

# Denotation & Compositionality

Peter Bosch

Institute of Cognitive Science

University of Osnabrück

Workshop 'Semantic Theory and Empirical Evidence'  
Osnabrück, 18. - 19. Sept., 2009

## Why compositional formal semantics for natural language?

because it holds the promise of helping to understand linguistic productivity:

*how can speakers of a language produce and understand previously unheard complex expressions in their language?*

Neuroscience & psychology have no comparable theory. Notions of wellformedness, interpretability, productivity – central to semantics – are alien to them.

Which makes them miss important properties of cognition.

## Empirical basis

"It has always been recognized that model-theoretic constructs cannot be identified with units of psychological processing [...] nor can translations into a formal language that serve as the (dispensable) intermediary correspond to what is in the head either; ultimately the only empirical test of a model-theoretic account of natural language semantics is the characterization of entailments among sentences it gives."

(Dowty 2007:40)

## A problem about observation

Truth and entailment are tied to propositions

- the relation between propositions and *observable utterance events* is unclear.
- it is clear from experimental work that already words, morphemes, even phonemes are in a clear sense *meaningful*, i.e., *have observable and immediate effect on action and perception*.
- How can we relate such experimental findings to formal semantic theories, which are about proposition?

A couple of illustrations from our own experimental work:

# Eye tracking for reference resolution (1)

Karabanov, Bosch, König 2007

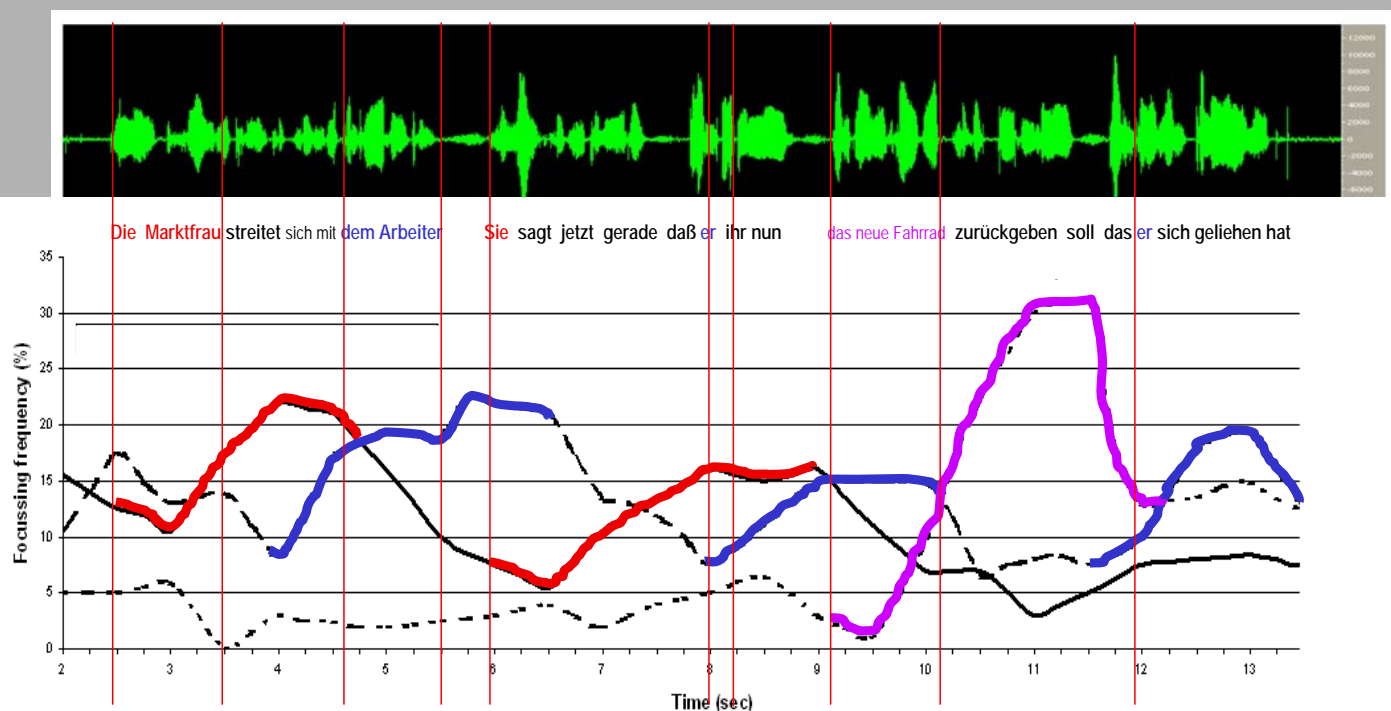


Heute ist Markt im Dorf. Die Marktfrau streitet sich mit dem Arbeiter. Sie sagt jetzt gerade, dass er ihr nun das neue Fahrrad zurückgeben soll, das er sich geliehen hat.

It's market day in the village. The market woman is quibbling with the worker. She's just saying that he should give the new bike back to her that he borrowed.

