
A Structured Context Model for Grammar Learning

Nancy Chang and Eva Mok

University of California, Berkeley

International Computer Science Institute



Everyday language relies heavily on context

- Mother: are **they** clean yet ?
- Eve: **they** clean
- Mother: **they**'re pretty clean .
- Eve: let me wash **them**
- Mother: just wash a little bit right in **there** .

What is Eve washing?

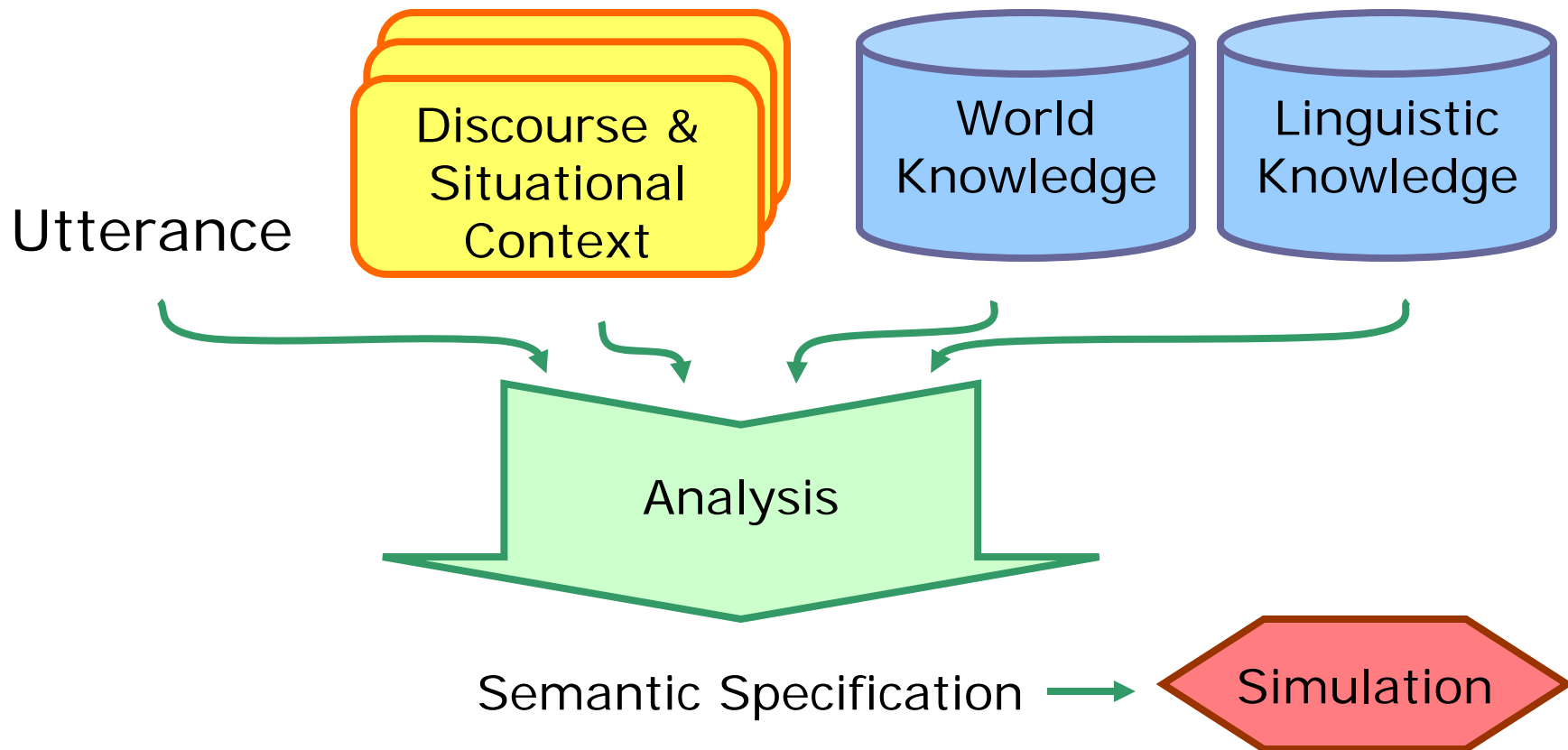
Modeling context for language understanding and learning

