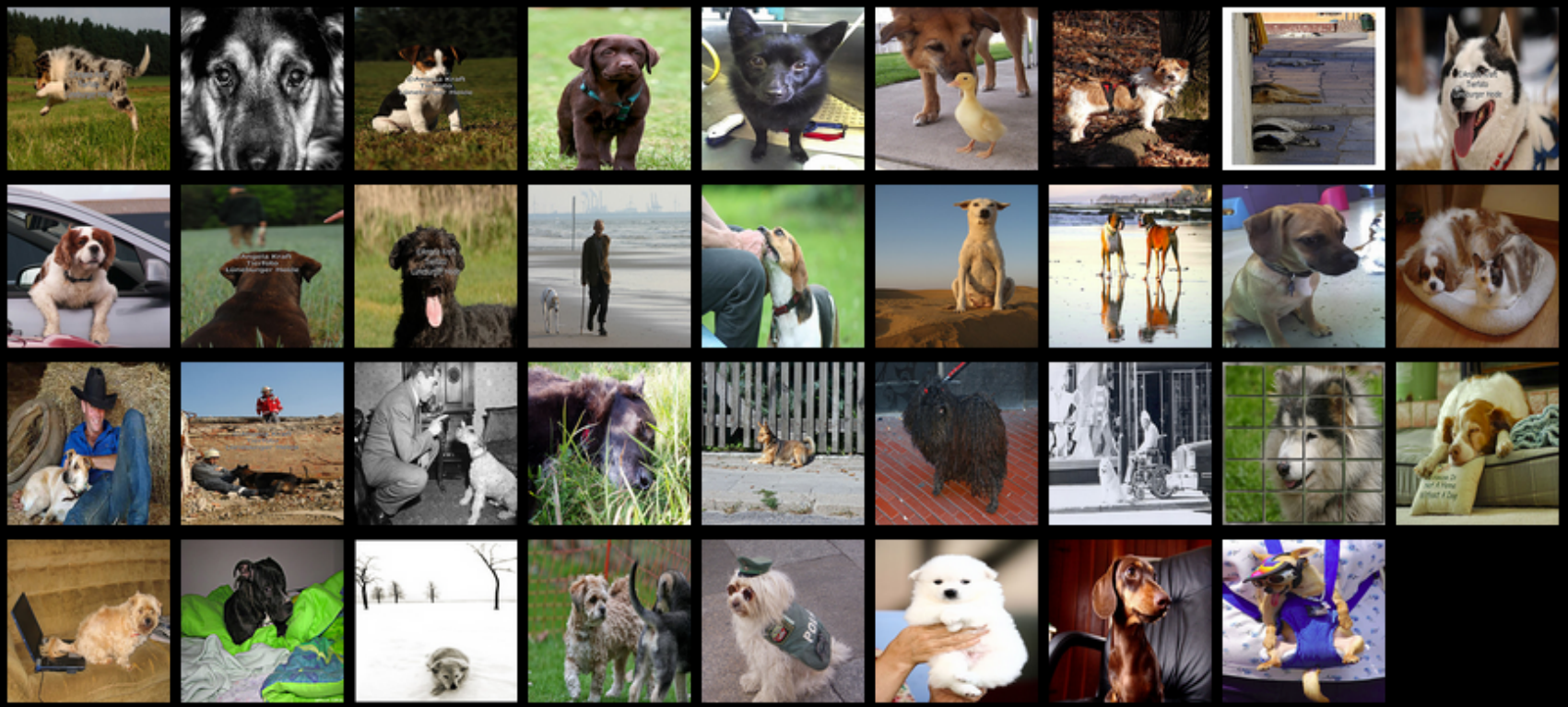




(Lund & Burgess 1996, Landauer & Dumais 1997, Baroni & Lenci 2010)

Image-based space

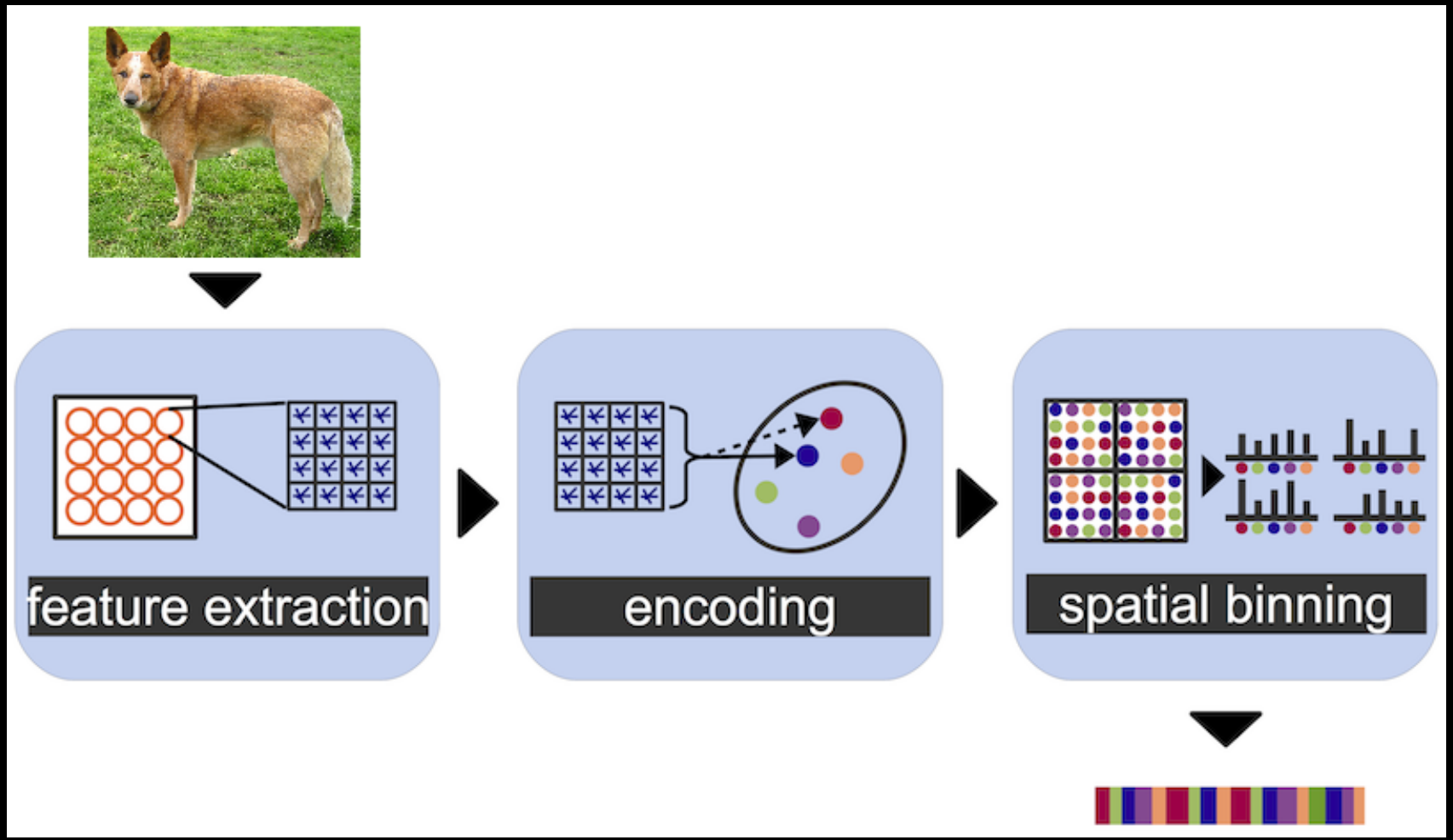
That's where perception kicks in!