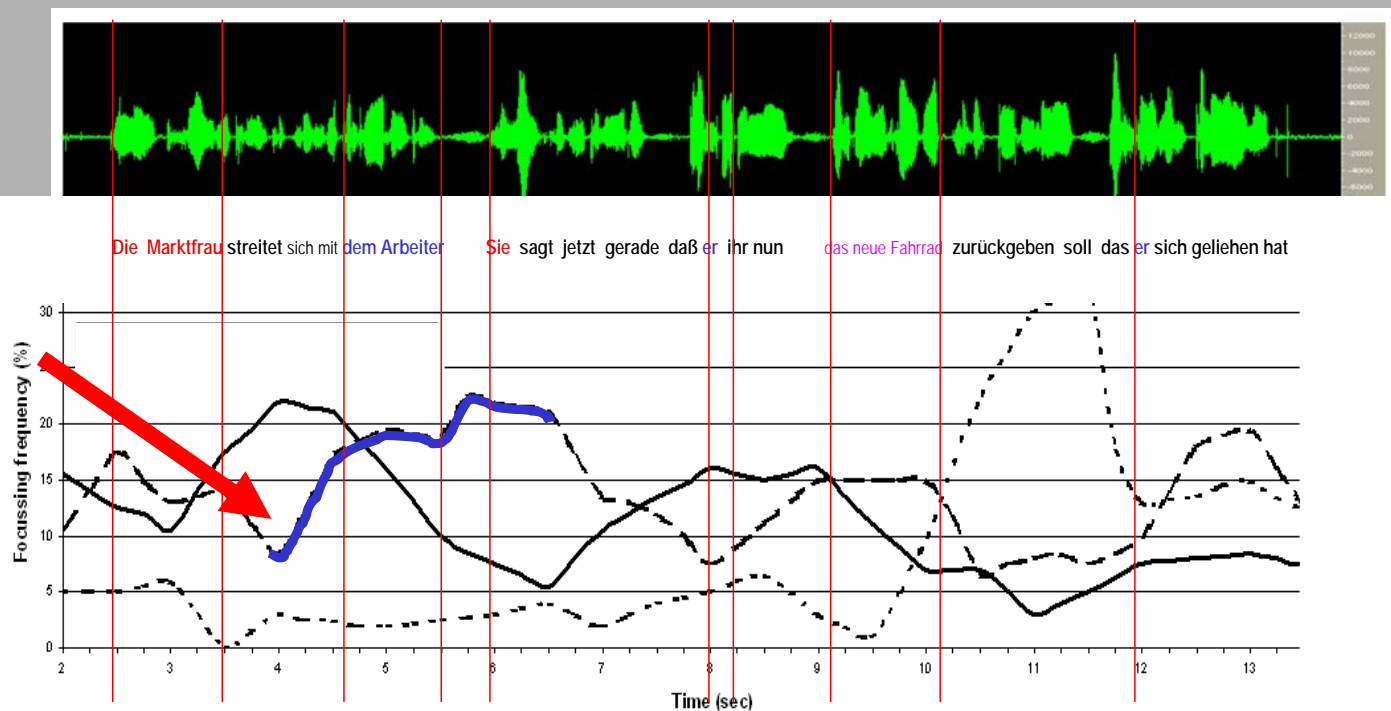
# Eye tracking for reference resolution (2)

Karabanov, Bosch, König 2007



# Eye tracking for reference resolution (2)

Karabanov, Bosch, König 2007



no referent anticipation plausible  
unless "streitet" has already been understood

## Processing effects of the definite determiner?

German default descriptions:

*der*<sub>[masc]</sub> *gelbe Hut*<sub>[masc]</sub>

*the yellow hat*

*die*<sub>[fem]</sub> *blaue Rakete*<sub>[fem]</sub>

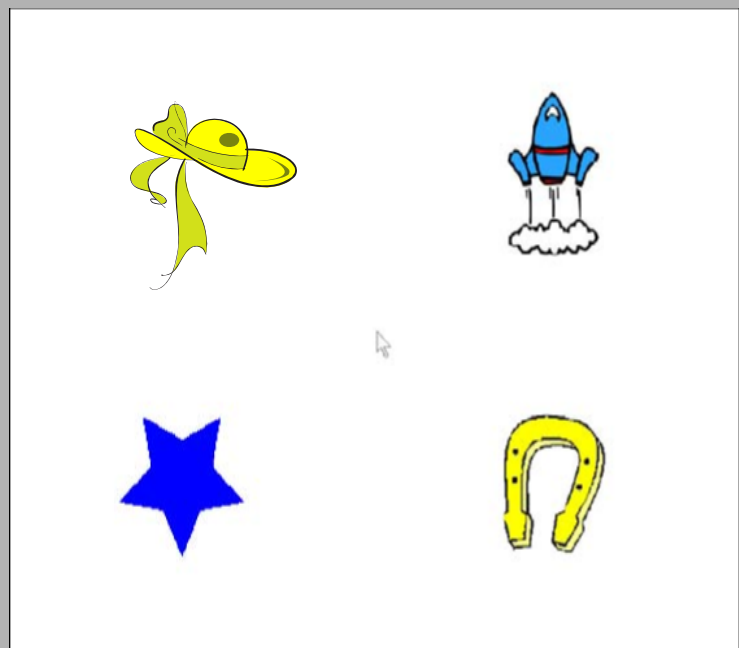
*the blue rocket*

*der*<sub>[masc]</sub> *blaue Stern*<sub>[masc]</sub>

*the blue star*

*das*<sub>[neut]</sub> *gelbe Hufeisen*<sub>[neut]</sub>

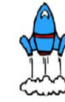
*the yellow horse shoe*



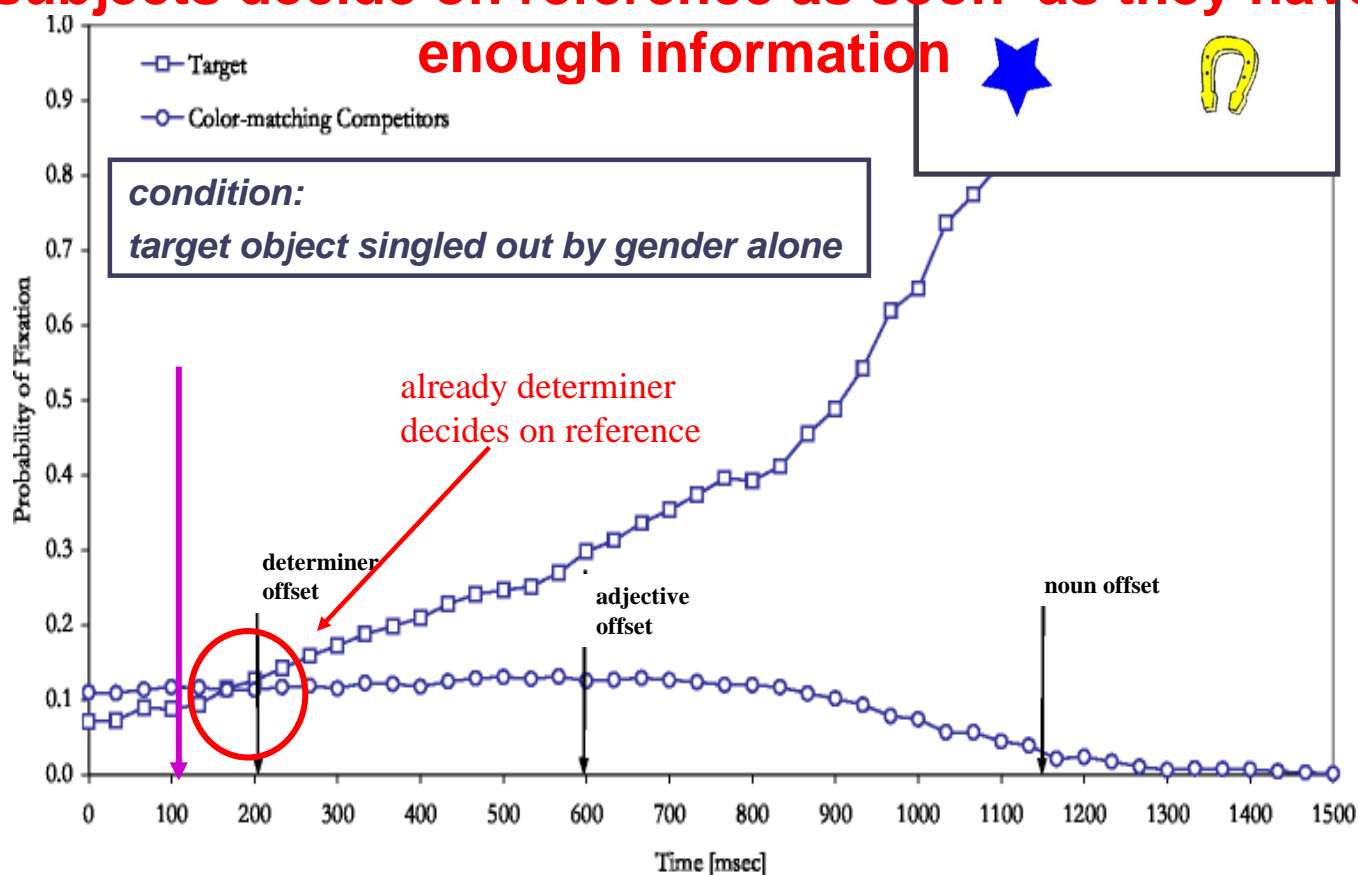
*Klicken Sie auf die blaue Rakete.*

*click on ... [followed by a def. determiner, adjective, and noun]*

# Klicken Sie auf die blaue Rakete



subjects decide on reference as soon as they have enough information



## summing up some experimental findings

The meaningfulness of the linguistic sign in utterance events is characterized by

### Crossmodality

inclusion of information from all sources available

### Incrementality

successive use of all information in the sequence in which it becomes available, to determine current processor action

### Immediacy

immediate use of all information that can be used

## too little or too much in linguistic knowledge?

The effects we saw cannot be attributed to linguistic knowledge alone, but result from an interaction with other resources.

Determiner:

- semantics: unique reference
- morphology: gender
- vision: recognition of objects in domain
- ling. experience: preferred nominals

Anticipated referent identification:

- preceding referent identification
- argument domain of verbal concept
- interpretation of visual scene

How far can we get with purely linguistic knowledge?

How does it interface with with other cognitive functions?

## too little or too much in linguistic knowledge?

The problem is familiar e.g. from polysemy.

Suppose you want to explain (Dowty style?) why one would conclude from an utterance of *the clock is running* that *the clock is working*, but from an utterance of *Pete is running* one would not conclude that *Pete is working*.

If you credit the lexicon with sufficiently many different entries for *run* and *work*, and code semantic differences between *clock* and *Pete* in the lexicon, you may end up putting all possible contexts of use into the lexicon.

Even if this were possible, it would defeat the purpose of an explanation of productivity. And it may even yield wrong predictions when *Pete* is the name of your file server.