- Linguistic structure reflects experiential structure
 - Discourse participants and entities
 - Embodied schemas:
 - action, perception, emotion, attention, perspective
 - Semantic and pragmatic relations:
 - spatial, social, ontological, causal
 - **‘Contextual bootstrapping’** for grammar learning
-

Goal: a structured representation of context

- Compatible with representation of
 - **linguistic knowledge**,
i.e. fits into a construction grammar framework
 - **embodied knowledge**,
i.e. supports simulation-based understanding
 - Consistent with neural representation
 - Computationally precise
-

Understanding an utterance in context: analysis and simulation



The basic unit of discourse is an utterance

The mother says to the child:
"You washed them"

Utterance

Discourse Segment:

DS01 speaker: Mother addressee: Eve attentional-focus: Hands content: { <i>"You washed them"</i> } speech-act: comment

The context model tracks accessible entities, events, and utterances



Discourse:

[Discourse01
participants: Eve , Mother
objects: Hands, ...
discourse-history: DS01
situational-history: Wash-Action]

Each of the items in the context model has rich internal structure

Discourse:

Participants:

Eve

category: child
gender: female
name: Eve
age: 2

Mother

category: parent
gender: female
name: Eve
age: 33

Objects:

Hands

category: BodyPart
part-of: Eve
number: plural
accessibility: accessible

Situational History:

Wash-Action

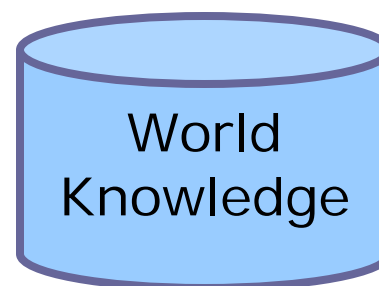
washer: Eve
washee: Hands

Discourse History:

DS01

speaker: Mother
addressee: Eve
attentional-focus: Hands
content: {"are they clean yet?"}
speech-act: question

The analyzer has access to embodied and ontological knowledge



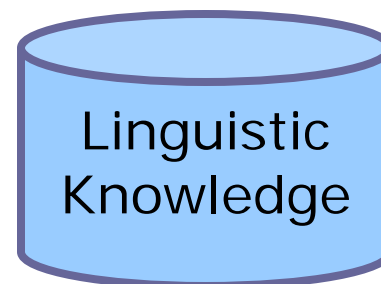
WASH-ACTION

washer

washee



Linguistic knowledge is represented as a set of **constructions**, i.e. <form, meaning> pairs



YOU

form: "you"

meaning: addressee

WASHED

form: "washed"

meaning: WASH-ACTION

Embodied Construction Grammar

(Bergen & Chang, 2005)

More complex constructions have constituent structure

YOU-WASHED-THEM

constituents:

YOU, WASHED, THEM

form:

YOU before WASHED

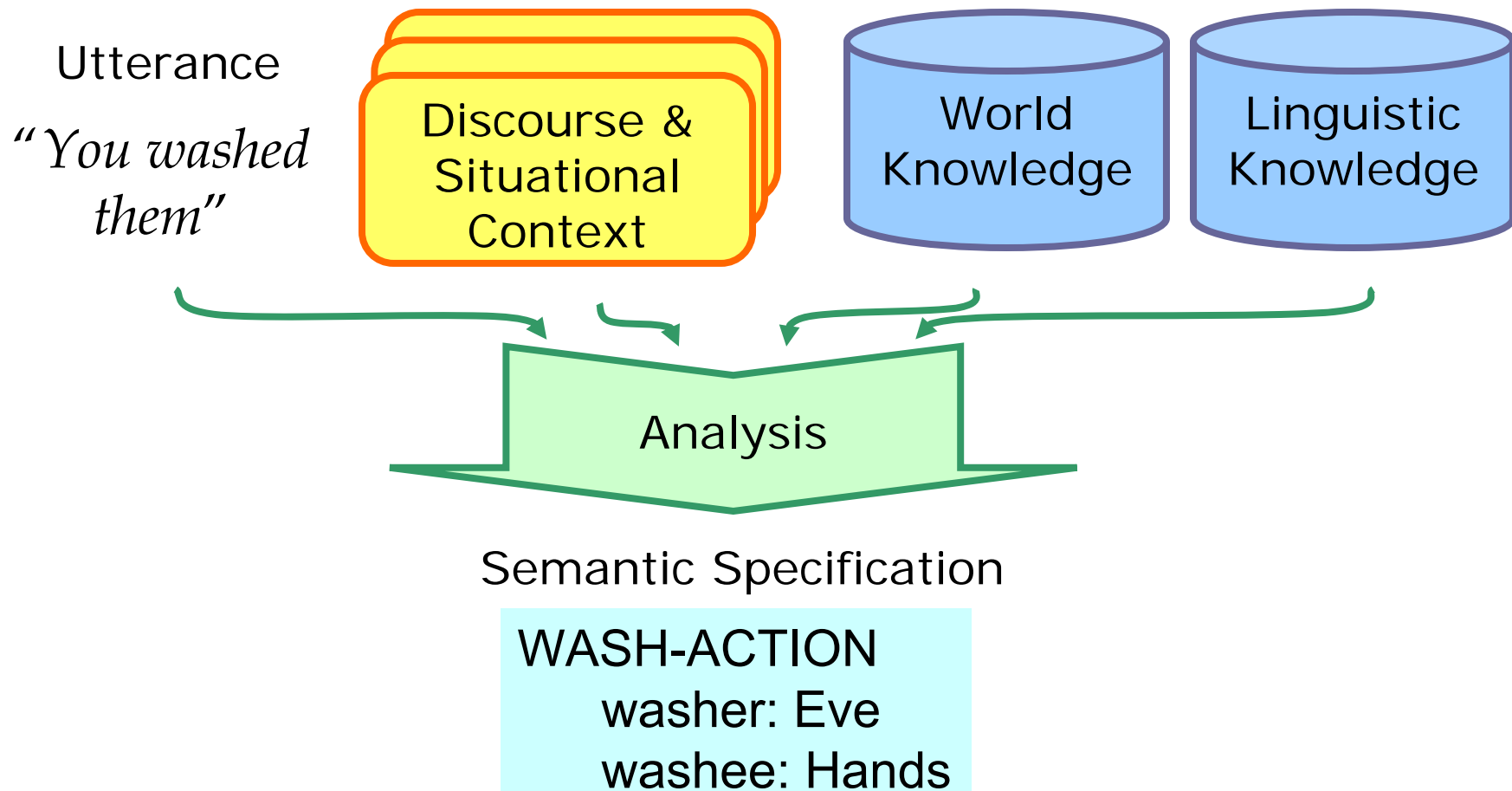
WASHED before THEM

meaning: WASH-ACTION

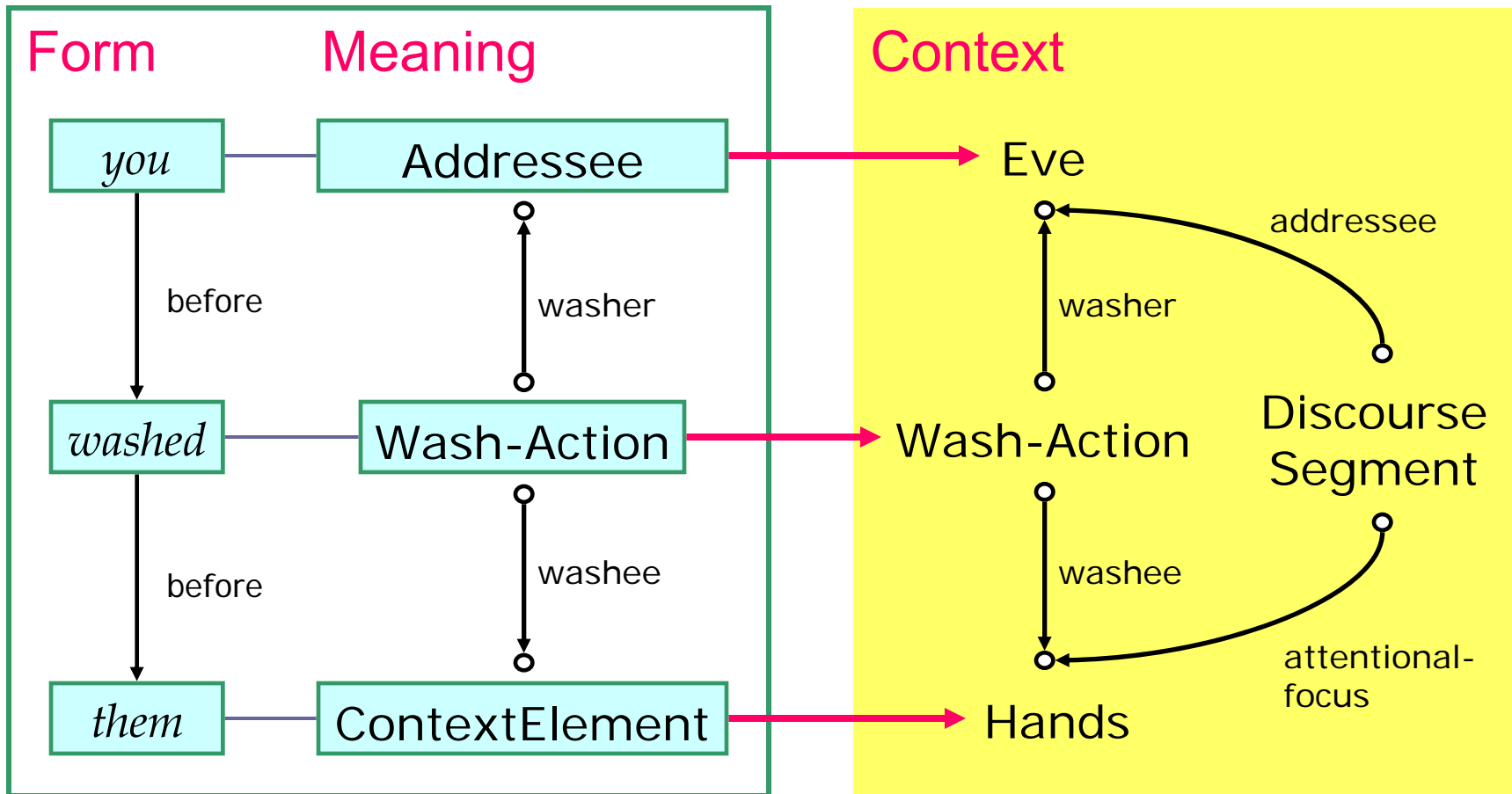
washer: addressee

washee: ContextElement

Analysis produces a semantic specification



