Generative Lexicon: Integrating Theoretical and Distributional Methods

James Pustejovsky
Brandeis University

Elisabetta Ježek
University of Pavia

July 17-21, 2017
5. Meaning Composition in GL
ESSLLI 2017
University of Toulouse

Course Outline

- July 17: Introduction to GL and Distributional Analysis
- July 18: Qualia Structure
- July 19: Event Structure
- July 20: Argument Structure
- July 21: Meaning Composition and Co-composition

Lecture 1: July 17

Introduction to GL and Distributional Analysis

- Basic concepts in GL
  - Notation and Language: typed feature structures
  - Qualia Structure
  - Events and their participants
  - Meaning Composition in GL
- Distributed meaning: Spreading the semantic load
- Polysemy in language
  - Types of contextual variations
  - Detecting copredications in corpora
- Evidence-based linguistics and distributional analysis

Lecture 2: July 18

Qualia Structure

- What is a Quale?
- What motivates Qualia?
- Default Qualia and context updating
- Methodology to identify Qualia
- Data for each Quale
- Qualia and Conventionalized Attributes
- Qualia and Type Systems

Lab on Qualia identification in corpora using SkE

Lecture 3: July 19

Event Structure

- Events as Structured Objects
- Event Types
  - States
  - Transitions
  - Point Verbs
  - Processes
- Events as Labeled Transition Systems
- Dynamic Event Models

Lab on identification of event type properties in corpora

Lecture 4: July 20

Argument Structure

- Argument Types in GL
  - True Arguments
  - Defaulted Arguments
  - Shadow Arguments
  - Hidden Arguments
- Argument Structure Representation
  - Semantic Types and Lexical Sets
  - Distributional Approach to Semantic Types
- Dynamic Argument Structure

Lab on Semantic Types and Lexical Sets

Lecture 5: July 21

Meaning Composition and Co-composition in GL

- Basic Assumptions
- Simple Function Application
- Coercion
- Subselection
- Co-composition
- The Lexicon-Pragmatics Interplay

Studies in evidence-based coercion

Meaning Composition in GL

Meaning composition

- Basic Assumptions
- Simple Function Application
- Coercion
- Data on Argument Typing and Coercion
- Co-composition

GL Type Theory

1. Natural Type: $g \vdash x : \alpha$
2. Artifactual Types:
   1. Artifact Entity: $x : e_N \otimes_a \sigma$
      x exists because of event σ
   2. Functional Entity: $x : e_N \otimes_t \tau$
      the purpose of x is τ
   3. Functional Artifactual Entity: $x : (e_N \otimes_a \sigma) \otimes_t \tau$
      x exists because of event σ for the purpose τ
3. Complex Types: $g \vdash x : \alpha \bullet \beta$

Natural Predicate Types

Predicates formed with Natural Entities as arguments:

1. *fall*: $e_N \rightarrow t$
2. *touch*: $e_N \rightarrow (e_N \rightarrow t)$
3. *be under*: $e_N \rightarrow (e_N \rightarrow t)$

a. $\lambda x{:}\, e_N[\ \mathit{fall}(x)]$
b. $\lambda y{:}\, e_N \lambda x{:}\, e_N[\ \mathit{touch}(x,y)]$
c. $\lambda y{:}\, e_N \lambda x{:}\, e_N[\ \mathit{be\text{-}under}(x,y)]$

Artifactual Entity Types

Entities formed from the Naturals by adding the AGENTIVE or TELIC qualia roles:

1. Artifact Entity: $x : e_N \otimes_a \sigma$
   x exists because of event σ
2. Functional Entity: $x : e_N \otimes_t \tau$
   the purpose of x is τ
3. Functional Artifactual Entity: $x : (e_N \otimes_a \sigma) \otimes_t \tau$
   x exists because of event σ for the purpose τ

a. *beer*: $(liquid \otimes_a brew) \otimes_t drink$
b. *knife*: $(phys \otimes_a make) \otimes_t cut$
c. *house*: $(phys \otimes_a build) \otimes_t live\_in$

Artifactual Predicate Types

Predicates formed with Artifactual Entities as arguments:

1. *spoil*: $e_N \otimes_t \tau \rightarrow t$
2. *fix*: $e_N \otimes_t \tau \rightarrow (e_N \rightarrow t)$

a. $\lambda x{:}\, e_A[\ \mathit{spoil}(\mathrm{x})]$
b. $\lambda y{:}\, e_A \lambda x{:}\, e_N[\ \mathit{fix}(\mathrm{x},\mathrm{y})]$

- The beer spoiled.
- Mary fixed the watch.

Complex Entity Types

Entities formed from the Naturals and Artifactuals by a product type between the entities, i.e., the dot, •.

1. a. Mary doesn't believe the book.
   b. John sold his book to Mary.
2. a. The exam started at noon.
   b. The students could not understand the exam.

Motivating Dot Objects

When a single word or phrase has the ability to appear in selected contexts that are contradictory in type specification.

If a lexical expression, α, where $\sigma \sqcap \tau = \bot$:

1. [___] $_{\sigma}$ X
2. [___] $_{\tau}$ Y

are both well-formed predications, then α is a dot object (complex type).

Dot Object Inventory: 1

1. Act•Proposition: promise, allegation, lie
   - I doubt John's promise of marriage.
   - John's promise of marriage happened while we were in Prague.
2. Attribute•Value: temperature, weight, height, tension, strength
   - The temperature is rising.
   - The temperature is 23.

Dot Object Inventory: 2

1. Event•Information: lecture, play, seminar, exam, quiz, test
   a. My lecture lasted an hour.
   b. Nobody understood my lecture.
2. Event•Music: sonata, symphony, song, performance, concert
   a. Mary couldn't hear the concert.
   b. The rain started during the concert.

Dot Object Inventory: 3

1. Event•Physical: lunch, breakfast, dinner, tea
   a. My lunch lasted too long today.
   b. I pack my lunch on Thursdays.
2. Information•Physical: book, cd, dvd, dictionary, diary, mail, email, mail, letter
   a. Mary burned my book on Darwin.
   b. Mary believes all of Chomsky's books.

Complex Predicate Types

Predicates formed with a Complex Entity Type as an argument:

1. *read*: $phys \bullet info \rightarrow (e_N \rightarrow t)$
2. Expressed as typed arguments in a λ-expression:
   $\lambda y{:}\, phys \bullet info\ \lambda x{:}\, e_N[read(x,y)]$
3. Mary read the book.

Modes of Composition

(1) a. PURE SELECTION (Type Matching): the type a function requires is directly satisfied by the argument;
b. ACCOMMODATION: the type a function requires is inherited by the argument;
c. TYPE COERCION: the type a function requires is imposed on the argument type. This is accomplished by either:
    i. *Exploitation*: taking a part of the argument's type to satisfy the function;
    ii. *Introduction*: wrapping the argument with the type required by the function.

Two Kinds of Coercion in Language

- Domain-shifting: The domain of interpretation of the argument is shifted;
- Domain-preserving: The argument is coerced but remains within the general domain of interpretation.

Domain-Shifting Coercion

1. Entity shifts to event:
   I enjoyed the beer
2. Entity shifts to proposition:
   I doubt John.

Domain-Preserving Coercion

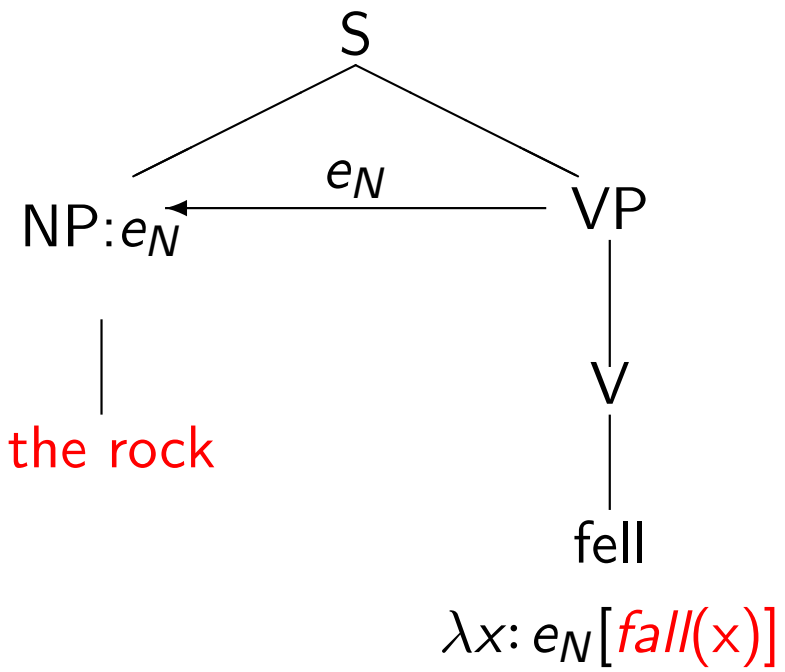
1. Count-mass shifting: There's chicken in the soup.
2. NP Raising: Mary and every child came.
3. Natural-Artifactual shifting: The water spoiled.
4. Natural-Complex shifting: She read a rumor.
5. Complex-Natural shifting: John burnt a book.
6. Artifactual-Natural shifting: She touched the phone.

