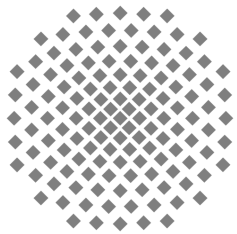
Implicit Events for Event/Entity-Ambiguous Nouns

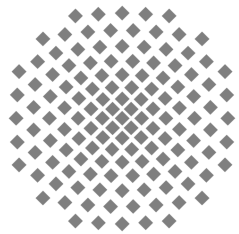
Alessandra Zarcone, Sebastian Padó

Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart
SFB 732 - D6

20/07/2010



Outline



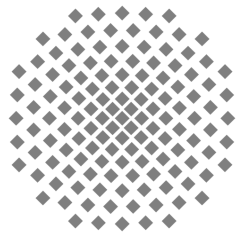
Covert events
Traditional account

Some verb-object pairs require the recovery of covert events (**CE**):

Entity-denoting objects (EN) vs. event-denoting objects (EV)

EV: begin the **afternoon**	EN: begin the **newspaper**
→ ✓ begin(afternoon)	→ × begin(newspaper)
→ × begin(**CE**(afternoon))	→ ✓ begin(**CE**(newspaper))
	→ begin **reading** the newspaper

1. **the trigger problem**
   - sortal trigger hypothesis: CEs are triggered by the ontological type of the object (EN vs. EV) (cfr. Pustejovsky 1995, Jackendoff 1997);
2. **the range problem**
   - qualia structure hypothesis: CEs are retrieved from the *qualia structure* of the noun's lexical entry (Pustejovesky 1995)
   - range: 1-2 CEs, e.g. *newspaper* → *writing* or *reading*



Covert events
Our questions

1. **the trigger problem**
   type clash only factor responsible for evoking CEs?

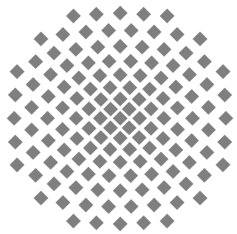
   Joe is a famous wrestler. He really enjoyed the fight → enjoy(fight)
   Joe is a wrestling fan. He really enjoyed the fight → enjoy(**watch**(fight))

   Joe began the translation → begin(translation), begin(**CE**(translation))

2. **the range problem**
   the qualia structure hypothesis undergenerates the CE range
   (Lascarides and Copestake 1998, Egg 2005)

   I need some help with my food → buying, finding, eating, cooking

   My goat eats anything. He really enjoyed your book.



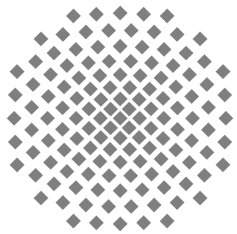
Covert events
Our agenda

1. **the trigger problem**
   Traxler et al. (2002), Self-paced reading (SPR) study
   - begin-verbs vs. spot-verbs and EN objects vs. EV objects
   - longer reading times at Obj. +1 position for CE condition: interaction Verb * Object:
     $RT_{begin,EN} > RT_{begin,EV} \approx RT_{spot,EV} \approx RT_{spot,EN}$

   1. the Obj + 1 problem: contexts are not uniform: *after school* / *the reporter told us* / *David said to Joan* / *and returned to her room*...
   2. ... and what about EN/EV type objects?

2. **the range problem**
   how to study recovered CEs in a behavioral experiment?

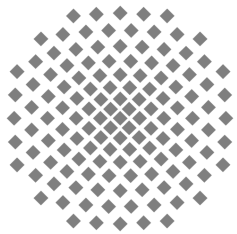
SPR	Web studies
+ on-line processing	− off-line measures
+ more natural tasks	− metalinguistic analysis
− CEs in absentia	+ CE elicitation



Covert events
Our experiments

1. **the trigger problem**:
   - a SPR study to verify the sortal trigger hypothesis (designed, to do)
     1. EN/EV objects
     2. the Obj + 1 problem (now even more crucial)
   - Web experiments (done)
     1. evaluate materials for SPR
     2. verify the sortal trigger hypothesis
2. **the range problem**:
   Web experiment 2 to study recovered CEs (done)

Web experiments delivered using the crowdsourcing paradigm (Snow et al. 2008), for fast and affordable collection of judgments over the web.



