Scope in an incremental context

Lecture 5: representation and computation

Asad Sayeed

University of Gothenburg

WARNING: Every linguist speculated about a phenomenon.

Part 1: the time course of processing

Linking theories

We've implicitly looked at two ways to characterize the "effort" of processing ("linking theories")

Linking theories

We've implicitly looked at two ways to characterize the "effort" of processing ("linking theories")

- The measurement – what factors *reflect* the processing effort experimentally.

Linking theories

We've implicitly looked at two ways to characterize the "effort" of processing ("linking theories")

- The measurement – what factors *reflect* the processing effort experimentally.
  - Reading time, ERP measurement, eye gaze, etc.

Linking theories

We've implicitly looked at two ways to characterize the "effort" of processing ("linking theories")

- The measurement – what factors *reflect* the processing effort experimentally.
  - Reading time, ERP measurement, eye gaze, etc.
- The model – what factors *account for* the observed effect.

Linking theories

We've implicitly looked at two ways to characterize the "effort" of processing ("linking theories")

- The measurement – what factors *reflect* the processing effort experimentally.
  - Reading time, ERP measurement, eye gaze, etc.
- The model – what factors *account for* the observed effect.
  - Operations on formal syntactic/semantic representations (the "Scope Fairy").

Linking theories

We've implicitly looked at two ways to characterize the "effort" of processing ("linking theories")

- The measurement – what factors *reflect* the processing effort experimentally.
  - Reading time, ERP measurement, eye gaze, etc.
- The model – what factors *account for* the observed effect.
  - Operations on formal syntactic/semantic representations (the "Scope Fairy").
  - Access to world knowledge/memory (the "Pragmatics Fairy").

Linking theories

We've implicitly looked at two ways to characterize the "effort" of processing ("linking theories")

- The measurement – what factors *reflect* the processing effort experimentally.
  - Reading time, ERP measurement, eye gaze, etc.
- The model – what factors *account for* the observed effect.
  - Operations on formal syntactic/semantic representations (the "Scope Fairy").
  - Access to world knowledge/memory (the "Pragmatics Fairy").
  - Language model expectation (less frequent = more effort; a . . . Statistics Fairy?).

Incremental quantifier scope ambiguity

Like most linguistic problems, becomes more clear when we add context.

Incremental quantifier scope ambiguity

Like most linguistic problems, becomes more clear when we add context.

Every

Incremental quantifier scope ambiguity

Like most linguistic problems, becomes more clear when we add context.

Every child

Incremental quantifier scope ambiguity

Like most linguistic problems, becomes more clear when we add context.

Every child climbed

Incremental quantifier scope ambiguity

Like most linguistic problems, becomes more clear when we add context.

Every child climbed a

Incremental quantifier scope ambiguity

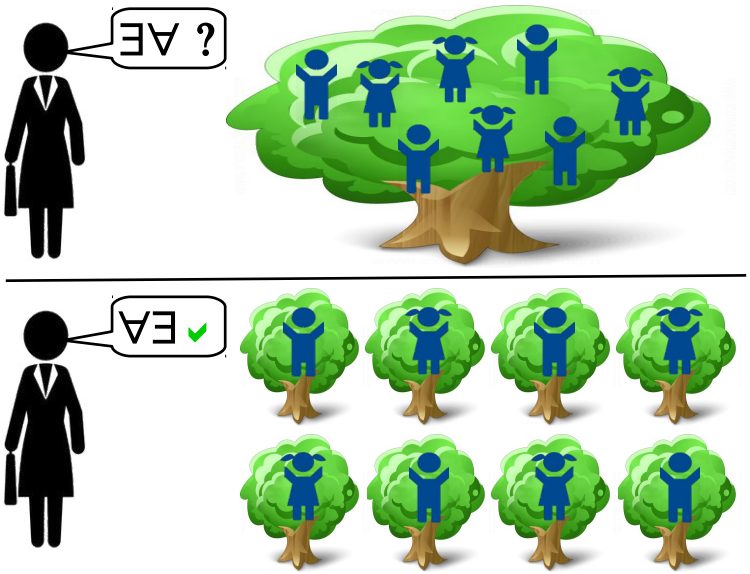
Like most linguistic problems, becomes more clear when we add context.

Every child climbed a tree.