Direct Argument Selection

- The spokesman denied the statement (PROPOSITION).
- The child threw the ball (PHYSICAL OBJECT).
- The audience didn't believe the rumor (PROPOSITION).

Natural Selection

1. The rock fell.



Natural Selection

(2) a. "fall" is of type $phys \rightarrow t$;
b. "the rock" is of type *phys* (modulo GQ type shifting);
c. Function Application (TM) applies;
$\Longrightarrow$ fall(the-rock)

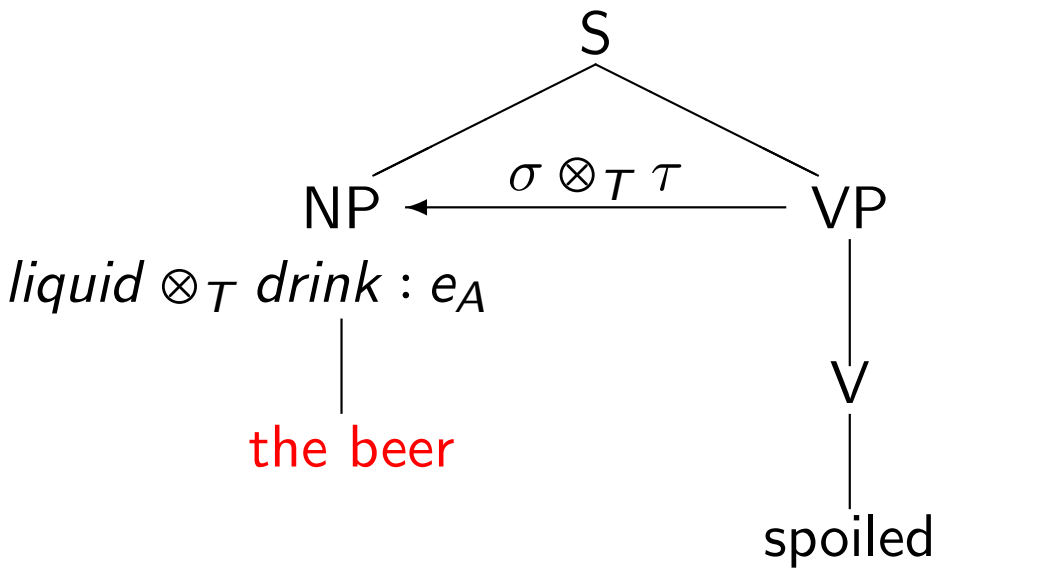
(3) Some water fell on the floor.

This results in the derivation shown in (4):

(4) a. "fall" is of type $phys \rightarrow t$;
b. "some water" is of type *liquid* (modulo GQ type shifting);
c. Accommodation Subtyping applies, $liquid \sqsubseteq phys$:
$\Longrightarrow$ "some water" is of type *phys*:
d. Function Application (TM) applies;
$\Longrightarrow$ fall(some-water)

Pure Selection: Artifactual Type

1. The beer spoiled.



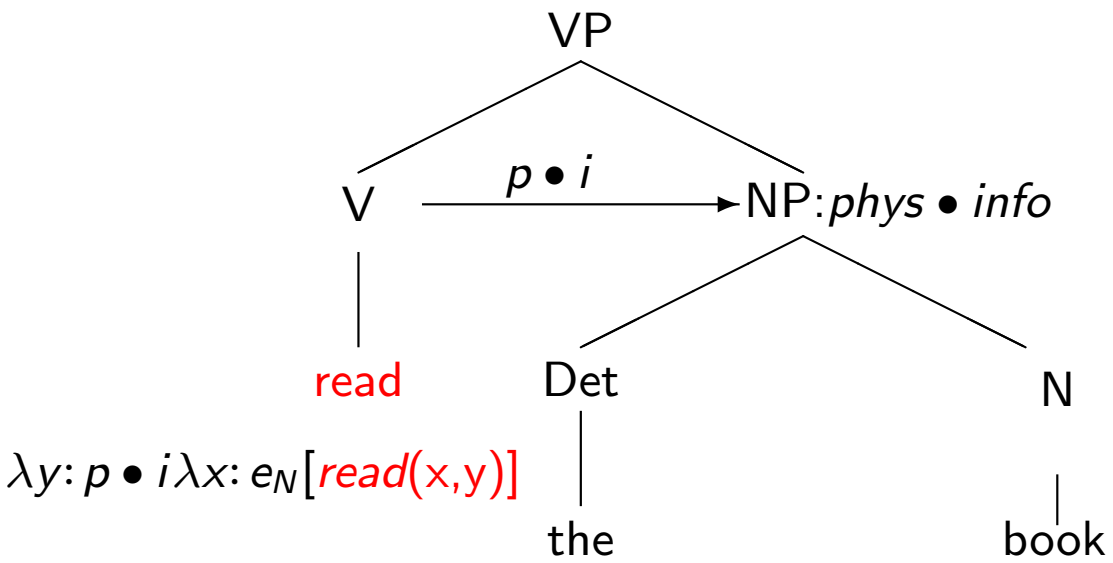
$\lambda x{:}\, e_A[spoil(x)]$

Pure Selection: Artifactual Type

(5) a. "spoil" is of type $phys \otimes_T \tau \rightarrow t$;
b. "the beer" is of type $liquid \otimes_T drink$ (modulo GQ type shifting);
c. Accommodation Subtyping applies to the head, $liquid \sqsubseteq phys$:
$\Longrightarrow$ "the beer" has head type *phys*:
d. Accommodation Subtyping applies to the TELIC, $drink \sqsubseteq \tau$:
$\Longrightarrow$ "the beer" has TELIC type τ
e. "the beer" has type $phys \otimes_T \tau$;
f. Function Application (TM) applies;
$\Longrightarrow$ spoil(the-beer)

Pure Selection: Complex Type

1. John read the book.



Pure Selection: Complex Type

The derivation of this example is fairly direct, and is shown in (6).

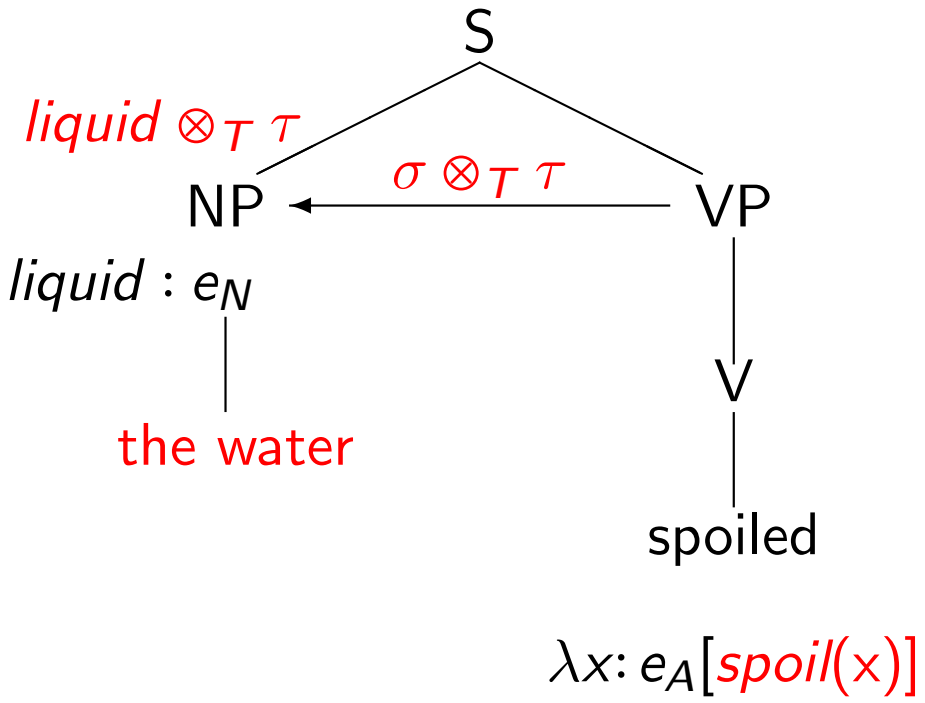
(6) a. "read" is of type $p \bullet i \rightarrow (e_N \rightarrow t)$;
b. "the book" is of type $p \bullet i$ (modulo GQ type shifting);
c. Function Application (TM) applies;
$\Longrightarrow \lambda x$ [read(x,the-book)]