(Bruni et al. 2012, 2014)

Image-based space

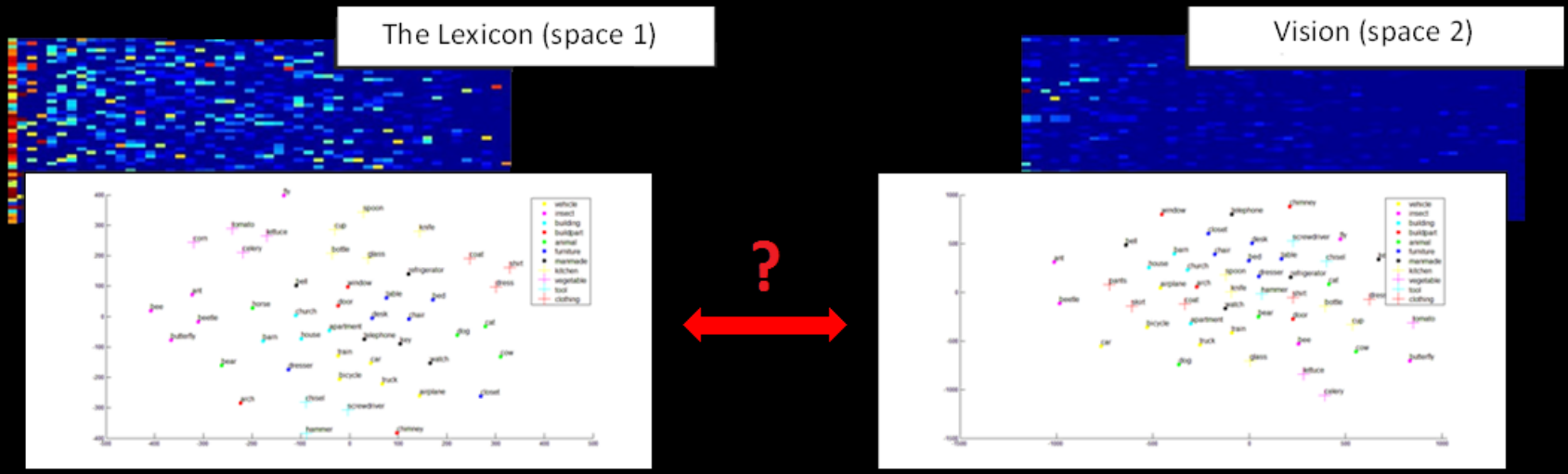
A vector space for visual meaning



(Bruni et al. 2012, 2014)

Relating spaces

comparing and combining



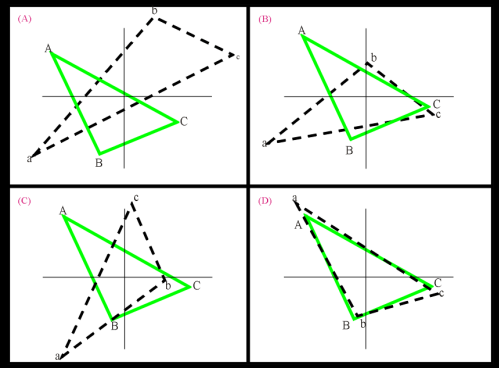
...same for any other possible perceptual modality...

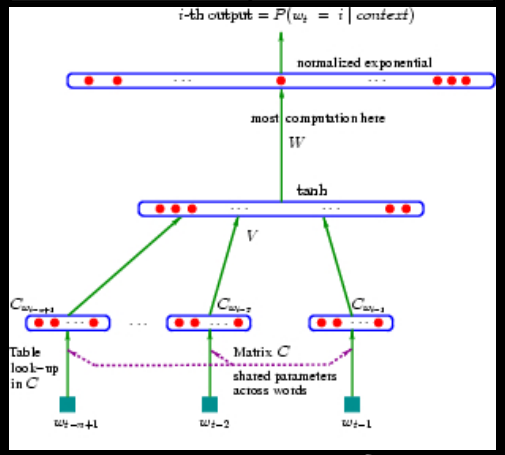
a problem of cross-modal learning (?)

Relating spaces

some solutions from the IT department

- linear mapping based on geometrical transformation
- linear similarity analysis based on correlation (eg: RSA)
- learning joint representations using Artificial Neural Networks
- ...