Incremental quantifier scope ambiguity

Like most linguistic problems, becomes more clear when we add context.

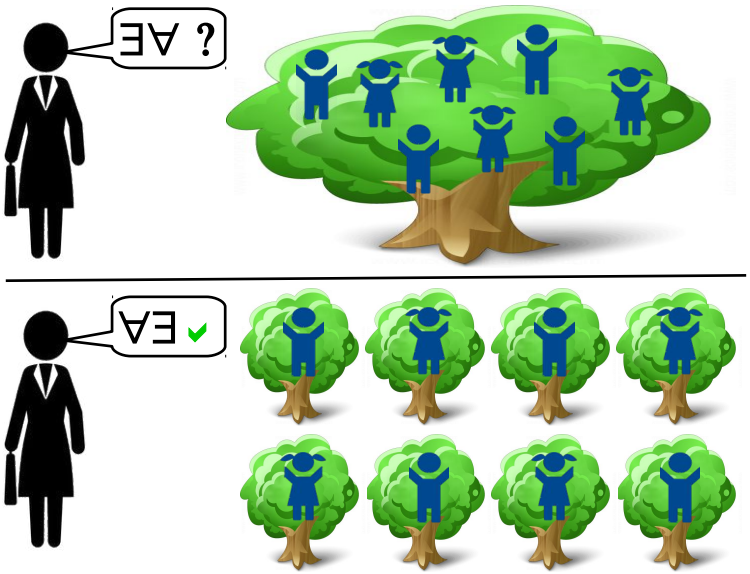
Every child climbed a tree.



Incremental quantifier scope ambiguity

Like most linguistic problems, becomes more clear when we add context.

Every child climbed a tree.



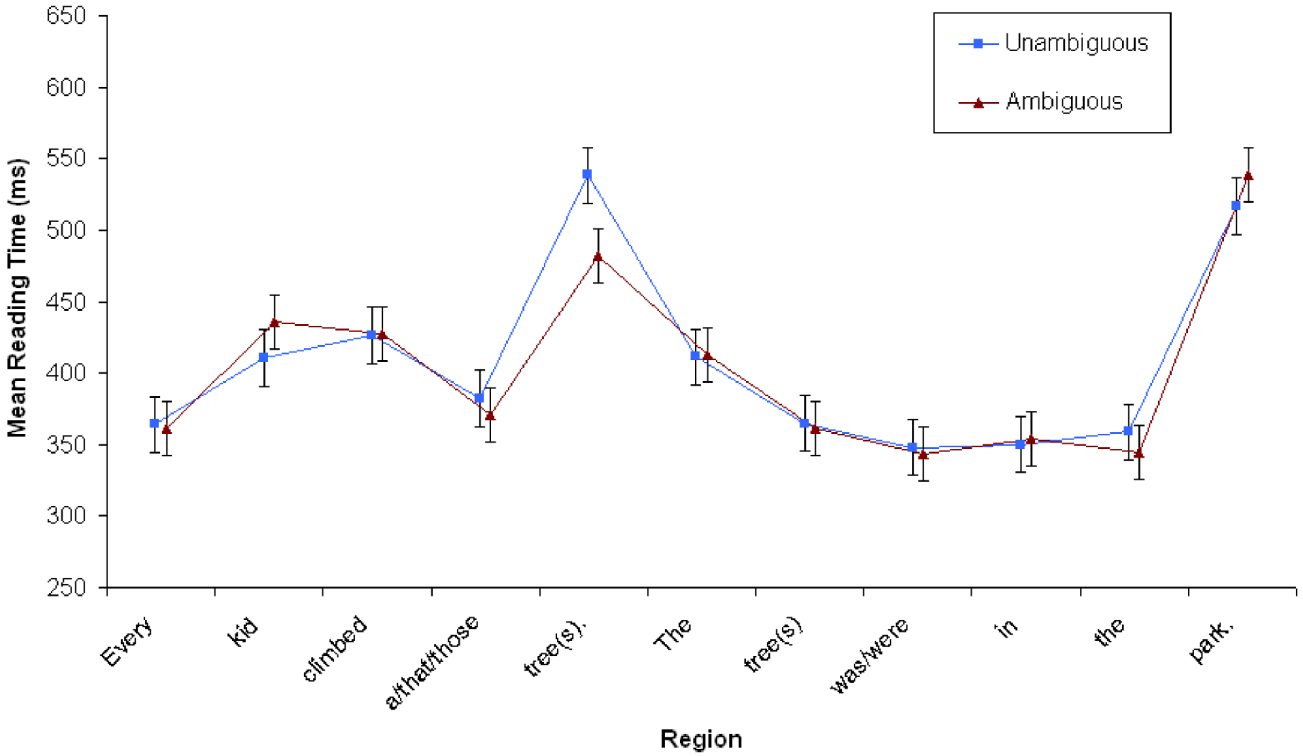
Perhaps we can use the creation of expectations about set cardinality to investigate what is "really going on".