Coercion of Arguments

- The president denied the attack.
  EVENT → PROPOSITION
- The White House denied this statement.
  LOCATION → HUMAN
- This book explains the theory of relativity.
  PHYS • INFO → human
- d. The Boston office called with an update.
  EVENT → INFO

Type Coercion: Qualia-Introduction

1. The water spoiled.

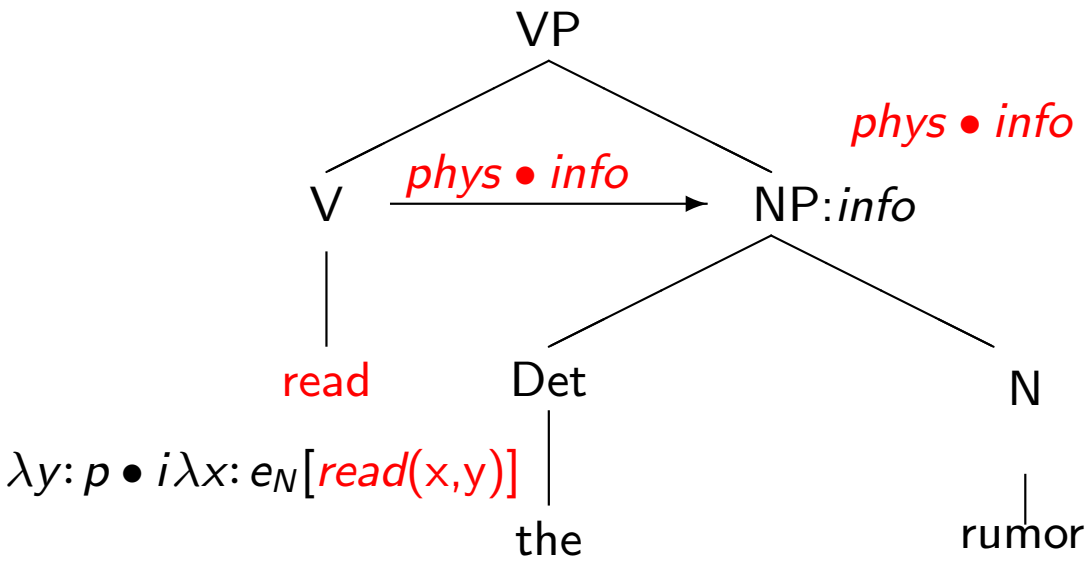


Type Coercion: Qualia-Introduction

(7) a. "spoil" is of type $\mathit{phys} \otimes_T \tau \rightarrow t$;
b. "the water" is of type *liquid* (modulo GQ type shifting);
c. Accommodation Subtyping applies to the head, $\mathit{liquid} \sqsubseteq \mathit{phys}$:
$\Longrightarrow$ "the water" has type *phys*;
d. Coercion by Qualia Introduction (CI-Q) applies to the type *phys*, adding a TELIC value τ:
$\Longrightarrow$ "the water" has type $\mathit{phys} \otimes_T \tau$;
e. Function Application applies;
$\Longrightarrow$ spoil(the-water)

Type Coercion: Natural to Complex Introduction

John read the rumor.

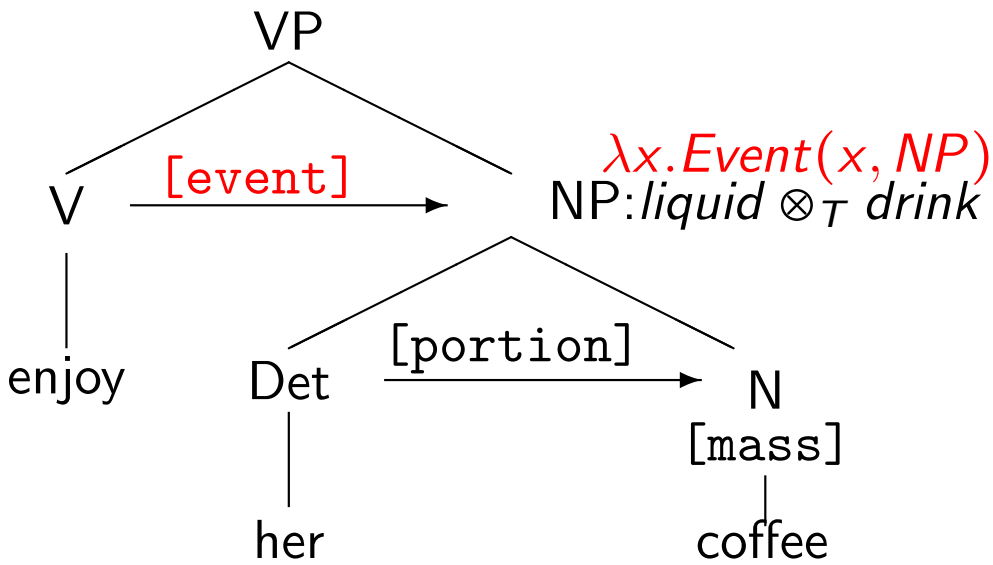


Type Coercion: Natural to Complex Introduction

(8) a. "read" is of type $p \bullet i \rightarrow (e_N \rightarrow t)$;
b. "the rumor" is of type i, $i \sqsubseteq t$ (modulo GQ type shifting);
c. Coercion by Dot Introduction (CI-•) applies to the type i, adding the missing type value, p, and the relation associated with the •:
$\Longrightarrow$ "the rumor" has type $p \bullet i$;
e. Function Application applies;
$\Longrightarrow \lambda$x[read(x,the-rumor)]