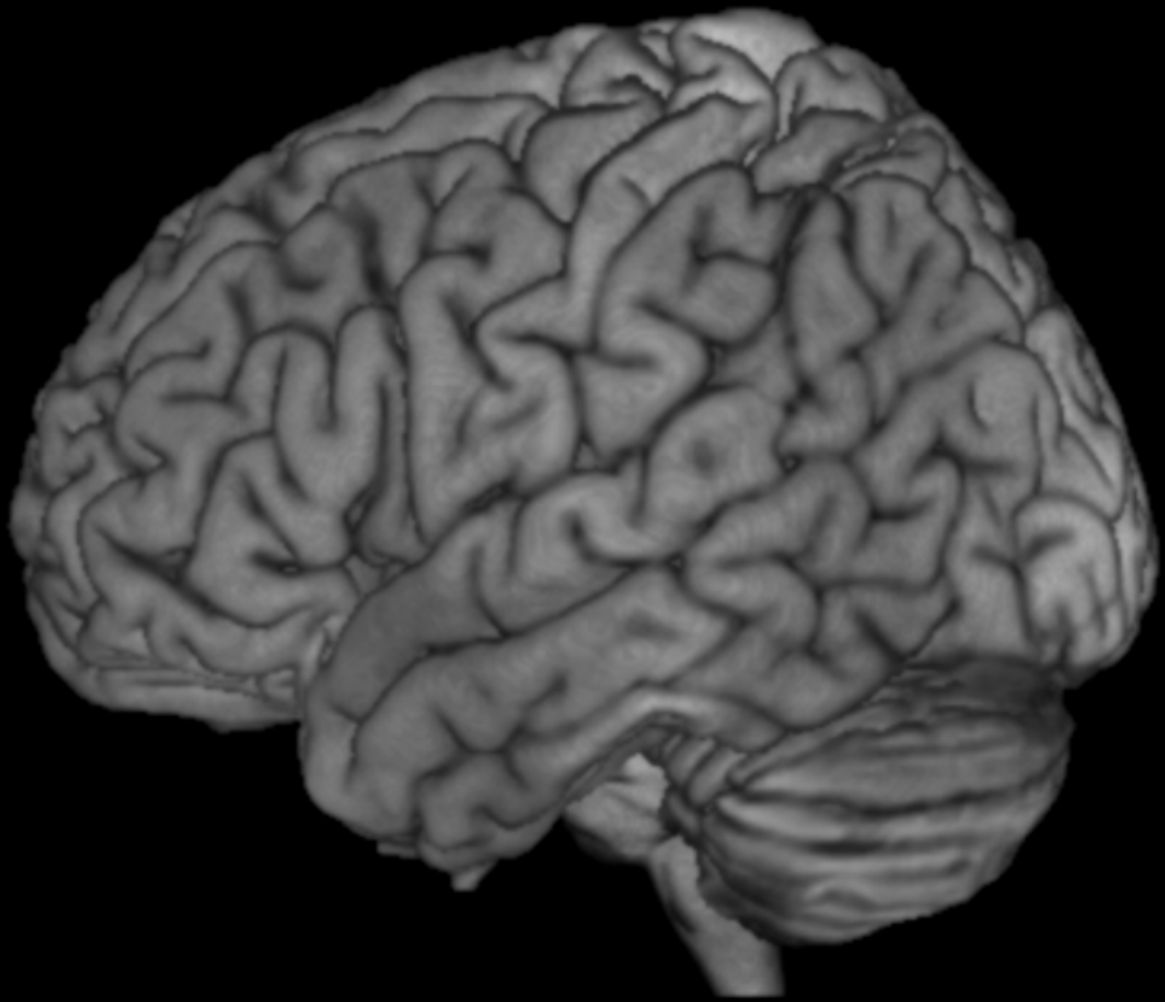




many possible approaches, but are they *good* models of cognition?... and the brain?