## Denotation instead of meaning

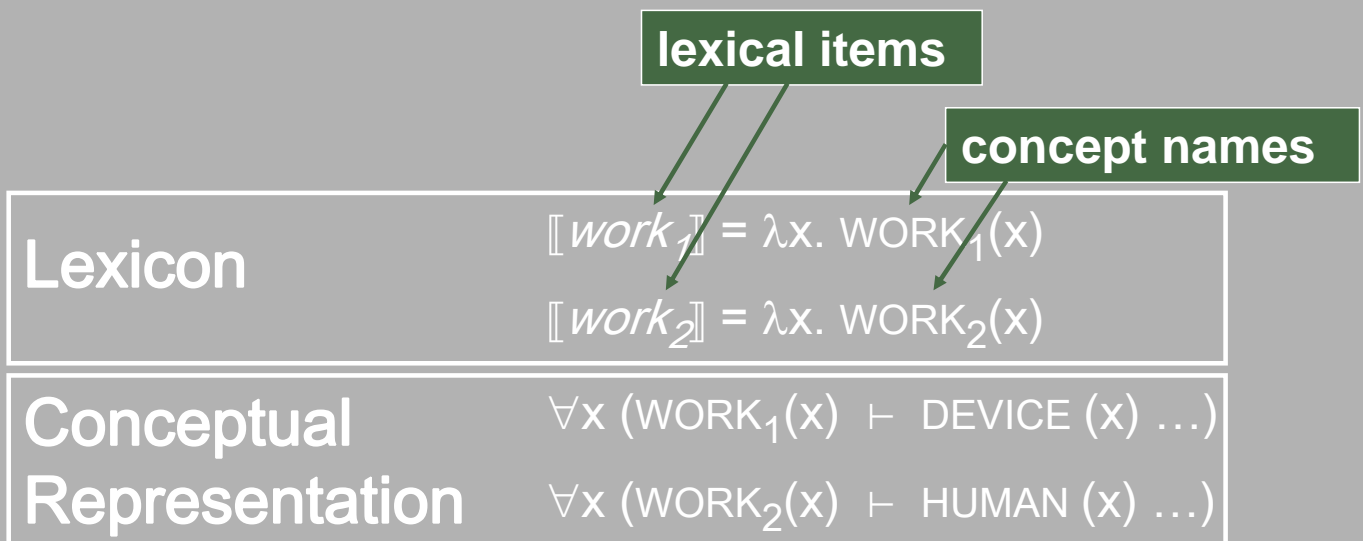
What is needed, is not inflated linguistic knowledge, but information on intended reference, or *denotation*.

Quite possibly we can have a compositional account for phenomena of polysemy

- by composing not *lexical meanings* of constituents,
- but their *denotations* – which would include all relevant knowledge concurrently available.

## Lexical disambiguation needs access to denotations

14



## Denotation instead of meaning

Denotation, however, unlike lexical meaning, varies with the utterance context.

David Kaplan introduced this idea for indexical expressions: Although the lexical meaning (character) of an indexical is constant, its denotation (content) may be different in different utterance contexts.

The same holds not only for Kaplanian indexicals but also for *other expressions that contain a variable in their lexical entry that is evaluated at the utterance context*, such as perspectival expressions (*nearby, enemy*) or comparative adjectives (*long, fast*) and their variables for perspective or comparison class or standard of comparison.

## Contextual Concepts as denotations

I call the denotation of an indexical expression (in this extended sense) in a context a

### *Contextual Concept (CC)*

CCs are partial truth functions, defined only for a narrow domain within a context. - Also Frege called the Bedeutung (denotation) of a predicate a Begriff (concept)

We may represent CCs by feature structures, much like the CONT or SEM feature in HPSG's lexical entries, only more specified.

The concept names in these structures are read as pointers to a general ontology in the style of description logics.

## Polysemy as context dependence

Next to indexicality proper, which depends on variable saturation, also other expressions display a dependence on utterance context, typically in the form of a *dependence on denotations of their arguments*.

The polysemy of *run* and *work* is an example.

Also here the denotations are CCs that include all concurrently available knowledge from the utterance context.

## Polysemy as context dependence

If an the question

*Do you know where Peter is?*

is answered by

*Peter is working.*

the CC denoted by *is working* is assumed to give information on a location.

If the same sentence is uttered in reply to

*How can Peter afford these expensive holidays?*

the CC would not include location information; here *is working* denotes a CC that must include information on the relation between work and pay.

## CCs are the invariants in VP anaphora

CCs, as VP denotations, are what remains constant in VP anaphora:

*Peter is working and so is Fred.*

cannot be interpreted as saying that one is in paid employment and the other is at his office.

A CC may also make reference to specific referents:

*Peter is working for her and so is Fred.*

means that both work for the same person.

Working for X may be a different concept than working for Y – including whatever relevant knowledge is available.

## Shiftable indexicals

In the CCs that are the denotations of shiftable indexicals we find variables that can be locally bound.

cf. the *Reported Speech Test*, due to Graham Katz:

Ted (in town): *I went to a nearby pub*

Steve (in the country): *I went to a nearby bar*

Reported as: *Ted and Steve each said he went for a drink nearby.*

Peter (Berlin): *It is raining*

Graham (Washington): *It is raining*

Reported as: *Peter and Graham both said that it was raining*

Ted (about his new bike): *It's really fast*

Steve (about his new car): *It's really fast*

Reported as: *Ted and Steve each said he had a fast vehicle.*

## Associative anaphora

So much for an introductory plea for using CCs as an *interface* that integrates constraints from lexical meaning and concurrent non-linguistic knowledge from the utterance context.

-----

I now want to give a more detailed, but still brief, account of associative anaphora – showing that also here the richer representation in terms of CCs is extremely helpful.

## Associative anaphora - the common view

22

Associative anaphora is commonly viewed as a relation between a DP and a discourse referent that was introduced implicitly earlier in discourse.

- (1) A car stopped. The driver got out.  
The engine sounded funny.  
The wheels were all painted white.

The DP *a car* would introduce a discourse referent,

and – somehow implicitly – with it,

further discourse referents, like wheels, an engine, and a driver, which would become available as referents for anaphoric reference by definite NPs.

## Associative anaphora

### - problems with the common view

23

What is nice about this story is

- that it assimilates associative anaphora to regular anaphorically definite reference:

*A car stopped. {The car/ it} had a flat tyre.*

*A car stopped. The driver got out.*

What makes this view wholly unattractive is

- that it assumes the activation of an undefined amount of knowledge that may be connected to a discourse referent introduced by the antecedent expression and thus
- overcrowds the discourse representation with discourse referent artefacts.