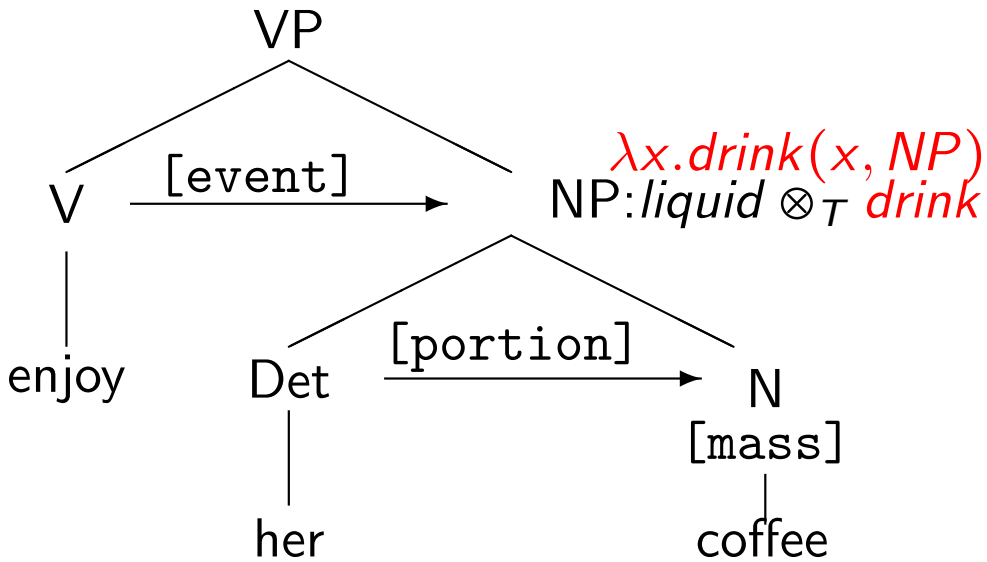
Type Coercion: Event Introduction

1. Mary enjoyed her coffee.



Type Coercion: Qualia Exploitation

1. Mary enjoyed her coffee.

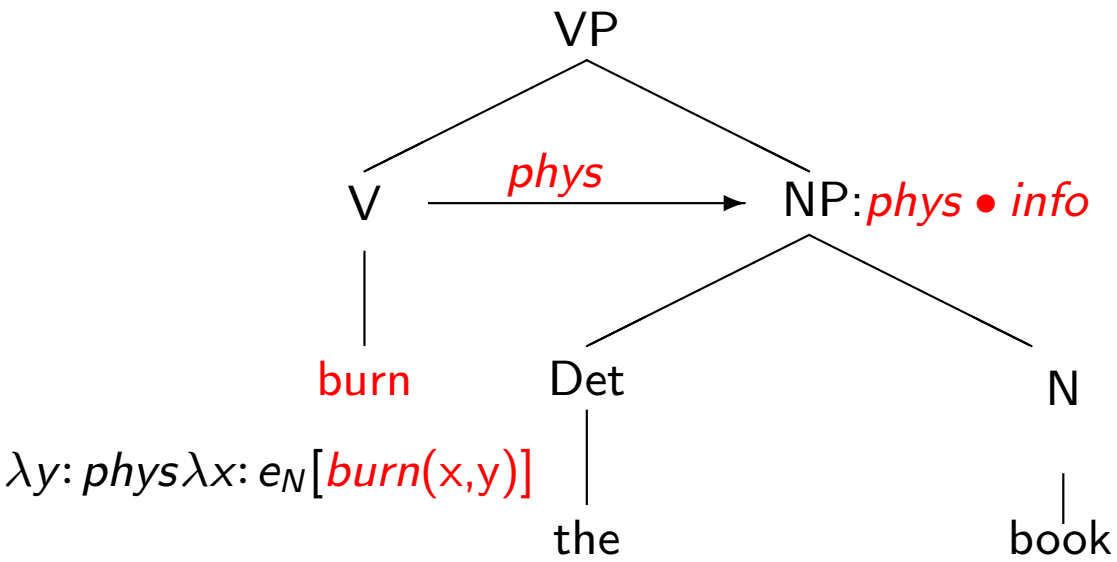


Type Coercion: Qualia Exploitation

(9) a. "enjoy" is of type $event \rightarrow (e_N \rightarrow t)$;
b. "her coffee" is of type $liquid \otimes_T drink$, (modulo GQ type shifting);
c. Coercion by Introduction (CI) applies to the type $liquid \otimes_T drink$, returning *event*:
$\Longrightarrow$ "her coffee" has type *event*;
d. Coercion by Qualia Introduction (CI-Q) applies to the type *event*, adding a value *drink* to the predicate, P:
$\Longrightarrow$ "her coffee" has type *event*, with P bound to *drink*;
e. Function Application applies;
$\Longrightarrow \lambda y$[enjoy(y,$\lambda x \exists e$[drink(e,x,her-coffee)]]

Type Coercion: Dot Exploitation

1. The police burned the book.
2. Mary believes the book.



Verb-Argument Composition Table

	Verb selects:		
Argument is:	**Natural**	**Artifactual**	**Complex**
Natural	Selection	Qualia Intro	Dot Intro
Artifactual	Qualia Exploit	Selection	Dot Intro
Complex	Dot Exploit	Dot Exploit	Selextion

Data on Argument Typing and Coercion

Pustejovsky and Jezek 2008

- Methodology inspired by Corpus Pattern Analysis (CPA)(Hanks 1994, Pustejovsky et al, 2004, Hanks and Pustejovsky 2005).
- Select a target verb in EnTenTen13 using SkE: *finish*, *last*, *attend*, *avoid*, *drink*, *leave*, *reach*, *smell*, *listen (to)*, *kill*, *ring*.
- Extract a sample of concordances.
- Use CPA list of types.
- Identify typing for specific argument positions in a specific verb sense by manually clustering the argument fillers into lexical sets (Hanks 1996).
- Identify type mismatches.

Data on Argument Typing and Coercion

Pustejovsky and Jezek 2008

(10) *ring* (Body: 'call by phone'; Arg: HUMAN)
Object

a. HUMAN: mother, doctor, Chris, friend, neighbour, director
b. INSTITUTION: police, agency, club
b. LOCATION: flat, house; Moscow, Chicago, London, place

'I rang the house a week later and talked to Mrs Gould'
'The following morning Thompson rang the police'
'McLeish had rung his own flat to collect messages'
'I said Chicago had told me to ring London.'

Data on Coercion: Dot Exploitation

Pustejovsky and Jezek 2008

(11) *house* (PHYS•LOCATION)
Object

a. PHYS: built, buy, sell, rent, own, demolish, renovate, burn down, erect, destroy, paint, inherit, repair
b. LOCATION: leave, enter, occupy, visit, inhabit, reach, approach, evacuate, inspect, abandon

'they built these houses onto the back of the park'
'the bus has passed him as he left the house'

Data on Coercion: Dot Exploitation

Pustejovsky and Jezek 2008

(12) *interview* (EVENT•INFORMATION)

Object

a. EVENT: conduct, give, arrange, attend, carry out, terminate, conclude, close, complete, end, hold, cancel, undertake, extend, control, continue, begin

b. INFORMATION: structure, discuss, analyze, describe

Subject

a. EVENT: last, go well, take place, follow, end, progress, begin, become tedious, precede, start, happen

b. INFORMATION: covers, centre on, concern, focus on

'Officials will be conducting interviews over the next few days'

'Let's discuss the interview'

Data on Qualia Exploitation

Pustejovsky and Jezek 2008

(13) *hear* (Body: 'perceive with the ear'; Arg:SOUND)
Object

a. SOUND: voice, sound, murmur, bang, thud, whisper, whistle

b. Q-E OF $phys \otimes_{telic} \tau$: siren, bell, alarm clock

'then from the house I heard the bell'
'you can hear sirens most of the time'
'the next thing he heard was his alarm clock'

Data on Type Introduction

Pustejovsky and Jezek 2008

(14) *read* (PHYS•INFORMATION)
Objects

a. *human* $\otimes_{telic}$ *write*: Dante, Proust, Homer, Shakespeare, Freud

'That is why I read Dante now'