Semantics in the brain

the point of reference for the reference machine

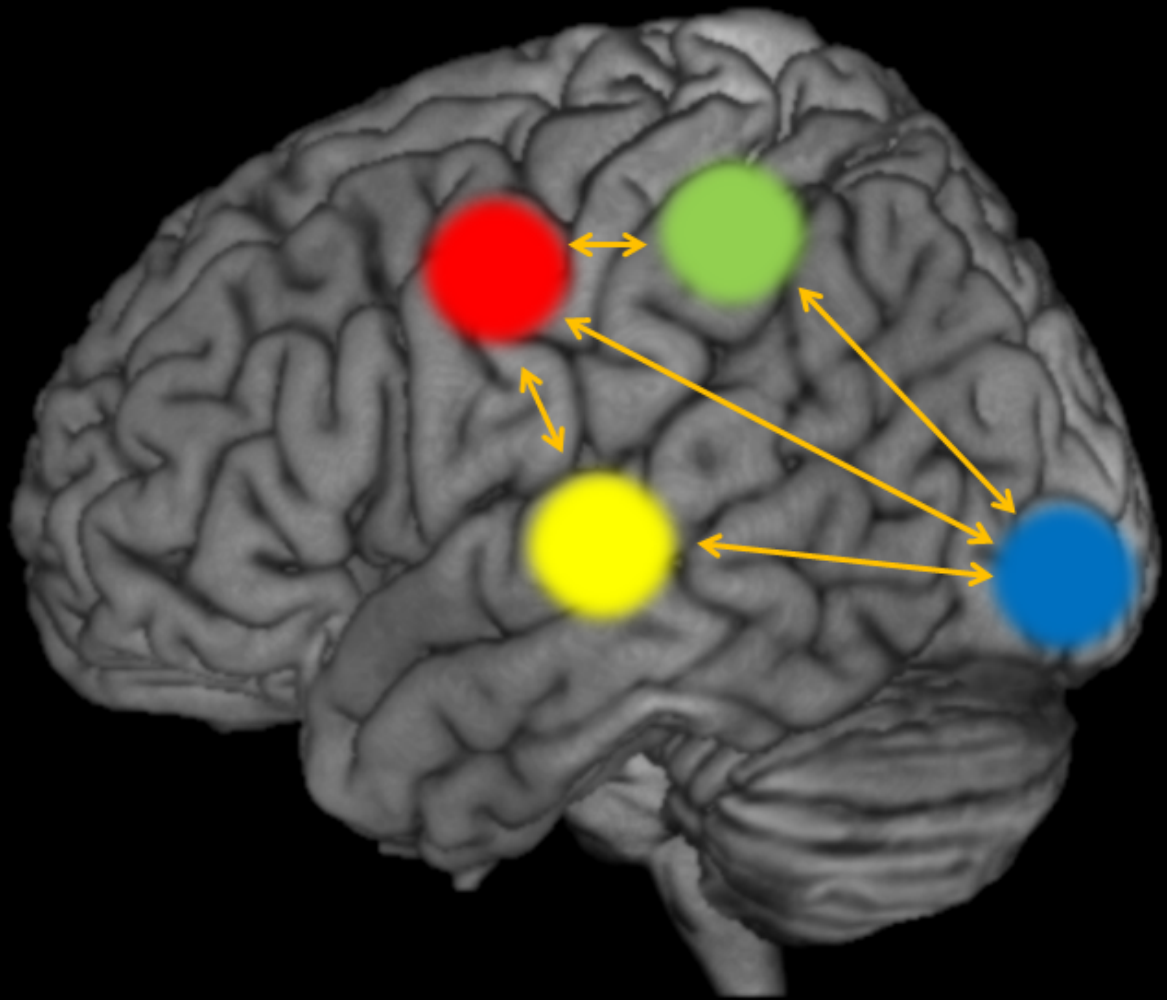


Semantics in the brain

modal areas

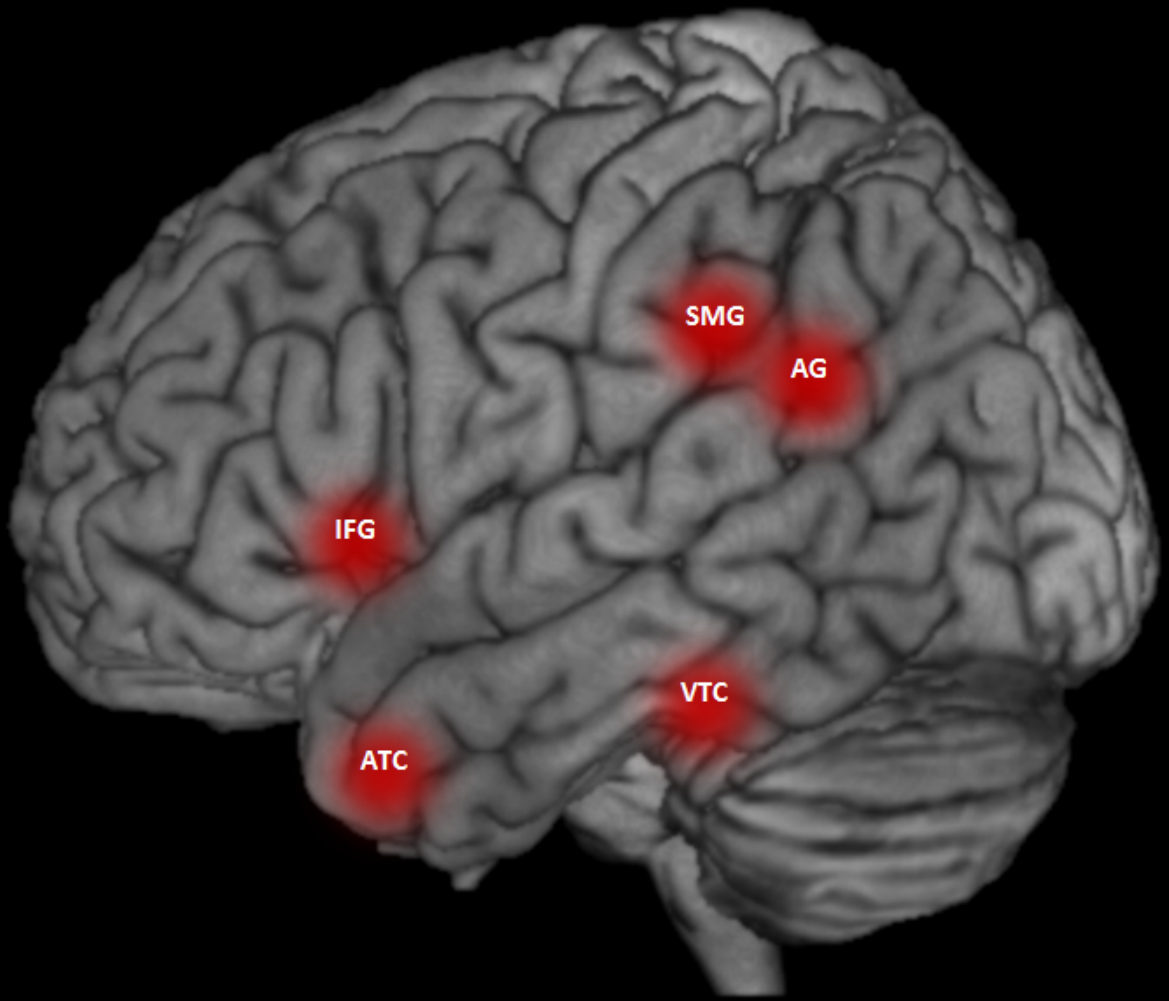
Semantics in the brain

multimodal integration



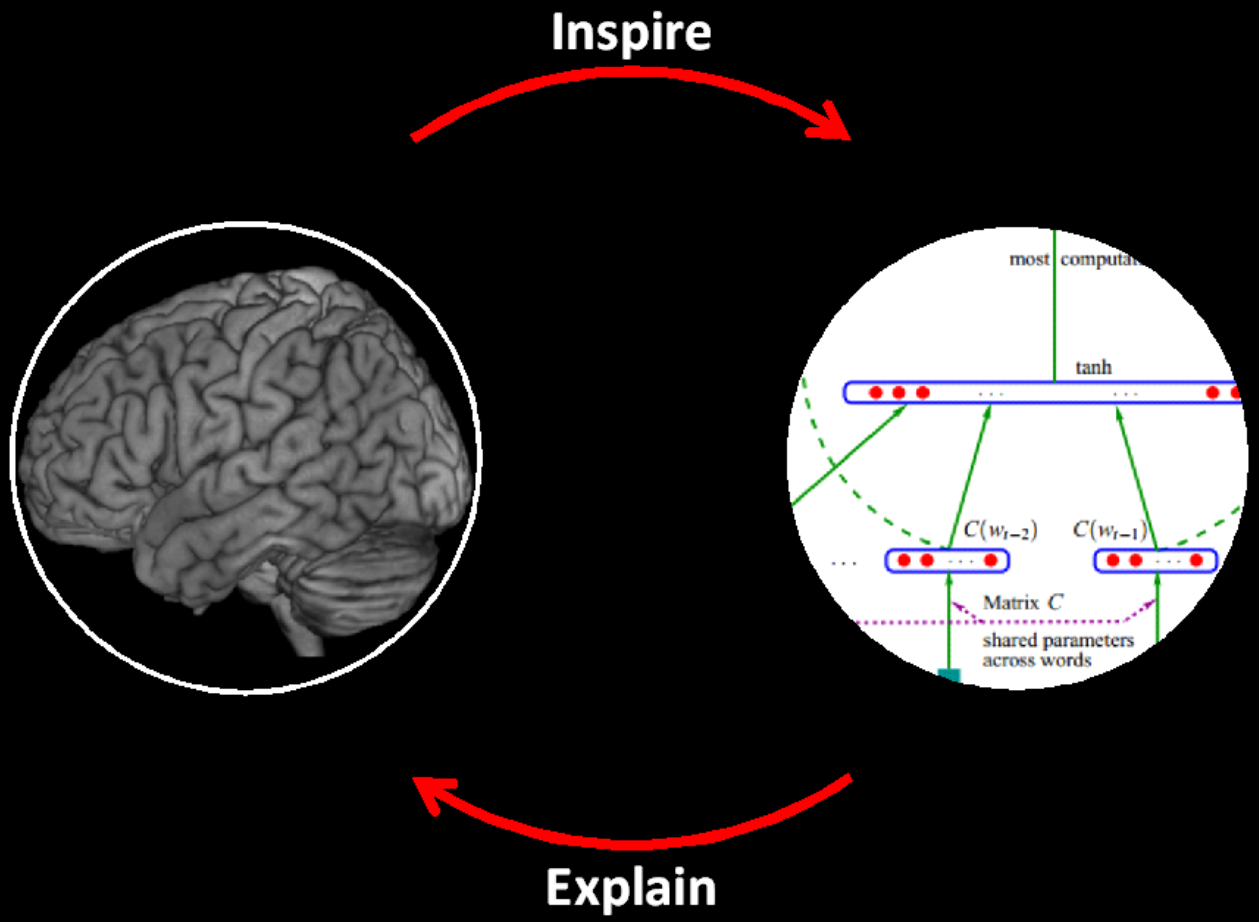
Semantics in the brain

language network



Semantics in the brain

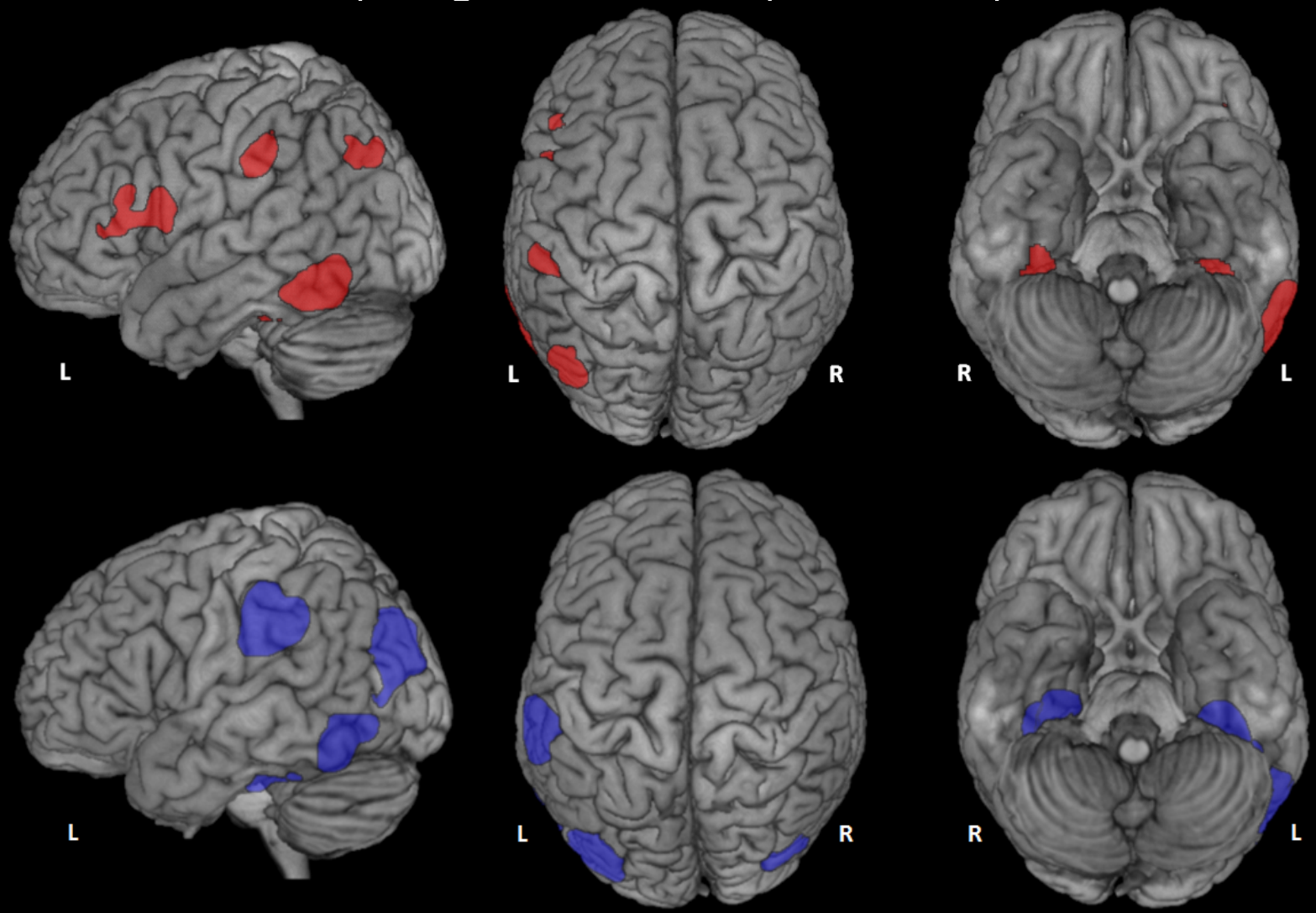
Why should we bother?



(Mitchell et al. 2008, Devereux et al. 2010, Chang et al. 2011)

Semantics in the brain

comparing brain and computational spaces



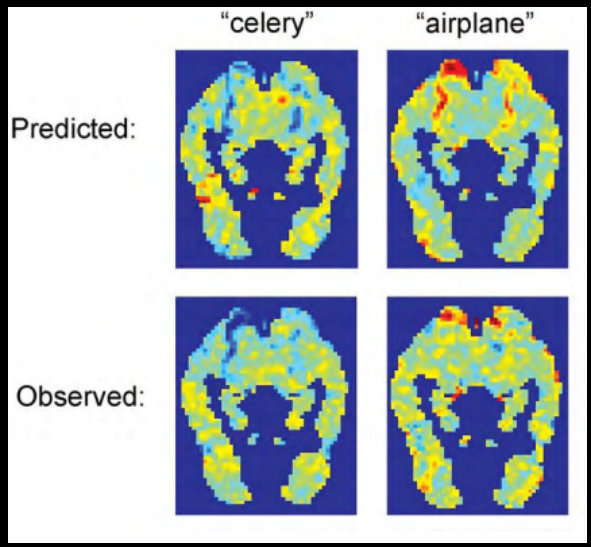
(Anderson et al. 2013)

Semantics in the brain

Open questions

Multimodality: are multimodal models *similar* to convergence zones?

Embodiment vs. disembodiment: is all about perception and integration or is there something else? Can lexical models tell something about it?