Analysis draws on constructions and context



The analysis process fills in the roles of the
WASH-ACTION schema

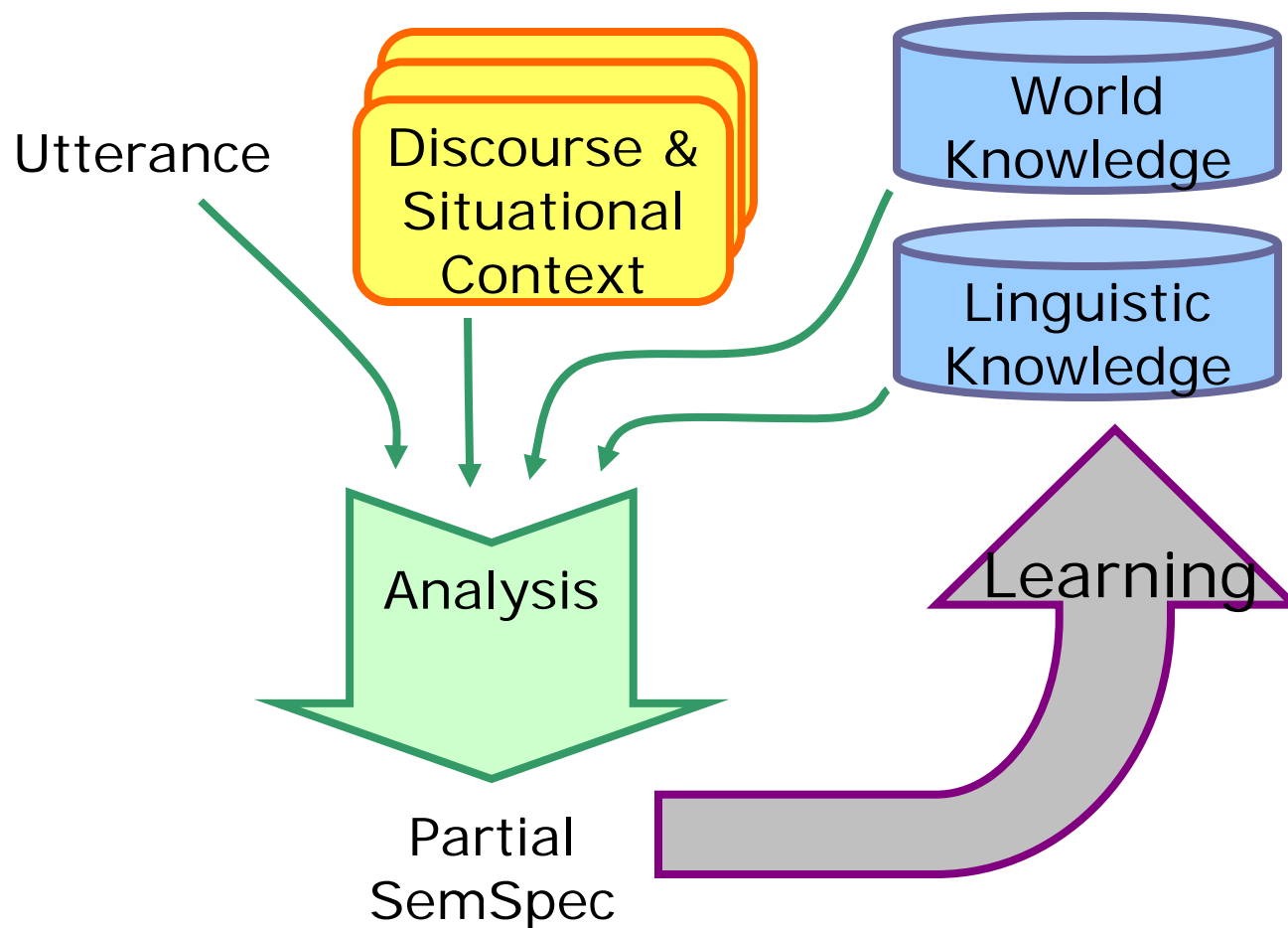
Semantic Specification

WASH-ACTION

washer: Eve

washee: Hands

Learning updates linguistic knowledge based on input utterances



Context supplements understanding: Incomplete grammars produce partial SemSpec

Form

Meaning

Context

you

Addressee

Eve

washed

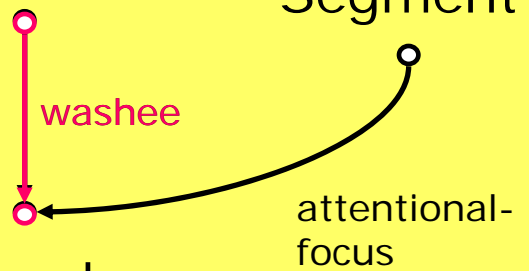
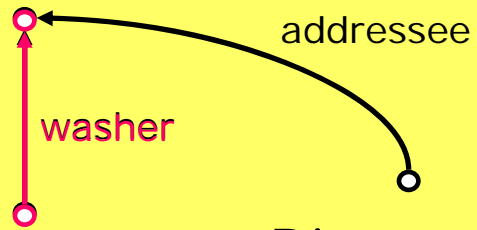
Wash-Action