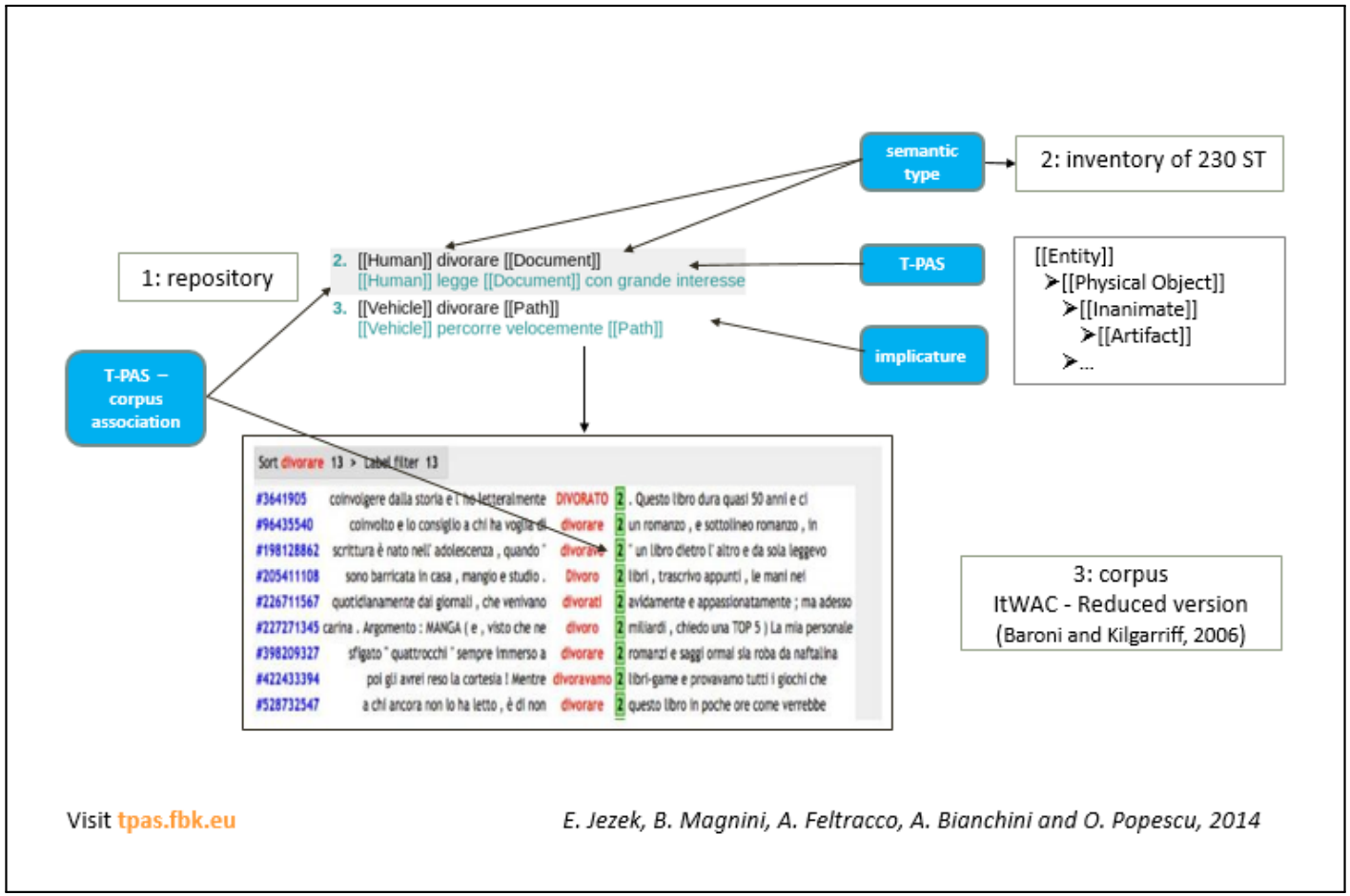
(15) *read* (PHYS•INFORMATION)
Objects

a. EVENT•INFO: story, description, judgement, quote, reply, speech, proclamation, statement, question, interview
b. SOUND•INFO: music

'I've read your speeches'
'I discovered he couldn't read music'

Typed Predicate-Argument Structure (T-PAS)

Jezek, Magnini, Feltracco, Bianchini, Popescu 2014



Methodology from Hanks, 2004, 2013.

Mismatch classification

- Verb classes (Levin 1993, VerbNet).

Mismatch classification

- Verb classes (Levin 1993, VerbNet).
- Targeted grammatical relation: SUBJ_OF, OBJ_OF, COMPL

Mismatch classification

- Verb classes (Levin 1993, VerbNet).
- Targeted grammatical relation: SUBJ_OF, OBJ_OF, COMPL
- Shift types: Artifact as Event, Artifact as Human, Artifact as Sound, Event as Location, Vehicle as Human.

Mismatch classification

- Verb classes (Levin 1993, VerbNet).
- Targeted grammatical relation: SUBJ_OF, OBJ_OF, COMPL
- Shift types: Artifact as Event, Artifact as Human, Artifact as Sound, Event as Location, Vehicle as Human.
- SemEval Coercion Task 7: Argument Selection and Coercion (Pustejovsky et al. 2010, Jezek and Quochi 2010).

Aspectual Verbs (Jezek, Magnini, Feltracco, Bianchini, Popescu 2014)

[[Human]-subj] interrompe [[Event]-obj]

- Arriva Mirko e interrompe la conversazione.
  'Mirko arrives and interrupts the conversation' (matching)
- Il presidente interrompe l'oratore.
  'The president interrupts the speaker' (HUMAN as EVENT; T=parlare 'speak')

Communication Verbs

[[Human]-subj] annuncia [[Event]-obj]

- Lo speaker annuncia la partenza.
  'The speaker announces the departure' (matching)

Communication Verbs

[[Human]-subj] annuncia [[Event]-obj]

- Lo speaker annuncia la partenza.
  'The speaker announces the departure' (matching)
- Il maggiordomo annuncia gli invitati.
  'The butler announces the guests' (HUMAN as EVENT, CA=arrivare 'arrive')

Communication Verbs

[[Human]-subj] annuncia [[Event]-obj]

- Lo speaker annuncia la partenza.
  'The speaker announces the departure' (matching)
- Il maggiordomo annuncia gli invitati.
  'The butler announces the guests' (HUMAN as EVENT, CA=arrivare 'arrive')
- L'altoparlante annunciava l'arrivo del treno.
  'The loudspeaker announces the arrival of the train'
  (ARTIFACT as HUMAN; T=usare 'use'(human, tool))

Communication Verbs

[[Human]-subj] annuncia [[Event]-obj]

- Lo speaker annuncia la partenza.
  'The speaker announces the departure' (matching)
- Il maggiordomo annuncia gli invitati.
  'The butler announces the guests' (HUMAN as EVENT, CA=arrivare 'arrive')
- L'altoparlante annunciava l'arrivo del treno.
  'The loudspeaker announces the arrival of the train' (ARTIFACT as HUMAN; T=usare 'use'(human, tool))
- Una telefonata anonima avvisa la polizia.
  'An anonymous telephone call alerted the police' (EVENT as HUMAN; AG=telefonare 'phone'(human1, human2))

Avoid Verbs

[[Human]-subj] evita [[Event]-obj]

- Abbiamo evitato l'incontro.
  'We avoided the meeting' (matching)
- Meglio evitare i cibi fritti.
  'It is best to avoid fried food' (ARTIFACT as EVENT;
  T=mangiare 'eat')

Forbid Verbs

[[Human]-subj] vieta [[Event]-obj]

- Nell'Italia di allora la legge vietava l'aborto.
  'At that time in Italy law prohibited abortion' (matching)
- La Francia vieta il velo a scuola.
  'France bans the headscarf in schools' (ARTIFACT as EVENT; T=indossare 'wear')

Verbs of Desire (Bos 2009)

[[Human]-subj] preferire [[Event]-obj]

- Preferisco bere piuttosto che mangiare.
  'I prefer drinking to eating' (matching)
- Preferisco la birra al vino.
  'I prefer beer to wine' (ARTIFACT as EVENT; T=bere 'drink')

Perception Verbs

[[Human]-subj] ascolta [[Sound]-obj]

- Rilassarsi ascoltando il rumore della pioggia.
  'Relax while listening to the sound of rain' (matching)

Perception Verbs

[[Human]-subj] ascolta [[Sound]-obj]

- Rilassarsi ascoltando il rumore della pioggia.
  'Relax while listening to the sound of rain' (matching)
- Ascoltava la radio con la cuffia.
  'He listened to the radio with his earphones' (ARTIFACT as SOUND: T=produrre_suono 'produce_sound')

Perception Verbs

[[Human]-subj] ascolta [[Sound]-obj]

- Rilassarsi ascoltando il rumore della pioggia.
  'Relax while listening to the sound of rain' (matching)
- Ascoltava la radio con la cuffia.
  'He listened to the radio with his earphones' (ARTIFACT as SOUND: T=produrre_suono 'produce_sound')
- Rimasi a lungo ad ascoltare il suo respiro.
  'I stayed for a long while listening to his breath' (EVENT as SOUND; NT=produrre_suono 'produce_sound')

Perception Verbs

[[Human]-subj] ascolta [[Sound]-obj]

- Rilassarsi ascoltando il rumore della pioggia.
  'Relax while listening to the sound of rain' (matching)
- Ascoltava la radio con la cuffia.
  'He listened to the radio with his earphones' (ARTIFACT as SOUND: T=produrre_suono 'produce_sound')
- Rimasi a lungo ad ascoltare il suo respiro.
  'I stayed for a long while listening to his breath' (EVENT as SOUND; NT=produrre_suono 'produce_sound')
- Non ho potuto ascoltare tutti i colleghi
  'I could not listen to all colleagues' (HUMAN as SOUND; CA=parlare 'speak')

Directed Motion Verbs 1/3

[[Human]-subj] raggiunge [[Location]-obj]

- Abbiamo raggiunto l'isola alle 5.
  'We reached the island at 5' (matching)
- Ho raggiunto il semaforo e ho svoltato a destra.
  'I reached the traffic light and turned right' (ARTIFACT as LOCATION; CA= essere_a 'be_at'(location))

Directed Motion Verbs 2/3

[[Human]-subj] arriva (Adv [[Location]])

- Alla fine, ormai col buio, sono arrivata a una radura.
  'Finally in the dark I came upon a clearing.' (matching)
- Gli invitati arrivano al concerto in ritardo.
  'The guests arrived late at the concert' (EVENT as LOCATION; CA=aver luogo_a 'take place_at'(location))

Motion using a Vehicle

[[Flying Vehicle]-subj] atterra ([Adv [Location]])

- Il nostro aereo atterra alle 21.
  'Our plane lands at 9pm' (matching)

Motion using a Vehicle

[[Flying Vehicle]-subj] atterra ([Adv [Location]])

- Il nostro aereo atterra alle 21.
  'Our plane lands at 9pm' (matching)
- Il pilota e' regolarmente atterrato senza problemi.
  'The pilot landed regularly with no problems' (HUMAN as VEHICLE; T=pilotare 'pilot'(human, vehicle))

Motion using a Vehicle

[[Flying Vehicle]-subj] atterra ([Adv [Location]])

- Il nostro aereo atterra alle 21.
  'Our plane lands at 9pm' (matching)
- Il pilota e' regolarmente atterrato senza problemi.
  'The pilot landed regularly with no problems' (HUMAN as VEHICLE; T=pilotare 'pilot'(human, vehicle))
- Tutti i voli civili sono atterrati.
  'All civilian flights landed' (EVENT as VEHICLE; *ArgStr* Exploitation?)