(Mitchell et al. 2008)

Part IV

Current and future work

Current work: sound

- SoundNet: creating a database of sounds
- Sound-based semantic modeling

Sound

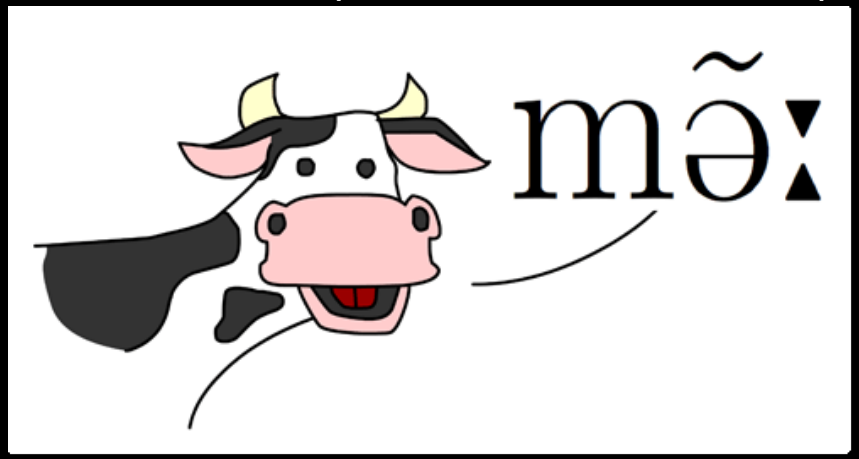
What do we know?

- There are many words for sounds:
  susurrus, blare, zing, clop, honk, sough, glug, splosh, blow, blast, tweet, tinkle, clitter, stridulate, thump, snap, whack, clank, skreak, crash, patter, song, beat, burble, clang, whistle, chirrup, snarl, strum, squeak, tick, knock, rustle, whish, plunk, cry, whir, bong, clack, gargle, sigh, crackle, whiz, chime, click, bang, purl, squelch, crunch, twang, bleep, grumble, step, murmur, sing, chink, quack, drum, rumble, roar, screech, brattle, buzz, clump, scranch, ding, splash, thrum, tap, ring, ting, beep, zizz, swoosh, jingle, bubble, chug, sizzle, hum, crack, mutter, toot, vroom

Sound

What do we know?