Wash-Action

them

ContextElement

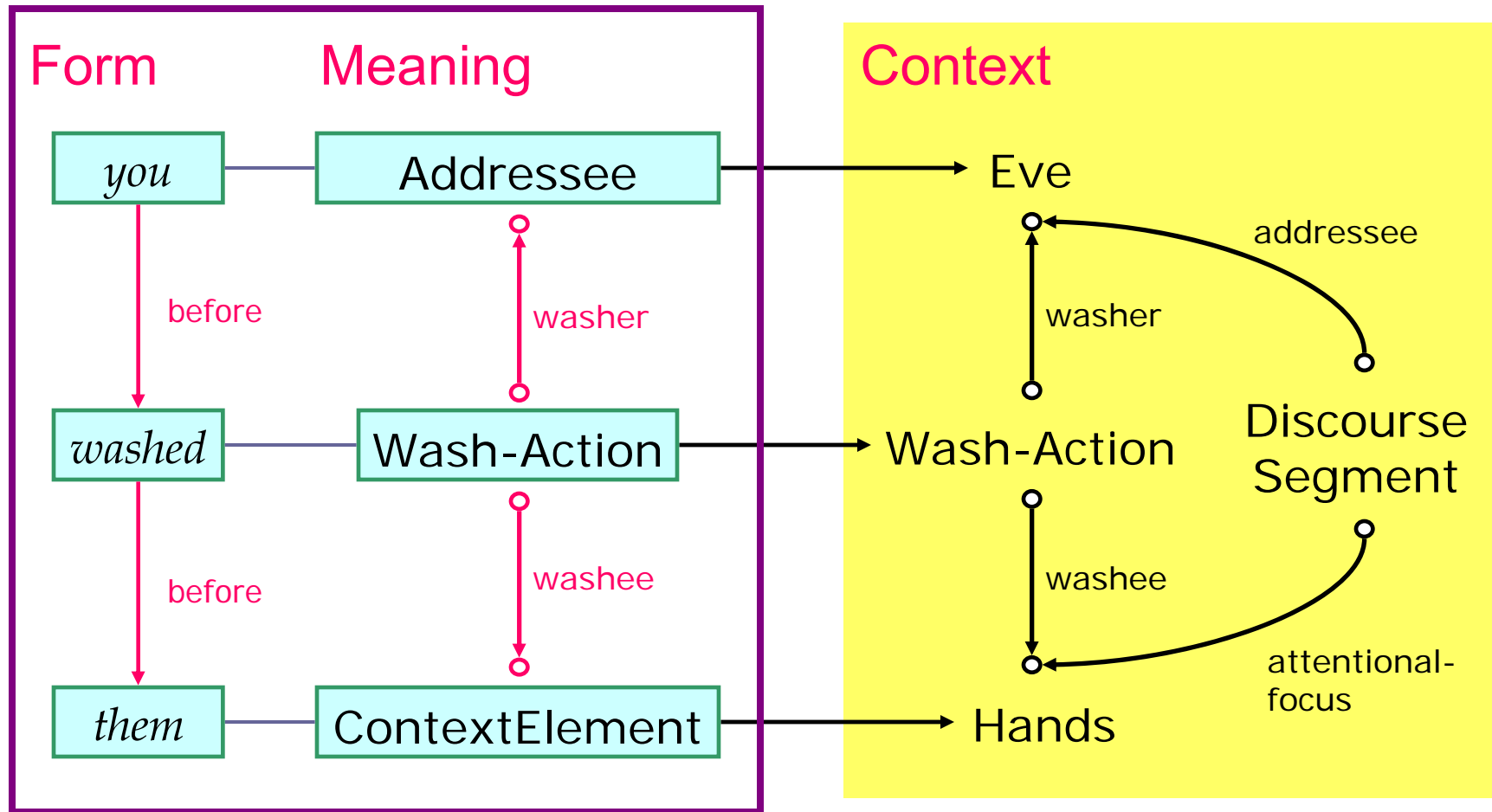
Hands



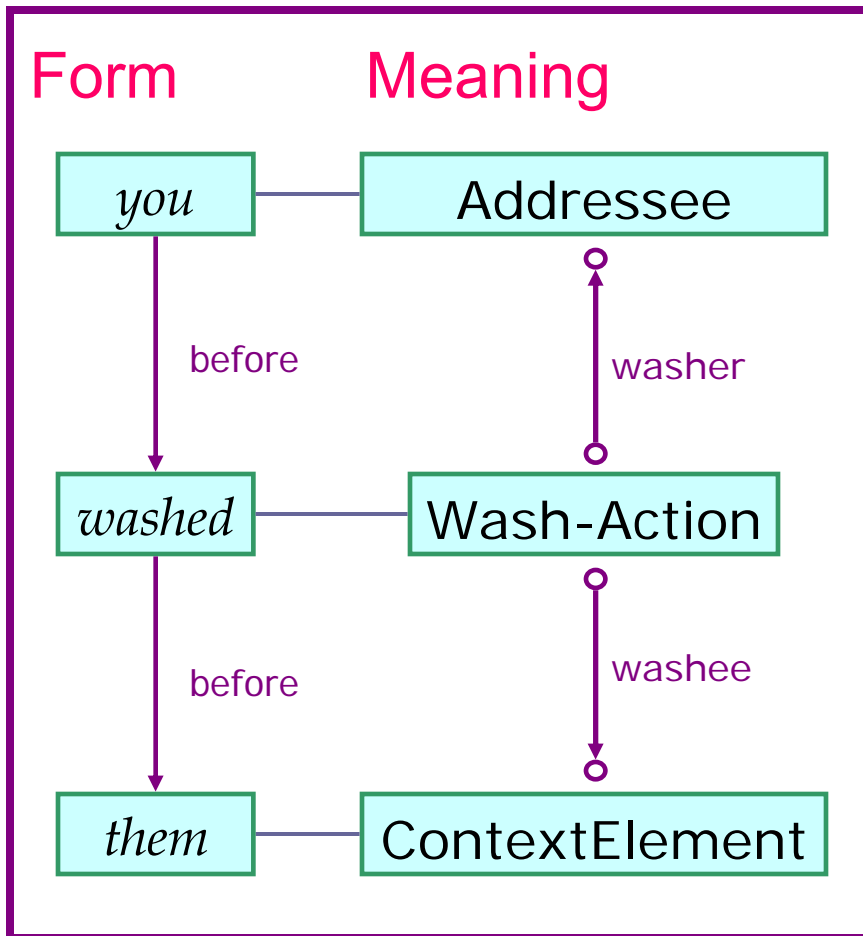
Discourse Segment

attentional-focus

Context bootstraps learning: new construction maps form to meaning



Context bootstraps learning: new construction maps form to meaning



YOU-WASHED-THEM
constituents:
 YOU, WASHED, THEM
form:
 YOU before WASHED
 WASHED before THEM
meaning: WASH-ACTION
washer: addressee
washee: ContextElement

The structured context model integrates directly into our existing systems

existing systems:

Robust parsing system
for English sentences

(Bryant, 2005)

Learning word-island
constructions from
single utterances in
English

(Chang, 2004)

Integrated structured context model &
ECG extensions

The structured context model enables the learning of context-dependent constructions

Integrated structured context model & ECG extensions

ongoing work:

- Probabilistic best-fit analysis
 - Extensions for Chinese
(Mok & Bryant, 2006)
 - Learning general constructions from a stretch of discourse
 - Handling argument omission
-

Eve, just 5 months earlier, produces little more than two-word combinations

- Mother: oh you dropped the celery .
 - Mother: let me get it for you .
 - Mother: Mommy'll have to wash it off .
 - Eve: oh foot
 - Mother: your foot ?
 - Mother: where is your foot ?
 - Eve: foot xxx chair
-

Context is necessary for understanding
this dialogue in Mandarin Chinese

- *CHI: zang1. (dirty)
- *MOT: zang1 le . (dirty ASPECT)
- *CHI: ei xi3+xi3 . (INTERJ wash+wash)