## Associative anaphora

### - Functional Concepts

24

Löbner (1998/2003) tried to improve on the common view:

The nominals of the anaphoric DPs, *driver*, *engine*, etc. denote functional concepts, like  $[[\text{driver}]] = \lambda y \lambda x. \text{DRIVER}(x, y)$ , containing an unarticulated argument  $y$ , thus "driver of  $y$ ".

Also by Katz's Reported Speech Test, we can show that there is a bindable unarticulated argument:

A says: *A Toyota stopped and the driver got out.*

B says: *A BMW stopped and the driver got out.*

Reported as: *Both A and B said that a car stopped and the driver got out.*

## Associative anaphora

### - Functional Concepts

25

Our VP anaphora test shows that the denotation of *the driver* in such contexts is a CC:

*PC Plonk talked to the driver after he'd stopped a car on Banbury Road and so did PC O'Booze when he stopped one on Bicester Road.*

meaning that each policeman talked to the driver of the car that he stopped.

The CC thus must include the unarticulated bound variable.

## Associative anaphora

### - a problem?

26

Löbner's story thus sounds like a good story.

But there must be more to it, because of examples like the following...

(2) A Lotus 2-Eleven was sold on ebay.

b. The price was ridiculous.

b'. #The sound was terrific.

(3) A Lotus 2-Eleven was passing.

b. #The price was ridiculous.

b'. The sound was terrific.

just in case you don't know the Lotus 2-Eleven...



(6) A car was sold.      The price was ridiculous.  
(7) A car was passing.    #The price was ridiculous.

28

Suppose our functional concept is this:

$[[\text{price}]] = \lambda y \lambda x. \text{PRICE}(x, y)$

and this function is applied to the discourse referent introduced by the antecedent, *a car*

- then, in (7), this discourse referent is apparently not in the domain of our function, while in (6) it is
- so, apparently, the discourse referent is represented differently in (6) and (7) - not simply as "a car":

$[[\text{car}]] = \lambda x \in D_e. \text{CAR}(x)$ .

How does this difference come about?

## Domain restrictions

29

The domain of the function `[[price]]` is limited, apparently, to things that are represented as *things that have a price*; call such things *commodities*, i.e., instances of the concept COMMODITY.

- The feature PRICE is a feature *only of concepts that are subsumed by the concept* COMMODITY.

(8) A car was sold.      #The sound was terrific.  
(9) A car was passing.      The sound was terrific.

30

Similarly for the SOUND *concept*.

Also the function  $\lambda y \lambda x. \text{SOUND}(x,y)$  is defined only for a specific domain: let's suppose that only *physical events* "have" a sound.

- The function SOUND is a feature only of concepts that are subsumed by the concept PHYS\_EVENT.

# Differences in representation

31

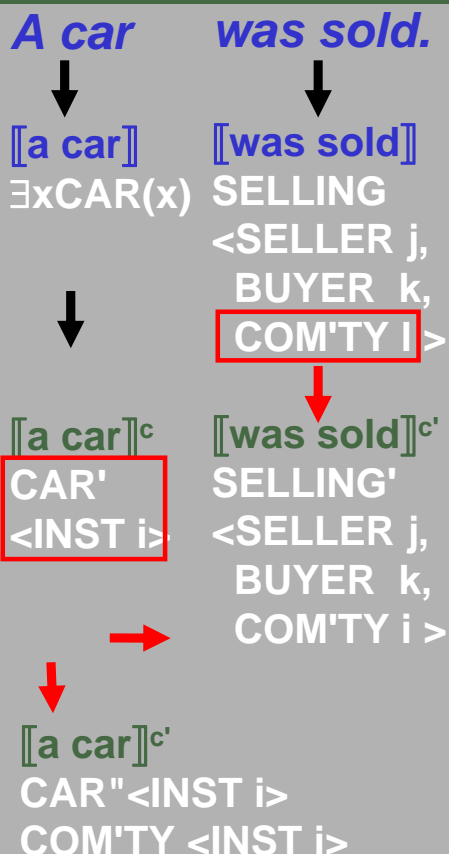
The representation of our discourse referents then comes about in two steps:

- our referent is introduced by the DP *a car* as a plain instance of the sort CAR
- the VP *was sold* modifies this representation and subsumes our discourse referent under COMMODITY, or PHYS\_EVENT or some other concept
  - which follows from a constraint on arguments of the VP *was sold*.

In detail...