- There are many words for sounds
- Sounds are categorized by:
  - their source (e.g. animal sounds)



Sound

What do we know?

- There are many words for sounds
- Sounds are categorized by:
  - their source (e.g. animal sounds)
  - the action producing them (*hit*, *thump*)
  - their structure (next slides)

Sound structure

'vertical'

parameter	value	explanation	examples
intensity	high	wide amplitude	bang, roar
	low	narrow amplitude	rustle, crackle
timbre	bright	high pitch or high overtones	click, clink
	dark	low pitch or low overtones	clack, clank
quality	musical	periodic vibration	ring, whistle
	noise	aperiodic vibration	hiss

(Lehmann 2003)

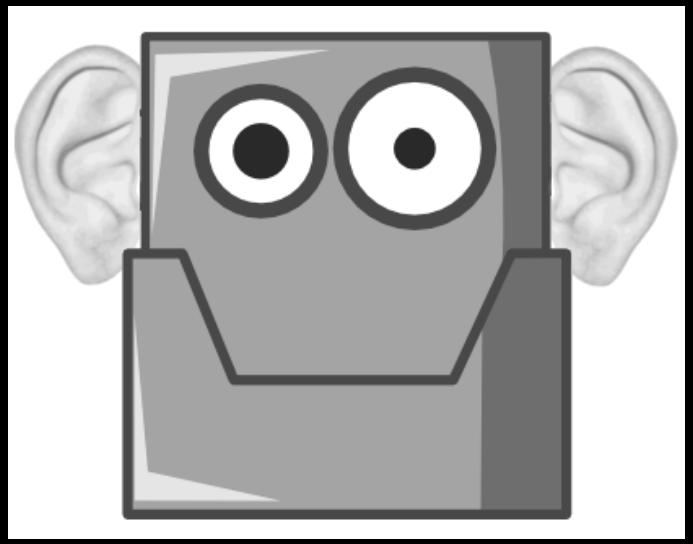
Sound structure

'horizontal'

temporal structure	examples
single impulse	crack, bang, clang
single impulse, repeated by default	bark, knock, ring
iteration of distinct instances	cluck, clack
high iteration frequency	rattle, jingle, crackle
vibrating	buzz, drone, creak
continuous, simplex	hiss, whistle

(Lehmann 2003)

Going digital



SoundNet

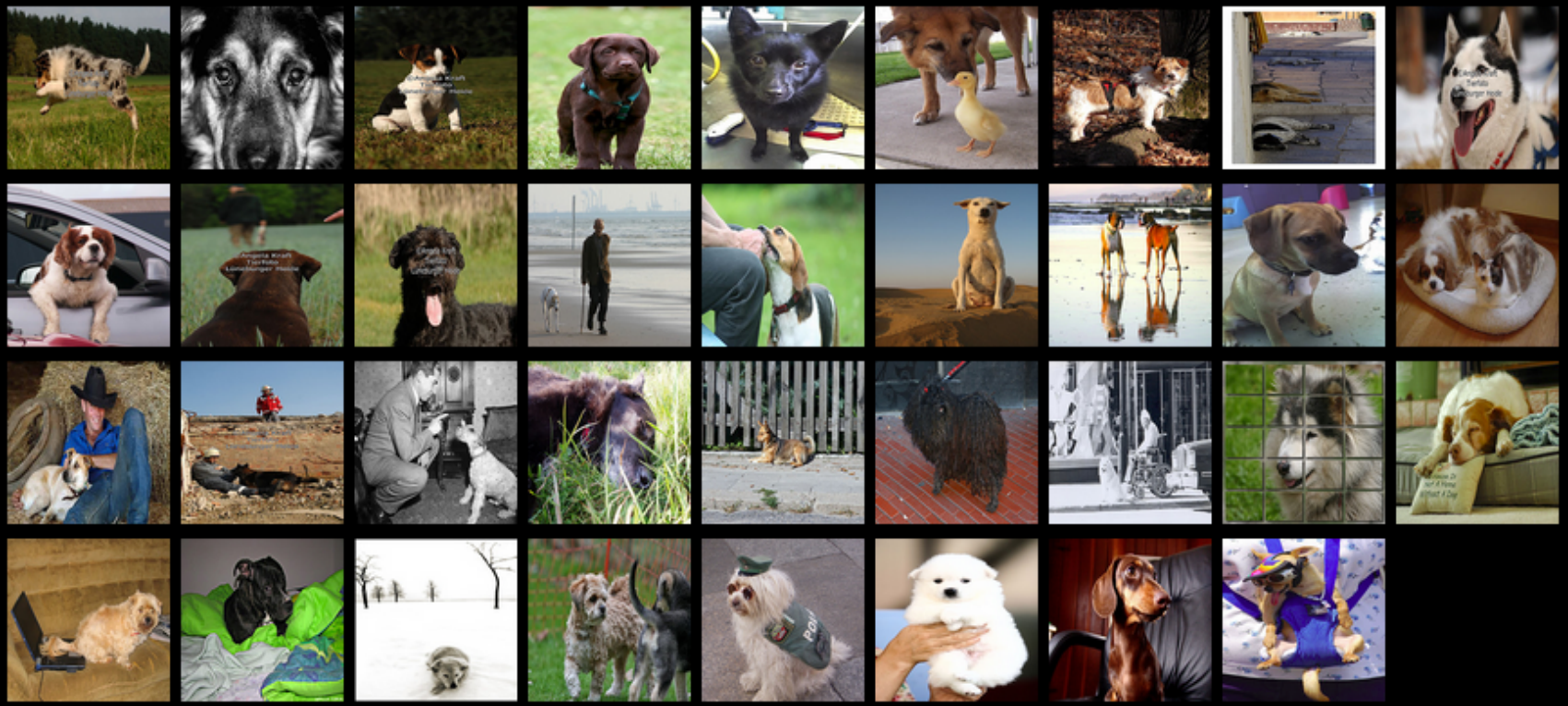
A linguistically annotated sound database

Goals:

- Get a representative set of sounds
- Provide uniform annotations
- Link sounds to related (WordNet) concepts