- *MOT: en xi3 xi3 . (INTERJ wash+wash)
- *MOT: xing2 . (alright)

Context is represented with feature structures

schema Discourse

roles

participants: ContextElement

objects: ContextElement

situational-history: Event

discourse-history: DiscourseSegment

schema DiscourseSegment

roles

speaker: ContextElement

addressee: ContextElement

attentional-focus: ContextElement

content

speech-act

Representing the hand-washing scenario

Discourse:

Discourse01
participants: Eve (1) , Mother (2)
objects: Hands (3)
situational-history: Wash-Action (4)
discourse-history: DS01 (5)

Situational History:

Wash-Action (4)
washer: Eve (1)
washee: Hands (3)

Discourse History:

DS01 (5)
speaker: Mother (2)
addressee: Eve (1)
attentional-focus: Hands (3)
content: {"are they clean yet?"}
speech-act: question

Participants:

Eve (1)
category: child
gender: female
name: Eve
age: 2

Mother (2)
category: parent
gender: female
name: Eve
age: 33

Objects:

Hands (3)
category: BodyPart
part-of: Eve (1)
number: plural
accessibility: accessible

Linguistic knowledge is represented with Embodied Construction Grammar (Bergen & Chang, 2005)

- Construction Grammar:
 - Form-Meaning pairs
 - Captures both compositional and non-compositional meaning
 - e.g. past tense *-ed*, *What's X doing Y?*
 - Embodied Construction Grammar:
 - schematic meaning representation grounded in sensori-motor experience
 - part of the Neural Theory of Language project
-

ReferentDescriptor is a bookkeeping structure that allows constructions access to context