## Incremental construction of a representation A car was sold. The price...

32



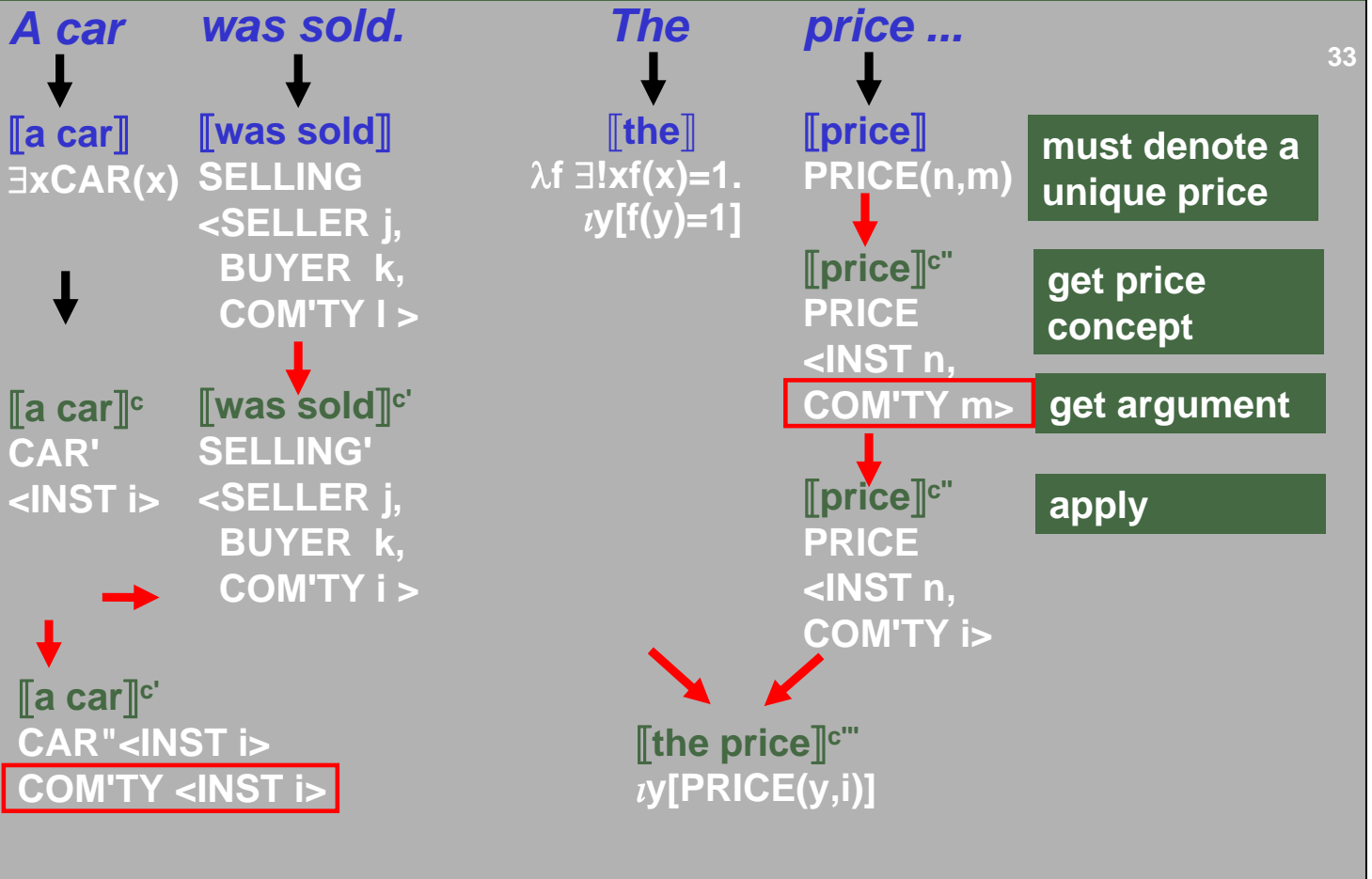
### unify spec & arg3

CAR' <INST i>  
COM'TY <SELLER x, BUYER y,  
                    PRICE z,...>

—————  
CAR'' <SELLER x, BUYER y,  
                    PRICE z, ..., INST i>

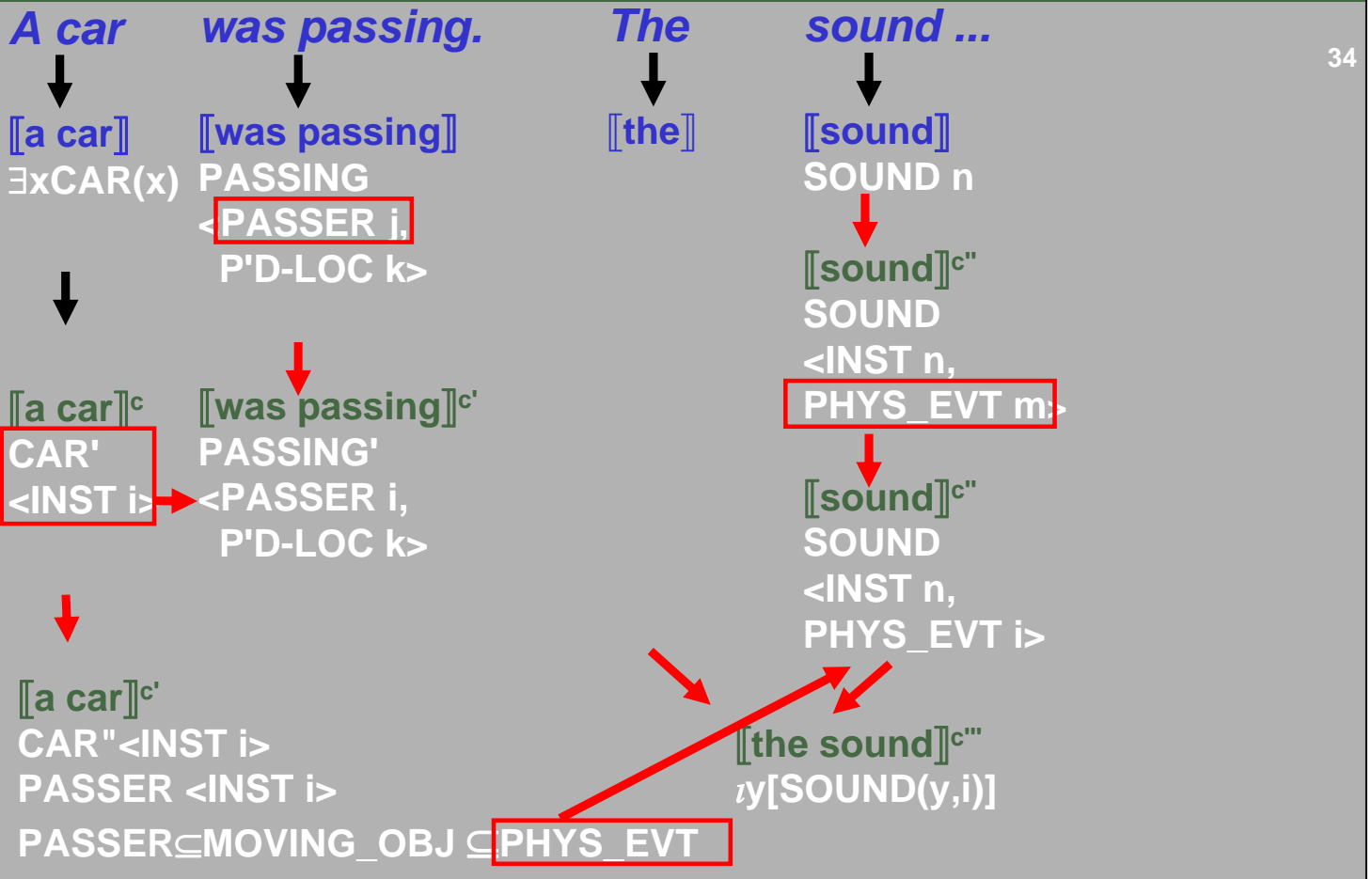
# Incremental construction of a representation