Self-paced reading study
Description and materials

aim: 1. **the trigger problem**:
verify the sortal trigger hypothesis

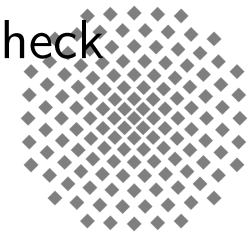
materials: 10 noun triplets x 6 conditions = 60 sentences:

EN: Jim **began/spotted the magazine from the camp on the hill**.
EV: Al **began/spotted the ceremony from the camp on the hill**.
EN/EV Nick **began/spotted the conquest from the camp on the hill**.

design: 2 (begin-verbs and spot-verbs) x 3 (EN, EV, EN/EV)

triplet selection: triplets balanced for

- object length and frequency (Francis and Kucera 1967)
- co-occurrence frequency with begin-verbs and spot-verbs (ukWaC corpus, Ferraresi et al. 2008)
- Web experiment 0: threefold expert annotation check
  - Krippendorff's $\alpha = .71$ (very good agreement)
  - weighted $\alpha = .79$



Web experiment 1
Description

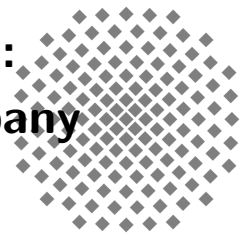
aim: evaluation of materials for SPR

- check for non-expert annotation of objects (EN, EV, EN/EV)
- effect of Obj + 1 position on sortal categorization

participants: 14 participants from the US

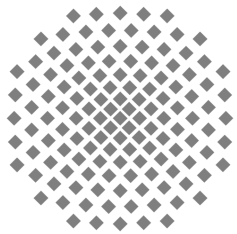
procedure: web-based annotation experiment
Jan enjoyed **the automobile**
(possible answers: EN, EV; they could check either or both)

materials: 60 sentences for the SPR study, in 3 "Obj + 1" contexts:
no Obj + 1:
Jan enjoyed the automobile **short Obj + 1:**
Jan enjoyed the automobile **on the premises full Obj + 1:**
Jan enjoyed the automobile **on the premises of the company**



Web experiment 1
Results

- reasonably good agreement (weighted $\alpha = .52$)
- very good agreement with the Gold Standard ($\alpha = .70$, weighted $\alpha = .79$)
- ruled out effect of the Obj + 1 context and of the verb on the sortal type assigned to the object
  - Binomial logistic regression model 1: *entity* $\sim$ *context* + *verb*
    Context: binomial $p = .3621 \rightarrow$ no effect
    Verb: $z = 1.491, p = .1359 \rightarrow$ no effect
  - Binomial logistic regression model 2: *event* $\sim$ *context* + *verb*
    Context: binomial $p = .6138 \rightarrow$ no effect
    Verb: $z = -0.504, p = .614 \rightarrow$ no effect



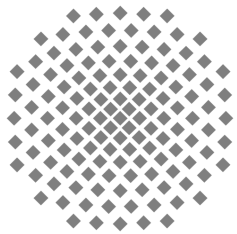
Web experiment 2
Description

aim:
1. **the trigger problem** evaluate correlation between EN/EV and CE/noCE
2. **the range problem** elicit CEs and explore their range

participant: 15 participants from the US

procedure: Jan enjoyed **the automobile**
does the sentence involve an additional activity that is not mentioned in the sentence?
(answers: *additional activity* or *no additional activity*)
when they answered *additional activity*,
participants were asked to provide examples

materials: same of Web experiment 1 and SPR



Web experiment 2
Results

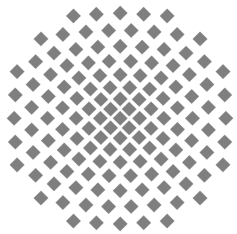
Agreement and accuracy:

- rather low agreement
  ($\alpha = .35$, ruling out EN/EV ambiguous objects, $\alpha = .36$)
- good agreement with the Gold Standard ($\alpha = .60$)