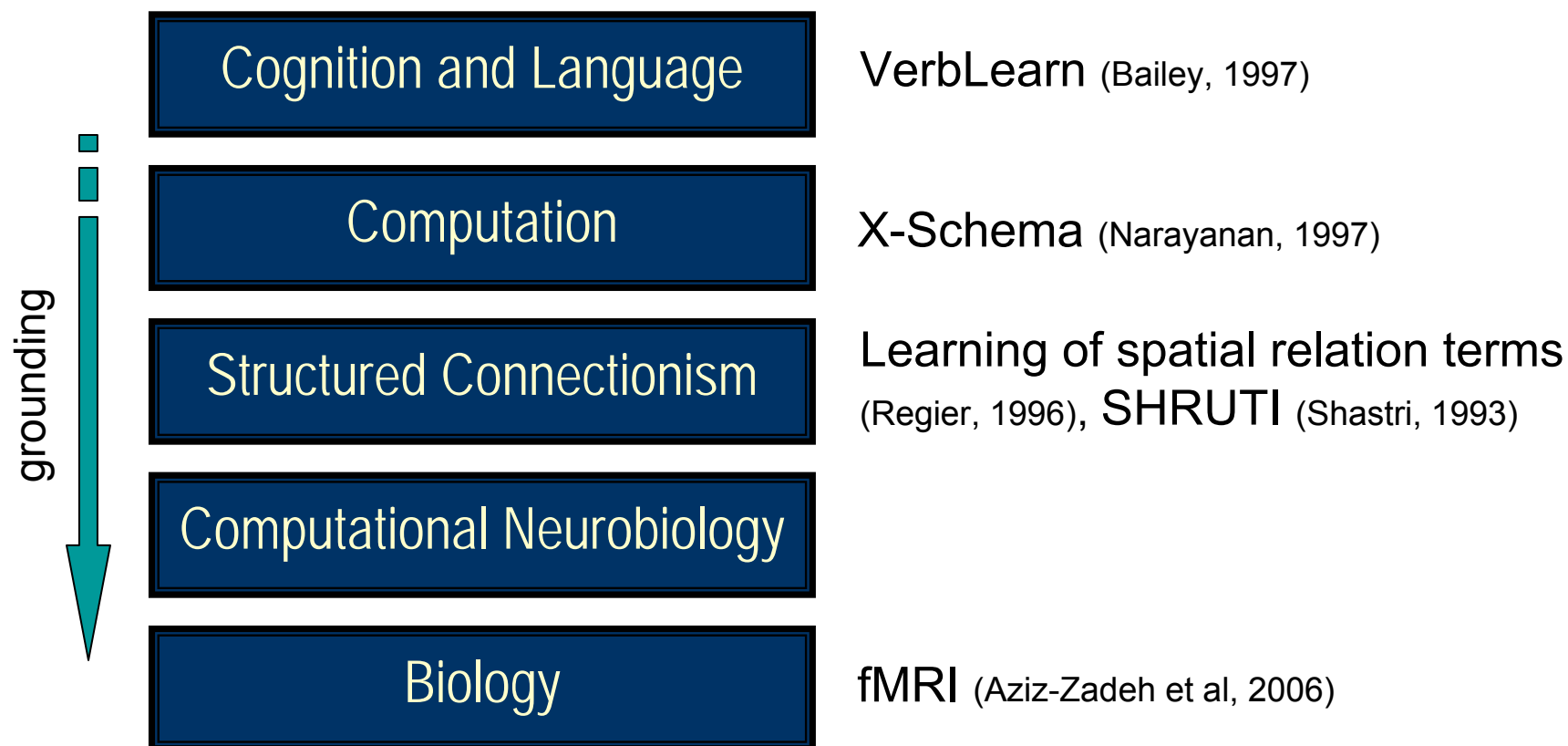
schema ReferentDescriptor
roles
number
grammatical-gender
ontological-category
case
accessibility
modifiers
resolved-ref: ContextElement

construction Them
form
self_f.orth ← "them"
meaning: ContextElement
evokes ReferentDescriptor **as** rd
self_m ↔ rd.resolved-ref
rd.number ← plural
rd.accessibility ← accessible
rd.case ← object

The structured context model enables the learning of context-dependent constructions

- ECG extensions, integrated context model
 - Language understanding:
 - existing: robust parsing on English (Bryant, 2005)
 - ongoing: better resolution, Chinese (Mok & Bryant, 2006)
 - Language learning:
 - existing: English child-directed utterances (single utterances)
 - ongoing: discourse, argument omission
-

Neural grounding in NTL is achieved by layers of abstraction



MIT Press

Jerome A. Feldman

From Molecule to Metaphor:
A Neural Theory of Language