## A car was sold. The price...



# Incremental construction of a representation

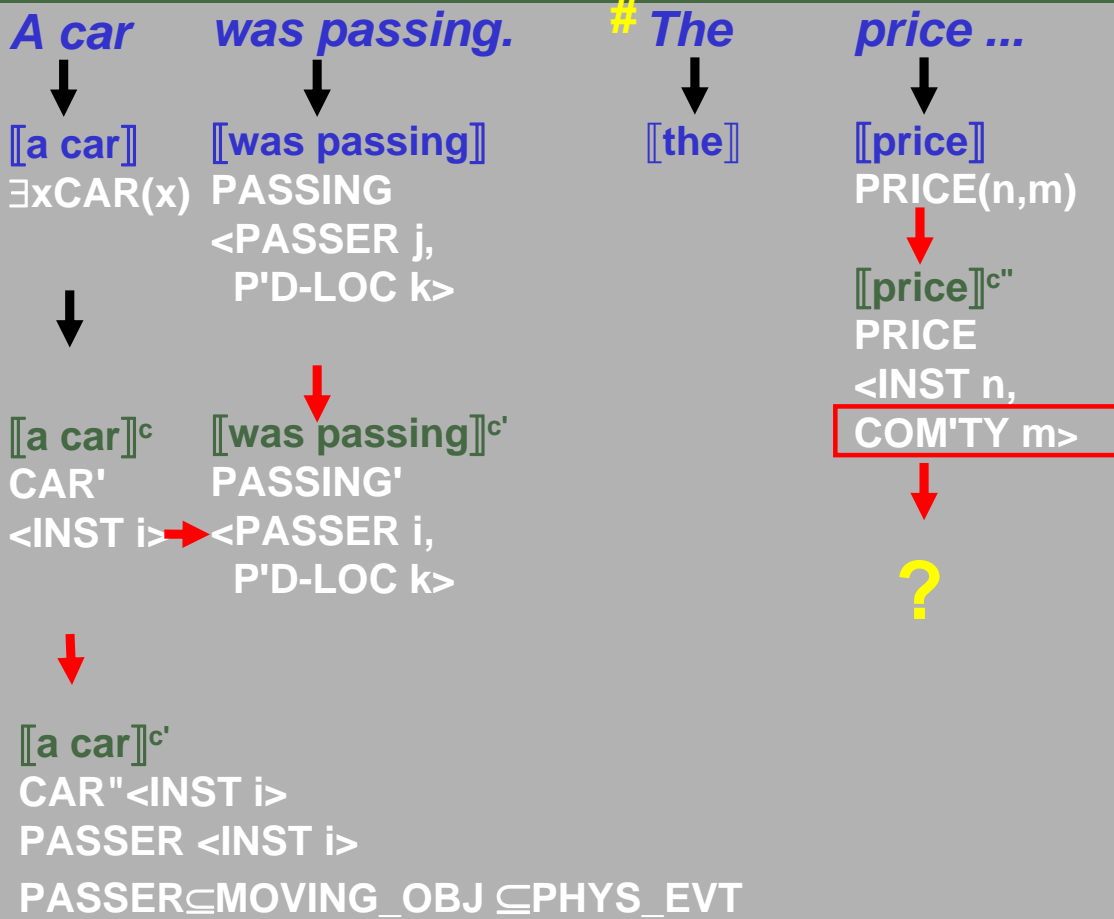
## A car was passing. The sound...