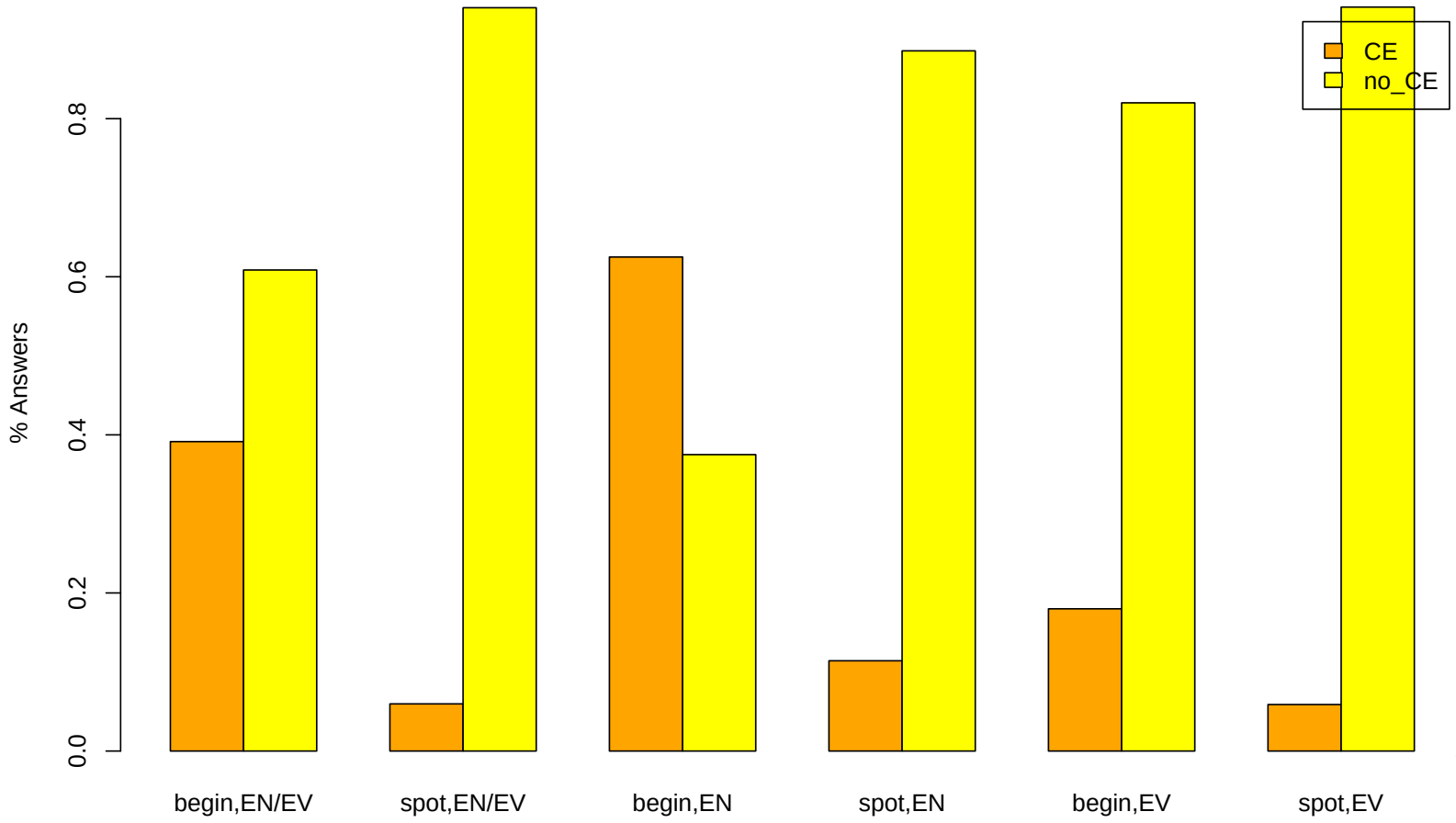
Binary decision *additional activity*/*no additional activity* (CE/noCE):