Vehicle Verbs

[[Human]-subj] parcheggiare ([[Vehicle]-obj])

- Luca ha parcheggiato sotto casa.
  'Luca parked near the house' (matching)
- L'ambulanza ha parcheggiato lontano.
  'The ambulance parked far away' (VEHICLE as HUMAN;
  T=guidare 'drive'(human, vehicle))

Theory Meets Data

Pustejovsky and Rumshisky (2008)

- Theory predicts phenomena generally by generative rules
- Evidence-based analysis often up-ends the theoretical predictions
- Argument Preferences and Type Selection

Verbs Selecting for Artifactual Entities

(16) a. NATURAL VERBS: touch, sleep, smile
b. ARTIFACTUAL VERBS: fix, repair, break, mend, spoil

(17)
$$\begin{bmatrix} \textbf{touch} \\ \text{ARGSTR} = \begin{bmatrix} \text{ARG1} = x : \textit{phys} \\ \text{ARG2} = y : \textit{phys} \end{bmatrix} \end{bmatrix}$$

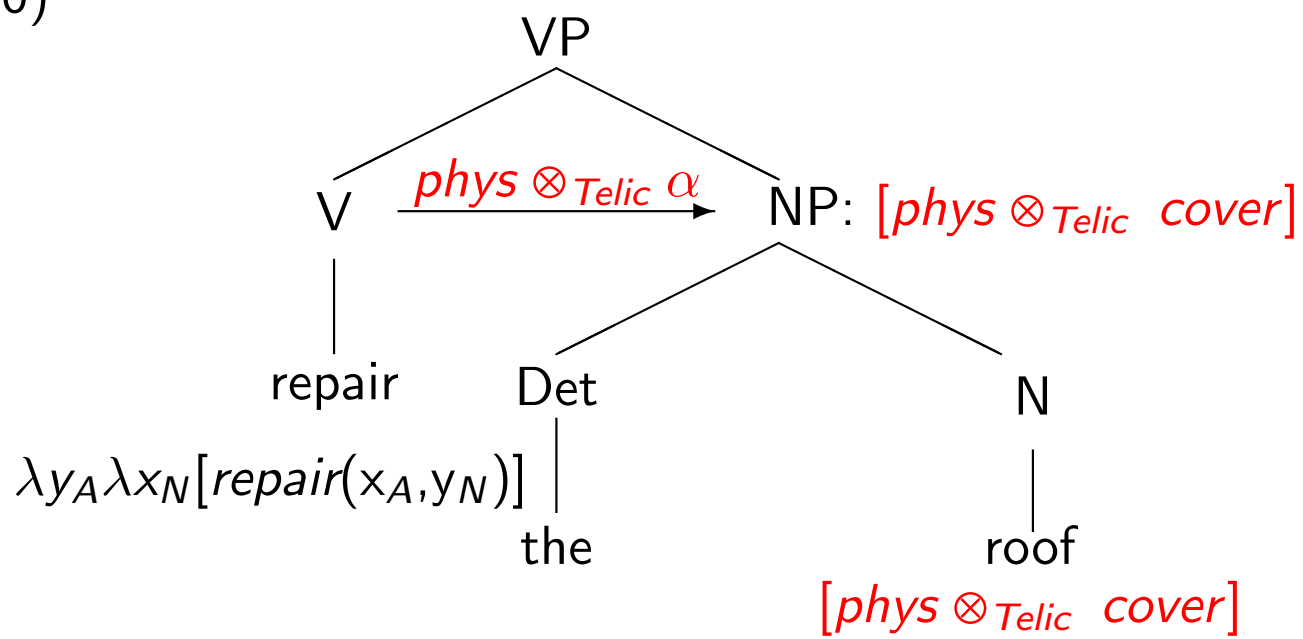
(18)
$$\begin{bmatrix} \textbf{repair} \\ \text{ARGSTR} = \begin{bmatrix} \text{ARG1} = x : \textit{human} \\ \text{ARG2} = y : \textit{phys} \otimes_{\textit{Telic}} \alpha \end{bmatrix} \end{bmatrix}$$

Examples of *repair*-verbs

(19) a. Mary repaired the roof.
b. John fixed the computer.
c. The plumber fixed the sink.
d. The man mended the fence.

Composition with *repair* and NP Object

(20)



Direct object complements for the *repair*-verbs

repair.v			fix.v			mend.v		
damage	107	42.66	pipe	9	11.83	fence	23	32.78
roof	16	20.27	gutter	4	11.45	shoe	10	19.01
fence	10	18.07	heating	5	9.66	puncture	4	18.91
gutter	5	15.87	car	19	9.43	clothes	11	18.68
ravages	4	15.76	alarm	5	9.13	net	8	18.01
hernia	4	15.61	bike	5	9.11	roof	8	16.99
car	23	15.39	problem	23	8.77	car	14	15.45
shoe	10	15.22	leak	3	8.58	way	20	14.26
leak	5	14.96	light	12	8.49	air-conditioning	2	12.71
building	17	14.02	boiler	3	7.96	damage	6	12.71
crack	6	13.99	roof	5	7.27	hole	5	11.38
wall	14	13.77	motorbike	2	7.19	bridge	4	9.68
fault	7	13.56	fault	4	6.91	heart	5	9.6
puncture	3	13.53	jeep	2	6.79	clock	3	9.45
pipe	7	12.89	door	11	6.65	chair	4	9.36
bridge	8	12.19	chain	4	5.48	wall	5	9.27
road	13	12.19	bulb	2	5.15	chain	3	8.3

Selectional Behavior of *repair*-Verbs

(21) *fix.v*
object
a. ARTIFACTUAL: pipe, car, alarm, bike, roof, boiler, lock, engine; heart; light, door, bulb
b. NEGATIVE STATE (condition on the artifact): leak, drip
c. NEGATIVE STATE (general situation): problem, fault

(22) *repair.v*
object
a. ARTIFACTUAL: roof, fence, gutter, car, shoe, fencing, building, wall, pipe, bridge, road; hernia, ligament
b. NEGATIVE STATE (condition on the artifact): damage, ravages, leak, crack, puncture, defect, fracture, pothole, injury
c. NEGATIVE STATE (general situation): rift, problem, fault

(23) *mend.v*
object
a. ARTIFACTUAL: fence, shoe, clothes, roof, car, air-conditioning, bridge clock, chair, wall, stocking, chain, boat, road, pipe
b. ARTIFACTUAL (extended or metaphoric uses): matter, situation; relationship, marriage, relations
c. NEGATIVE STATE (condition on the artifact): puncture, damage, hole, tear

Corpus Evidence Suggests a Different Typing Structure

The verbs select for a negative state of an artifactual type.

(24) a. GENERAL NEGATIVE SITUATION: "fix the problem"
b. CONDITIONS OF THE ARTIFACT: "hole in the wall", "dent in the car".

When the negative relational state is realized, it can either take an artifactual as its object, or leave it implicitly assumed:

(25) a. *repair the puncture / leak*
b. *repair the puncture in the hose / leak in the faucet*

When the artifactual is realized, the negative state is left implicit by default.

(26) a. *repair the hose / faucet*
b. *repair the (puncture in) the hose / (leak in) the faucet*

Revised Typing for *repair*-Verbs

- Selectional properties for the verb *repair* need modification to reflect behavior witnessed from organic data;
- This can be accomplished by positing the negative state as the selected argument of a verb such as *repair*, and the artifactual posited as a *default argument*.