# Incremental construction of a representation

## A car was passing. #The price...

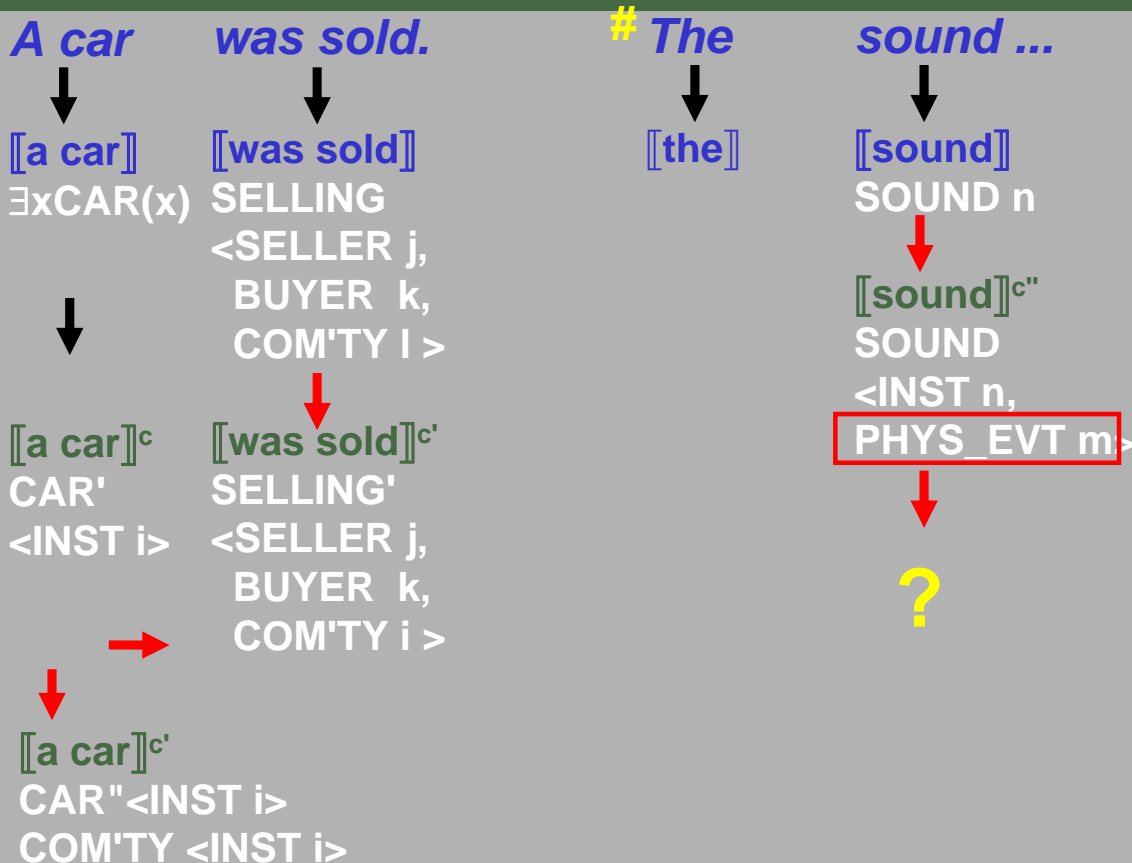
35



# Incremental construction of a representation

## A car was sold. #The sound...

36



## Changes in subsumption relations

37

Our discourse referent is an instance of CAR in any case, but - sometimes - also of

COMMODITY, or  $PASSER \subseteq MOVING\_OBJ \subseteq PHYS\_EVT$

The  $[[price]]$  and  $[[sound]]$  functions are defined for COMMODITY and PHYS\_EVT,

which explains that

- any instance of  $CAR \subseteq COMMODITY$  has a price, and
- any instance of  $CAR \subseteq PHYS\_EVT$  may produce a sound.

## Conceptual representation in discourse: CCs

38

The core of the current proposal is quite simple:

The conceptualization of a referent  $i$  in discourse, a CC,

- ❖ is constrained by the denotation of the DP by which  $i$  is introduced into the discourse
- ❖ and by further information available as, e.g., from predication.

## CCs are also at work in many other cases...

- in pronominal anaphora: *Cut the onion in small pieces. Then fry it.*
- in weak definites: She works in the university/hospital/supermarket. Sie arbeitet im/\*in dem Supermarkt.
- in definites: *Put the cube on the cylinder into the bowl* (in the presence of 2 cubes and 2 cylinders, on cube on one cylinder)
- vagueness: Each application of the general Sorites premise introduces its own comparison class for the gradable ("vague") adjective
- restrictive / non-restrictive modification

and, as already hinted,

- in VP anaphora & VP ellipsis: also the 7 readings of *John revised his paper before the teacher did, and Bill did too*
- in polysemy: *The show/watch/tap... is running*
- in relative adjectives: *Fred's a good doctor* (Fred, qua doctor, is good)

## Following David Kaplan

He distinguishes context of utterance from context of evaluation. He distinguishes, within Fregean *Sinn*, between *character* and *content*. Character being a function from context to content, and content being a function from worlds to extensions.

Surprisingly though, Kaplan exploits his distinction only for what he calls "demonstratives" (*I, you, here, now, that, actual, and present* are some of his examples) and at the same time he believes that there are "demonstrative-free" expressions that "have a constant character, i.e., that *express the same content in every context*" (Kaplan 1978:85)

## ... but not all the way

Given his focus on *propositions and entailment*, Kaplan must stick to the assumption of "demonstrative-free" expressions.

*"Utterances take time, and utterances of distinct sentences cannot be simultaneous (i.e., in the same context). But in order to develop a logic of demonstratives it seems most natural to be able to evaluate several premises and a conclusion all in the same context. Thus the notion of  $\varphi$  being true in [context]  $c$  and [LD structure]  $\mathfrak{A}$  does not require an utterance of  $\varphi$ ." (Kaplan 1978:91, cf. also 1989:584)*

There is no room for incrementality or any other discourse dependence, or dependence on concurrent knowledge. If entailment is the focus, context must remain constant for all propositions involved.

## Diverging from Kaplan

I want to diverge from Kaplan by replacing  
his focus on *propositions and entailment*  
by a focus on *utterances (as events in space and time)*

keeping logic, as it were, in the background and making room for an empirical theory of semantic processing that can, in principle, tackle what we observe as "meaningful" events in the course of utterances.

## Diverging from Kaplan

Nearly all we need to do in order to turn Kaplan's theory of indexicals into a theory of processing is

- ❖ to give up the axiom that the context does not change in the course of an utterance,
- ❖ make it an empirical question whether or not the denotation of a syntactic constituent varies with the utterance context, and
- ❖ include in the representation of context, in our CCs, all knowledge resources that demonstrably contribute to the truth conditions of "what is said".

This may actually pave the way to a compositional semantic theory as the backbone of a theory of utterance processing.



**Thank you!**