- significant effect of object_type and verb_type with interaction
  - Binomial logistic regression model: $answer \sim obj\_type * verb\_type$;
  - Obj_type: binomial $p < .001 \rightarrow$ significant effect
    Verb_type: $z = -8.322, p < .001 \rightarrow$ significant effect
    Interaction: binomial $p < .001 \rightarrow$ significant effect

... does this confirm the sortal trigger hypothesis?
$\rightarrow$ item-wise analysis



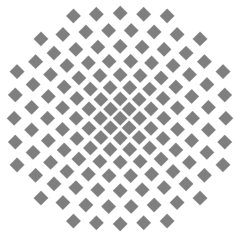
Web experiment 2
Results



Web experiment 2

A closer look: verbs

		CE	no CE
begin-verb	continued	0.71	0.29
begin-verb	began	0.62	0.38
begin-verb	finished	0.56	0.44
begin-verb	started	0.47	0.53
begin-verb	tried	0.47	0.53
begin-verb	enjoyed	0.38	0.62
begin-verb	preferred	0.32	0.68
spot-verb	considered	0.29	0.71
begin-verb	ended	0.22	0.78
spot-verb	remembered	0.18	0.82
begin-verb	endured	0.16	0.84
spot-verb	recalled	0.15	0.85
spot-verb	disdained	0.10	0.90
begin-verb	savored	0.08	0.92
spot-verb	discussed	0.04	0.96
spot-verb	approved	0.02	0.98
spot-verb	reviewed	0.00	1.00
spot-verb	organized	0.00	1.00
spot-verb	spotted	0.00	1.00
spot-verb	prepared	0.00	1.00

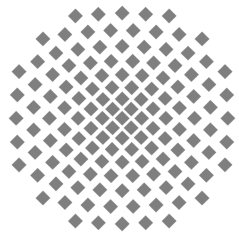


Web experiment 2

A closer look: VPs

type V	type obj	v	obj	CE	no CE
begin-verb	EN	began	the newspaper	0.89	0.11
begin-verb	EN/EV	began	the breakfast	0.81	0.19
begin-verb	EN	tried	the tent	0.73	0.27
begin-verb	EN	enjoyed	the automobile	0.50	0.50
begin-verb	EV	continued	the season	0.46	0.54
begin-verb	EN/EV	enjoyed	the translation	0.39	0.61
spot-verb	EN	remembered	the brandy	0.34	0.66
begin-verb	EV	enjoyed	the conference	0.24	0.76
spot-verb	EV	considered	the debate	0.22	0.78
spot-verb	EN/EV	disdained	the blessing	0.21	0.79
spot-verb	EV	remembered	the revolt	0.10	0.90
spot-verb	EN/EV	remembered	the shower	0.08	0.92
begin-verb	EN	savored	the butter	0.07	0.93
spot-verb	EN	disdained	the portrait	0.07	0.93
begin-verb	EN/EV	endured	the shower	0.07	0.93
begin-verb	EV	endured	the revolt	0.03	0.97
begin-verb	EV	ended	the ceremony	0.00	1.00
spot-verb	EV	discussed	the expedition	0.00	1.00
spot-verb	EN/EV	prepared	the harvest	0.00	1.00
spot-verb	EV	prepared	the holiday	0.00	1.00
spot-verb	EN	spotted	the magazine	0.00	1.00
spot-verb	EN	prepared	the package	0.00	1.00