(27)
$$\left[\begin{array}{l}\textbf{repair} \\ \text{ARGSTR} = \left[\begin{array}{l}\text{ARG1} = x : \textit{human} \\ \text{ARG2} = y : \textit{neg\_state}(z) \\ \text{D-ARG1} = z : \textit{phys} \otimes_{\textit{Telic}} \alpha\end{array}\right]\end{array}\right]$$

Co-compositionality

Pustejovsky (1995, 2013)

- A semantic property of a linguistic expression in which all constituents contribute functionally to the meaning of the entire expression.
- A characterization of how a system constructs the meaning from component parts.
- It is the set of computations within a specific system that should be characterized as co-compositional for those expressions.

Co-compositionality

(28) a. John ran.
b. John ran for twenty minutes.
c. John ran two miles.

(29) a. John ran to the store.
b. John ran the race.

There are two senses of *run* that emerge in context with these examples:

(30) a. run_1: manner-of-motion activity, as used in (28);
b. run_2: change-of-location transition, as used in (29);

Co-compositionality

(31) a. Mary *waxed* the car.
b. Mary *waxed* the car clean.

(32) a. John *wiped* the counter.
b. John *wiped* the counter dry.

(33) a. John *baked* the potato.
b. John *baked* the cake.

(34) a. Mary *fried* an egg.
b. Mary *fried* an omelette.

(35) a. John *carved* the stick.
b. John *carved* a statue.

Co-compositionality

- Informally, we can view co-compositionality as the introduction of new information to an expression by the argument, beyond what it contributes as an argument to the function within the phrase.
- Hence, it can be considered an ampliative operation, relative to the function application.

The Case of bake

(36) $\lambda y \lambda x \lambda e \left[\begin{array}{l} \textbf{bake} \\ \text{AS} = \left[\begin{array}{l} \text{A1} = x : \textit{phys} \\ \text{A2} = y : \textit{phys} \end{array} \right] \\ \text{ES} = \left[\text{E1} = e : \textit{process} \right] \\ \text{QS} = \left[\text{A} = \textit{bake}(e, x, y) \right] \end{array} \right]$

(37)

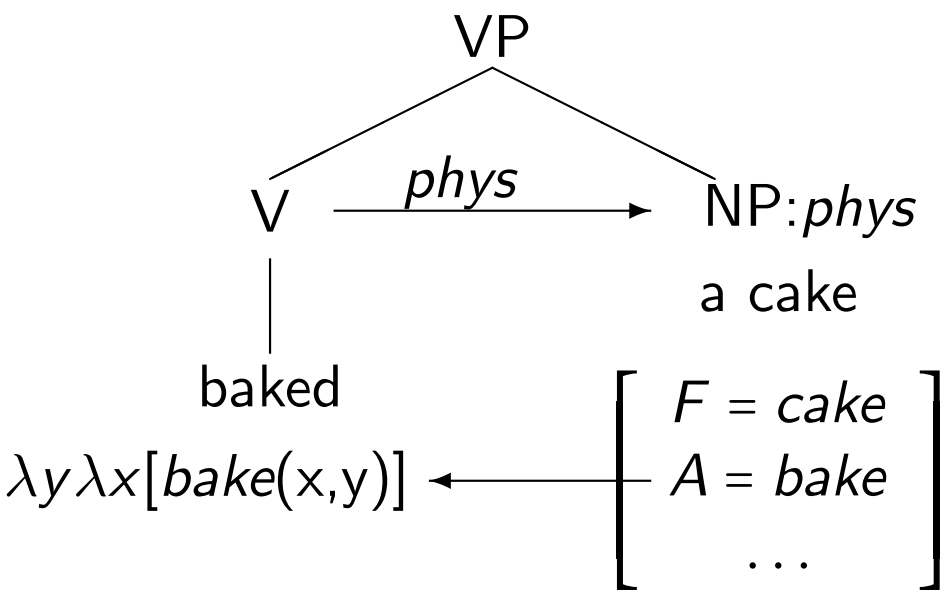
$$\lambda x \exists y \left[\begin{array}{l} \textbf{cake} \\ \text{AS} = \left[\begin{array}{l} \text{ARG1} = x : \textit{phys} \\ \text{D-ARG1} = y : \textit{mass} \end{array} \right] \text{QS} = \left[\begin{array}{l} \text{F} = \textit{cake}(x) \\ \text{C} = \textit{made\_of}(x, y) \\ \text{T} = \lambda z, e[\textit{eat}(e, z, x)] \\ \text{A} = \exists w, e[\textit{bake}(e, w, y)] \end{array} \right] \end{array} \right]$$

The Agentive for *cake* makes reference to the process within which it is embedded in the sentence (i.e., *bake a cake*), which is a case of cospecification.

Co-compositionality

- The direct object cospecifies the verb selecting it, since its type structure makes reference to the governing verb, *bake*.

(38)



Co-compositionality

From the underlying *change-of-state* sense of *bake*, the *creation* sense emerges when combined with the NP *a cake*.

$\exists e_1, e_2, x, y[bake(e_1, j, y) \wedge cake(e_2, x) \wedge made\_of(x, y) \wedge e_1 \leq e_2]$

The operation of co-composition results in a qualia structure for the VP that reflects aspects of both constituents. These include:

(A) The governing verb *bake* applies to its complement;

(B) The complement co-specifies the verb;

(C) The composition of qualia structures results in a derived sense of the verb, where the verbal and complement AGENTIVE roles match, and the complement FORMAL quale becomes the FORMAL role for the entire VP.

Co-compositionality

The derived sense is computed from an operation called *qualia unification*, introduced in Pustejovsky (1995). The conditions under which this operation can apply are stated in (39) below:

(39) FUNCTION APPLICATION WITH QUALIA UNIFICATION: For two expressions, α, of type <a,b>, and β, of type a, with qualia structures QS_α and QS_β, respectively, then, if there is a quale value shared by α and β, $[_{QS_\alpha} \ldots [Q_i = \gamma\ ]]$ and $[_{QS_\beta} \ldots [Q_i = \gamma\ ]]$, then we can define the qualia unification of QS_α and QS_β, $QS_\alpha \sqcap QS_\beta$, as the unique greatest lower bound of these two qualia structures. Further, $\alpha(\beta)$ is of type b with $QS_{\alpha(\beta)} = QS_\alpha \sqcap QS_\beta$.

Co-compositionality

The composition in (38) can be illustrated schematically in (40) below.

(40) $\left[\text{V} \;\; A = bake \;\right] \sqcap \begin{bmatrix} \text{NP} & F = cake \\ & A = bake \end{bmatrix} = \begin{bmatrix} \text{VP} & F = cake \\ & A = bake \end{bmatrix}$

Properties of Co-compositional Derivations

- Within an expression, α, consisting of two subexpressions, α_1 and α_2, i.e., $[_\alpha\ \alpha_1\ \alpha_2]$, one of the subexpressions is an *anchor* that acts as the primary functor;
- Within the argument expression, there is explicit reference to the anchor or the anchor's type (that is, the complement co-specifies the functor);
- The composition of lexical structures results in a derived sense of the functor, within α.

General Co-compositionality

- The derivation for an expression α, is *co-compositional* with respective to its constituent elements, α_1 and α_2, if and only if one of α_1 or α_2 applies to the other, $\alpha_i(\alpha_j)$, $i \neq j$, and $\beta_j(\alpha_i)$, for some type structure β_j within the type of α_j, i.e., $\beta_j \sqsubseteq \mathit{type}(\alpha_j)$.
- $[[\alpha]] = \alpha_i(\alpha_j) \sqcap \beta_j(\alpha_i)$.

The more general characterization of co-compositionality allows us to analyze a number of constructions as co-compositional: *subject-induced coercion* and certain light verb constructions, e.g., *functionally dependent verbs*.

Induced Agency

Wechsler

(41) a. The storm killed the deer.
b. An angry rioter killed a policeman.

(42) a. The glass touched the painting.
b. The curious child touched the painting.

(43) a. The ball rolled down the hill.
b. John rolled down the hill as fast as he could.

(44) a. The room cooled off quickly.
b. John cooled off with an iced latte.

Induced Agency

- Let us characterized "agency", in terms of Qualia Structure, as referring to the potential to act towards a goal.
- For a cognitive agent, such as a *human*, this amounts to associating a set of particular activities, $\mathcal{A}$, as the value of the Agentive role, and
- A set of goals, $\mathcal{G}$, associated with the Telic role in the Qualia for that concept.