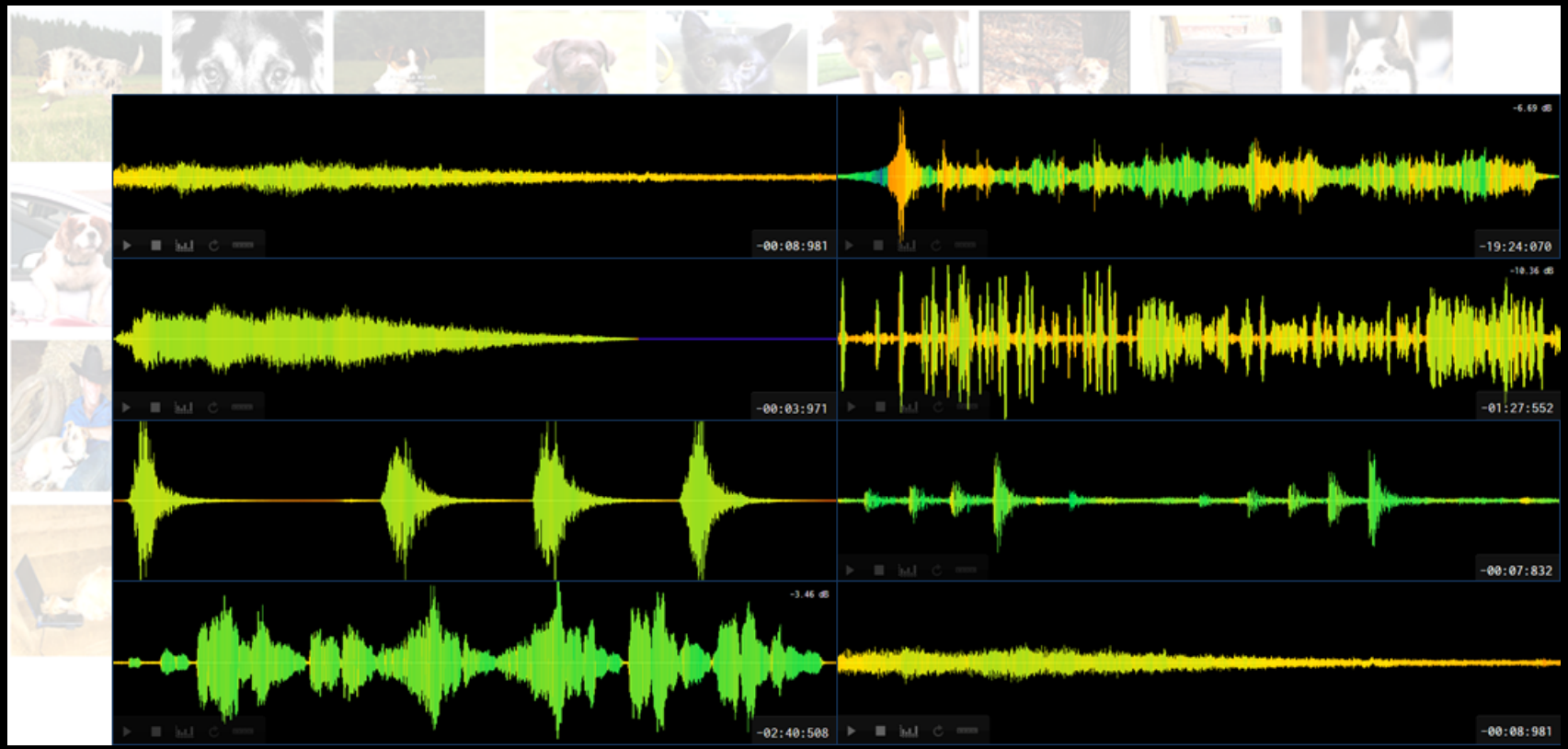
A sound-based space

images...



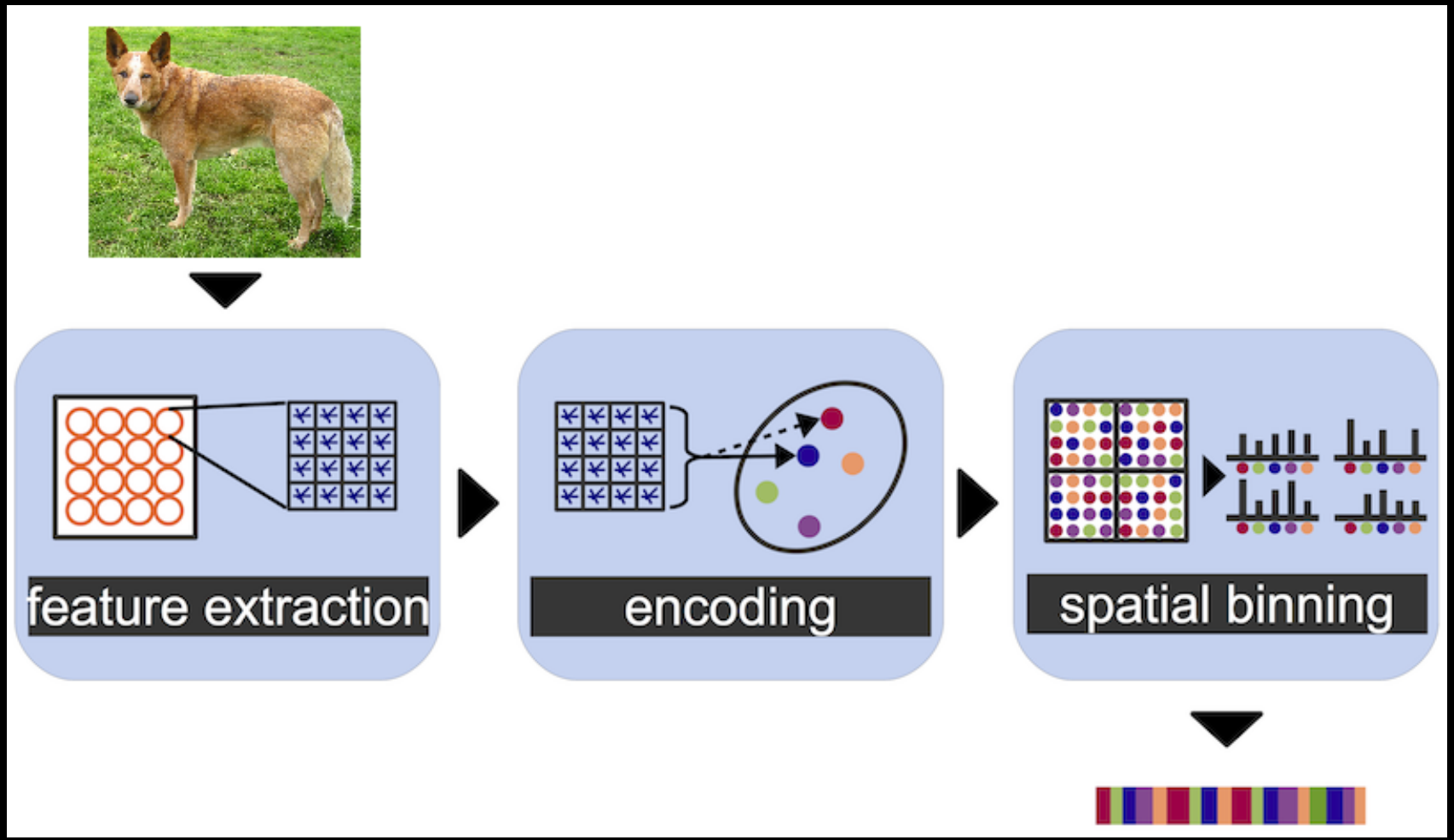
A sound-based space

and now sounds!