type V	type obj	v	obj	CE	no CE



Web experiment 2

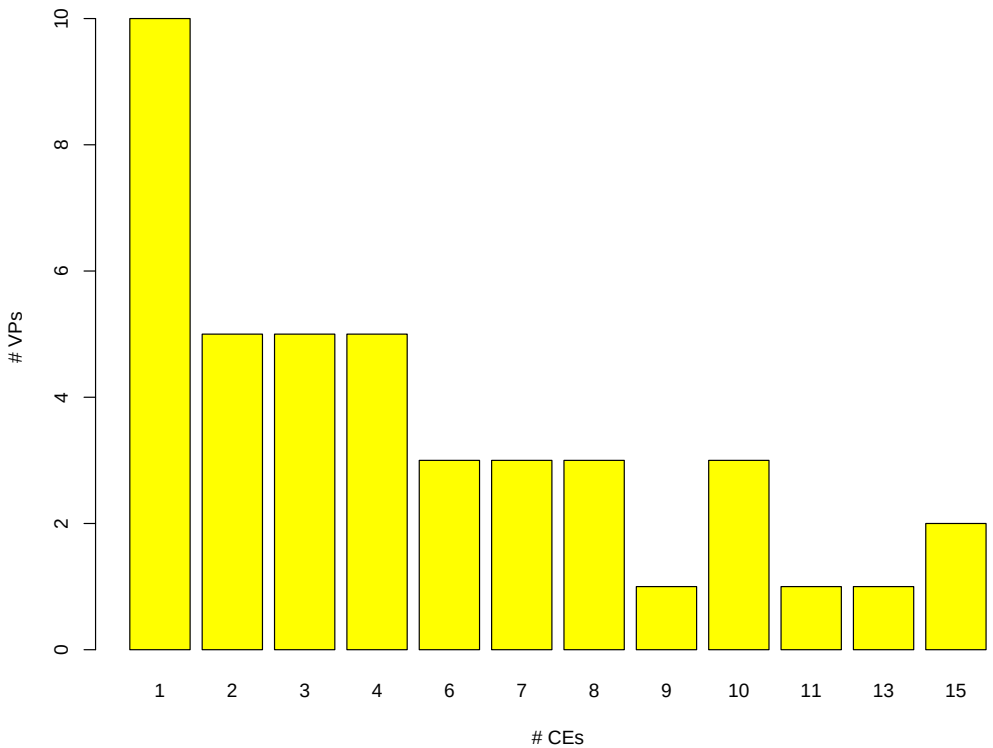
Range of elicited CEs

CEs elicited per VP item: range 1-15, mean 5

EN: start the portrait → 10 CEs: paint (x20), draw (x4), critique (x3), hang (x2), model (x2), sketch (x2), admire, fix, pose for, review

EN/EV: finish the harvest → 15 CEs: gather (x5), collect (x4), plan (x3), reap (x3), sell (x3), load (x2), store (x2), cook, eat, enjoy, jar, package, pick, pull, ship