(45)
$$\lambda x\left[\begin{array}{l}\textbf{human\_agent}\\ \text{QS} = \left[\begin{array}{l}\text{F} = \mathit{human}(x)\\ \text{T} = \lambda e'[\mathcal{G}\textbf{(e',x)}]\\ \text{A} = \lambda e[\mathcal{A}\textbf{(e,x)}]\end{array}\right]\end{array}\right]$$

Induced Agency

(46) $\lambda y \lambda x \lambda e_2 \lambda e_1 \left[\begin{array}{l} \textbf{kill} \\ \mathrm{AS} = \left[\begin{array}{l} \mathrm{A1} = x : \mathit{phys} \\ \mathrm{A2} = y : \mathit{phys} \end{array} \right] \mathrm{ES} = \left[\begin{array}{l} \mathrm{E1} = e_1 : \mathit{process} \\ \mathrm{E2} = e_2 : \mathit{state} \end{array} \right] \\ \mathrm{QS} = \left[\begin{array}{l} \mathrm{F} = \mathit{dead}(e_2, y) \\ \mathrm{A} = \mathit{kill\_act}(e, x, y) \end{array} \right] \end{array} \right]$

Functionally Dependent Verbs

(47) a. The door opened.
b. Mary opened the door.

(48)
$$\left[\begin{array}{l} \textbf{open} \\ \text{AS} = \left[\begin{array}{l} \text{A1} = x : \mathit{anim} \\ \text{A2} = y : \mathit{phys} \end{array}\right] \\ \text{ES} = \left[\begin{array}{l} \text{E1} = e_1 : \mathit{state} \\ \text{E2} = e_2 : \mathit{state} \\ \text{E3} = e_3 : \mathit{process} \end{array}\right] \\ \text{QS} = \left[\begin{array}{l} \text{F} = \mathit{open}(e_2, y) \\ \text{A} = \mathit{act}(e_3, x, y) \wedge \neg \mathit{open}(e_1, y) \end{array}\right] \end{array}\right]$$

Functionally Dependent Verbs

(49) a. Mary opened the book.
b. They opened the trail.
c. Mary opened the door.
d. Bill opened Microsoft Word.

(50)
$$\left[\begin{array}{l} \textbf{open} \\ \text{AS} = \left[\begin{array}{l} \text{A1} = x : \textbf{anim} \\ \text{A2} = y : \textbf{phys}\ [\text{TELIC} = \alpha] \end{array}\right] \\ \text{ES} = \left[\begin{array}{l} \text{E1} = e_1 : \textbf{state} \\ \text{E2} = e_2 : \textbf{state} \\ \text{E3} = e_3 : \textbf{process} \end{array}\right] \\ \text{QS} = \left[\begin{array}{l} \text{F} = \alpha(e_2, y) \\ \text{A} = act(e_3, x, y) \wedge \neg\alpha(e_1, y) \end{array}\right] \end{array}\right]$$

Lexicon and World Knowledge

Synthesis from Jezek 2016, The Lexicon: An Introduction, OUP, ch. 2

- Words denote classes of entities and are associated with conceptual categories, for example a *dog* denotes an *animal*, a *table* denotes an *artifact*, *bread* denotes a kind of *food*, a *park* denotes a *location*, *run* denotes a *process*, *love* denotes a *state*, and so forth.
- A conceptual category may be analyzed as a set of salient attributes or properties, for example the concept *dog* has properties: breathes, barks, wags its tail, has fur, and so forth (Baroni and Lenci, 2008, Poesio and Almuhareb 2008).
- But which properties of a concept are genuinely distinctive and enter into the **lexical make-up** of a word and which ones do not?

Lexicon and world knowledge

- There are deep controversies regarding what piece of information associated with a word should enter into its definition, and constitute what is called its **lexical information**.
- Traditionally, it is assumed that **world / commonsense / encyclopedic knowledge** should be excluded.
- This amounts to the large body of knowledge that people possess about the entities and events denoted by words as a result of their experience of the world.
- It has to do with the speaker's perception of the world, and the analogies speakers establish between objects and events, rather than with their linguistic knowledge.

Lexicon and world knowledge

- The distinction is very difficult to draw.
- According to some authors, it is not even necessary.
- Others believe it should be conceived as a continuum rather than a dichotomy.
- Opinions differ because there is **no consensus about what criteria** must be satisfied for a piece of information to qualify as encyclopedic knowledge instead of linguistic meaning, or vice versa.
- Those who make a distinction take different positions on the subject.

Minimalism

- According to the **minimalist position**, nothing of what we know about, say, the entity called *dog* is part of the lexical information associated with the word *dog*, except for those features that are necessary to define it as a domestic animal (as opposed to a wild one) and allow us to distinguish it from other entities falling into the same category.

Maximalism

- According to the **maximalist position**, the opposite is instead true, that is, the lexical information associated with the word *dog* incorporates our knowledge that dogs can be aggressive (and therefore bite and attack), that they have an acute sense of smell, that they like to chase cats, and so on.
- This additional knowledge about dogs is what we know from our individual experience.

No distinction

- A radical position is that taken by those who hold that the distinction between lexical information and world / encyclopedic / commonsense knowledge is artificial or useless, and should be eliminated.
- According to this position, words **give access to concepts**, and all the properties that enter into the constitution of a concept can in principle be exploited in language through the use of words.
- The contexts in which words are used determine which property/ies of the concept is/are activated in the specific case.

No distinction

- The lexicon is interpreted as the **access node** into the vast repository of information associated with conceptual categories.
- This position is dominant in cognitive semantics and pragmatics (Sperber and Wilson 1995; Carston 2002), where **context-dependency** is dealt with at the conceptual level (instead of at the lexical level).

Ad-hoc concepts (Barsalou 1983, 2010; Wilson and Carston 2007)
General extenders ("whales, candlelight and stuff like that" Overstreet 1999).

Meaning Eliminativism (ME)

- Extreme version of contextualism (Recanati, 2004).
- We don't need abstract scheme in the form of context-independent linguistic meaning as input to the composition process.
- This can proceed **without the help of conventionalized context-independent word meanings**.
- ME gets rid of abstract meaning in favour of observed occasion of particular uses.

Intermediate Position

- A third position is intermediate, and **linguistically motivated**.
- According to this position, the information encoded in a word amounts to those aspects that influence how the word behaves grammatically and how it may be interpreted in different contexts.
- This position is adopted *inter alia* in GL.

Distributional Methodology

- One way of identifying these aspects is to examine the distribution of words in context.
- For example, the expression *quick coffee* means 'coffee which is drunk quickly': the meaning of *coffee* **contributes information regarding the activity of drinking**.
- This appears not to be the case with the word *water* which, in the context of *quick* means '**that moves quickly**' rather than 'which is drunk quickly'.
- According to this methodology, if **a piece of knowledge is exploited** in our understanding of linguistic expressions, it is likely to be part of its lexical information.

Pragmatics of Contextualizing the Event

1. It's raining.
   here now
2. You're not going to die.
   soon, from your cold
3. I had a big breakfast.
   recently

Viewpoints

- Free enrichment: Any utterance may contain unarticulated constituents which are not part of the LF of the sentence, but are needed to determine a truth-theoretic interpretation. (Recanati, 2002, Carston, 2002)
- Pragmatic saturation: All truth-conditional effects of extra-linguistic context can be traced to logical form. (Stanley, 2000)
- Discourse Structure: A sentential LF embeds within a discourse structure, DRS, where constraints on licensing and accessibility of discourse referents are determined and computed. (DRT, SDRT, DPL)
- GL in Context: Combines parametric and non-parametric factors to built a context.

GL Enriches the Domain Contributing to Contextualized Meaning

- GL's multiple dimensions of semantic interpretation enhance traditional notions of compositional meaning;
- Qualia Structure and Event Structure provide presuppositional aspects of interpretation lacking in most model theoretic treatments of NL semantics;
- Coercion and co-composition can be seen as mechanisms operating at the discourse and text level.
- Corpus Data and evidence-based analysis can help reveal how these mechanisms play out in actual contexts, and point to inconsistencies in theoretical claims.
- Corpus-driven analysis and evidence-based theory construction drives more expressive and realistic frameworks for lexical resources.

Conclusion

Conclusion


danke
謝謝
ngiyabonga
teşekkür ederim
спасибо
dank je
gracias
thank you
tapadh leat
bedankt
hvala
mauruuru
mochchakkeram
dziękuje
go raibh maith agat
obrigado
sukriya
kop khun krap
terima kasih
arigalò
dakujem
grazie
merci
мерси
감사합니다