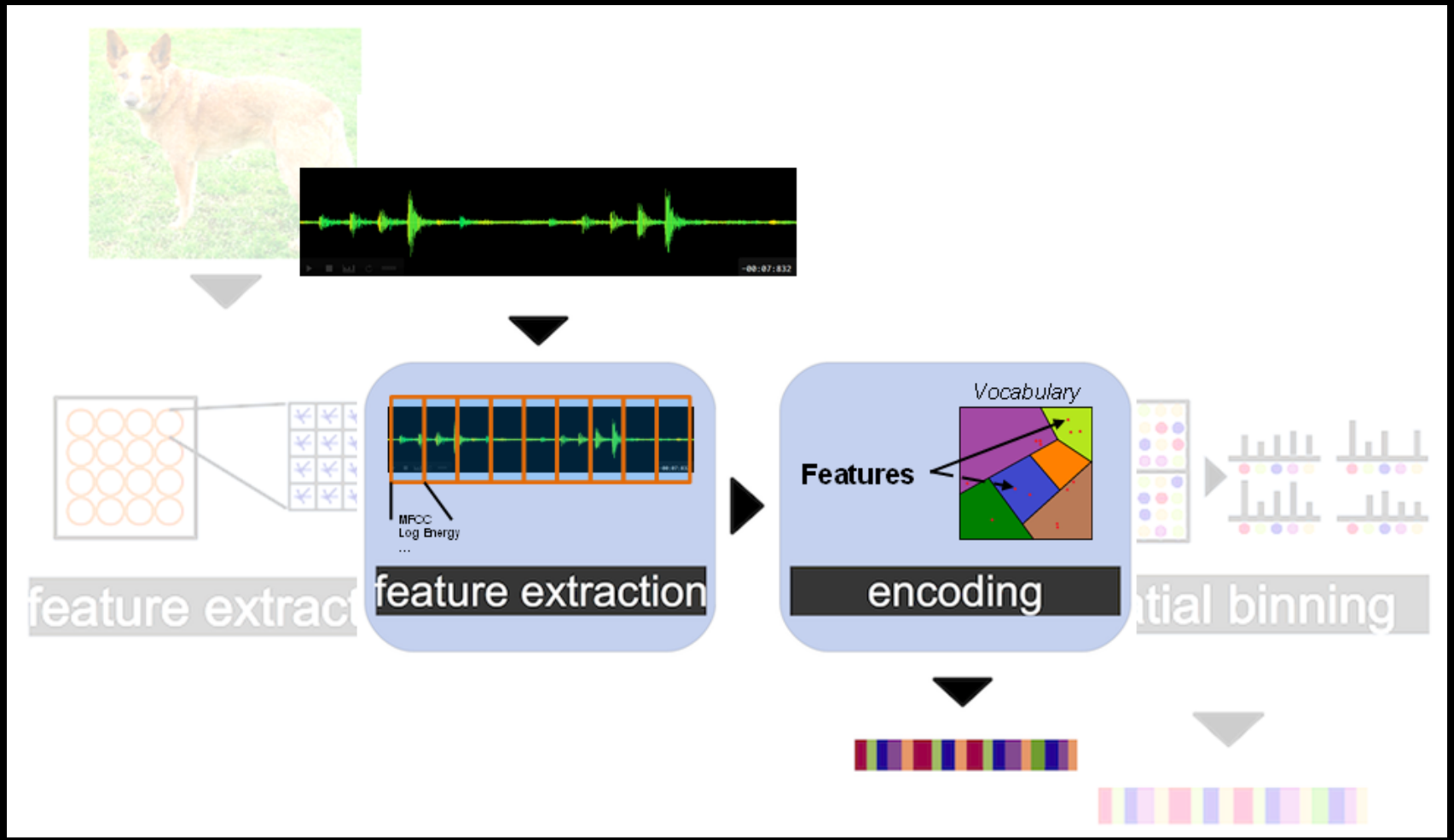
A sound-based space

images...



A sound-based space

and now sounds!



Plans for the future

- Video for multimodal semantics of motion and action
- Predicting concreteness:
  - from text
  - from multimodal distributional semantics
  - from the brain

Thank you!

References I

Anderson, Andrew J., Elia Bruni, Ulisse Bordignon, Massimo Poesio & Marco Baroni. 2013. Of words, eyes and brains: Correlating image-based distributional semantic models with neural representations of concepts. In *Proceedings of the 2013 conference on empirical methods in natural language processing*, 1960–1970. Seattle, Washington, USA: Association for Computational Linguistics. http://www.aclweb.org/anthology/D13-1202.

Baroni, Marco & Alessandro Lenci. 2010. Distributional memory: A general framework for corpus-based semantics. *Computational Linguistics* 36(4). 673–721.

References II

Bruni, Elia, Gemma Boleda, Marco Baroni & Nam-Khanh Tran. 2012. Distributional semantics in technicolor. In *Proceedings of the 50th annual meeting of the association for computational linguistics: Long papers-volume 1*, 136–145. Association for Computational Linguistics.

Bruni, Elia, Nam Khanh Tran & Marco Baroni. 2014. Multimodal distributional semantics. *Journal of Artificial Intelligence Research* 1. 1–47.

Chang, Kai-min, Tom Mitchell & Marcel Just. 2011. Quantitative modeling of the neural representation of objects: How semantic feature norms can account for fMRI activation. *NeuroImage* 56. 716–727.

References III

Devereux, Barry, Colin Kelly & Anna Korhonen. 2010. Using fmri activation to conceptual stimuli to evaluate methods for extracting conceptual representations from corpora. In *Proceedings of the naacl hlt 2010 first workshop on computational neurolinguistics* CN '10, 70–78. Stroudsburg, PA, USA: Association for Computational Linguistics. `http://dl.acm.org/citation.cfm?id=1866686.1866695`.

Gärdenfors, Peter. 2000. *Conceptual spaces: The geometry of thought*. MIT press, Cambridge MA.

Gärdenfors, Peter. 2004. Conceptual spaces as a framework for knowledge representation. *Mind and Matter* 2(2). 9–27.

Harnad, Stevan. 1990. The symbol grounding problem. *Physica D: Nonlinear Phenomena* 42(1-3). 335–346.

Kempton, Willett. 1984. Interview methods for eliciting fuzzy categories. *Fuzzy Sets and Systems* 14(1). 43–64.

References IV

Landauer, Thomas & Susan Dumais. 1997. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review* 104(2). 211–240.

Lehmann, Christian. 2003. On the linguistic categorization of sounds. Available at `http://www.christianlehmann.eu/publ/Ling_cat_sounds.pdf`.

Lund, Kevin & Curt Burgess. 1996. Producing high-dimensional semantic spaces from lexical co-occurrence. *Behavior Research Methods* 28. 203–208.

Mitchell, Tom, Svetlana Shinkareva, Andrew Carlson, Kai-Min Chang, Vincente Malave, Robert Mason & Marcel Just. 2008. Predicting human brain activity associated with the meanings of nouns. *Science* 320. 1191–1195.