EV: enjoy the conference → 4 CEs: attend (x3), hold (x2), participate in, watch

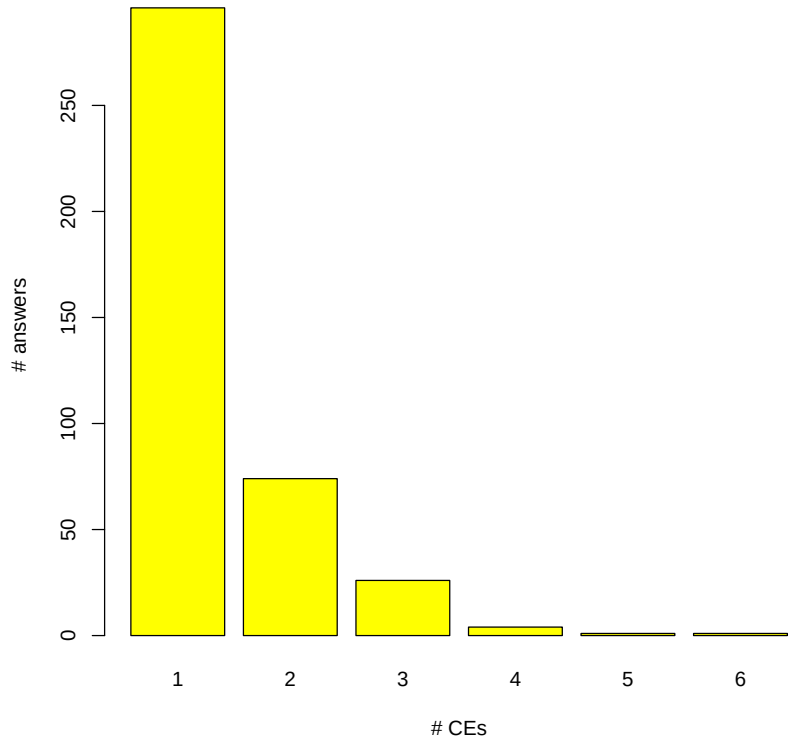


Web experiment 2

Range of elicited CEs

CEs elicited per VP item per participant: range 1-6, mean 1.4

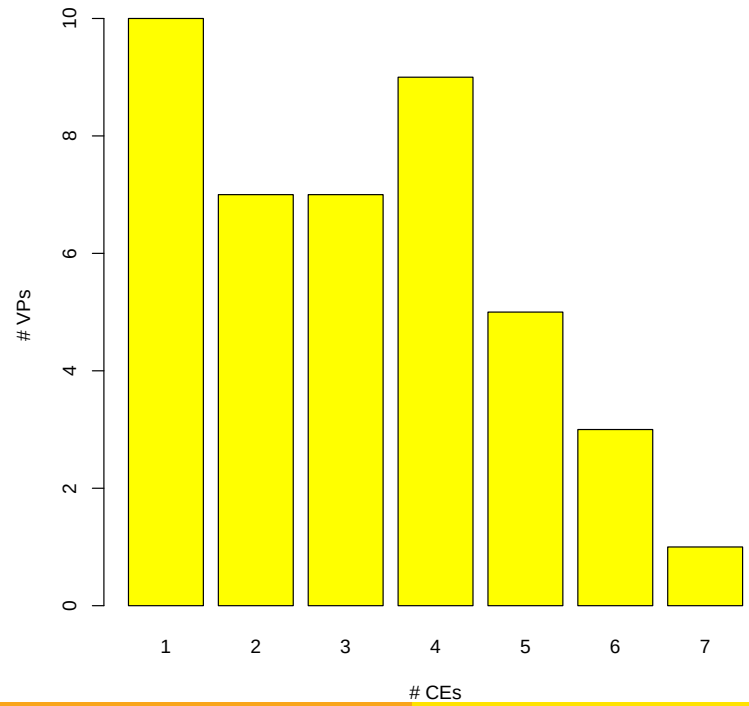
EN: begin the newpaper → 6 CEs: deliver, edit, print, read, sell, write
EN/EV: begin the breakfast → 3 CEs: cook, eat, serve
EV: continue the season → 3 CEs: coach, play, watch



Web experiment 2
Range of elicited CEs

CEs elicited per VP when participants gave only one answer:
range 1-7, mean 3.2;

EN: consider the butter → 6 CEs: eat (x4), add, buy, churn, cook with, eat, make, melt
EN/EV: prefer the collection → 6 CEs: view (x3), buy, discuss, polish, study, watch
EV: start the semester → 3 CEs: spend, teach, join

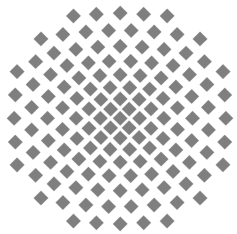


Conclusions

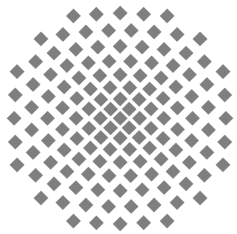
The sortal trigger hypothesis and the qualia structure hypothesis do not seem enough to explain CE recovery:

1. sortal trigger hypothesis:
   - CEs for begin-verbs with EV obj
   - also for EN/EV obj, the EV reading does not block CE recovery
   - no clear-cut distinction between begin-verbs and spot-verbs
2. qualia structure hypothesis:
   - wide range of CEs elicited per context
   - the range is also fairly wide when participants only give one answer
   - elicited CEs $>$ qualia structure events

Plausibility hypothesis: plausibility-driven CEs retrieval



Thank you!



Bibiliography

- Egg, M. (2005). *Flexible semantics for reinterpretation phenomena*. Stanford: CSLI Publications.
- Jackendoff, R. (1997). *The architecture of the language faculty*. Cambridge: MIT Press.
- Lascarides, A., & Copestake, A. (1998). Pragmatics and Word Meaning. *Journal of Linguistics, 34*(2), 55-105.
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge: MIT Press.
- Traxler, M. J., Pickering, M. J., & McElree, B. (2002). Coercion in sentence processing: evidence from eye-movements and self-paced reading. *Journal of Memory and Language, 47*, 530-547.