Two-quantifier experiments

Dwivedi [2013]: concludes that simple heuristics and world-knowledge dominates, algorithmic processing available as last resort.

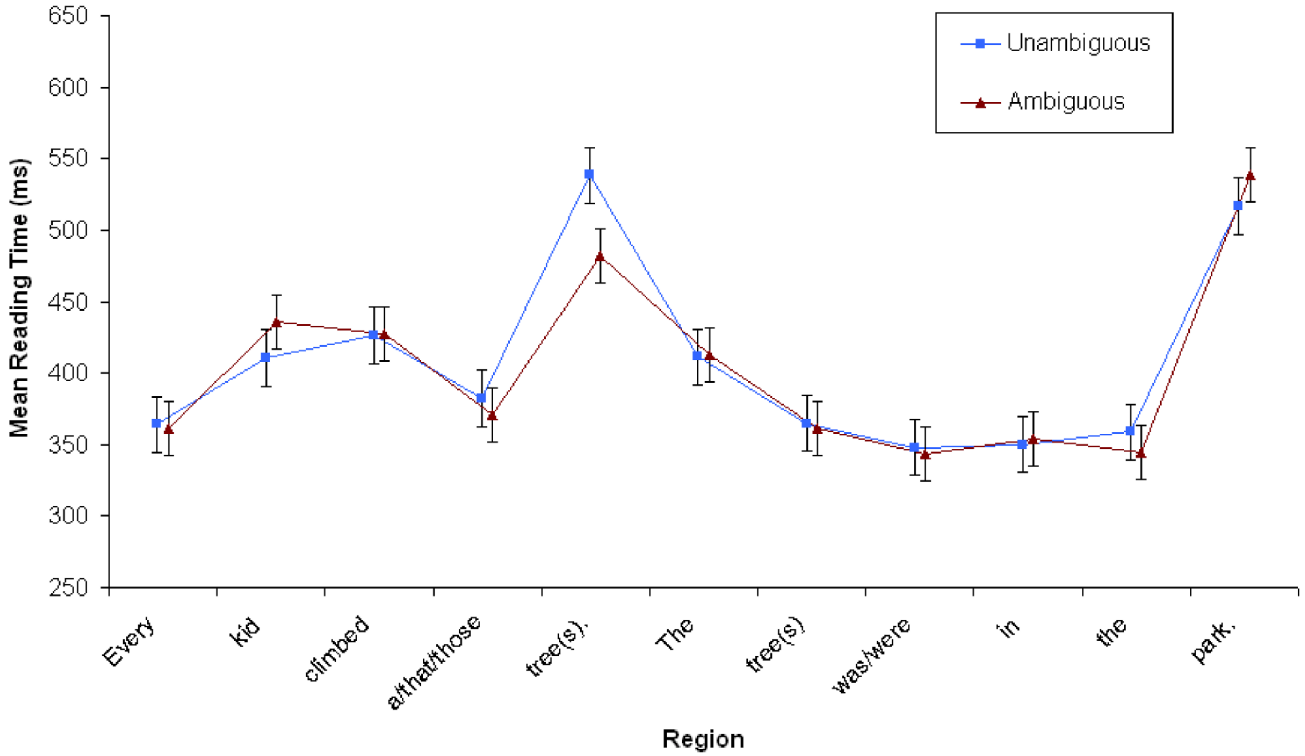
Two-quantifier experiments

Dwivedi [2013]: concludes that simple heuristics and world-knowledge dominates, algorithmic processing available as last resort.



Two-quantifier experiments

Dwivedi [2013]: concludes that simple heuristics and world-knowledge dominates, algorithmic processing available as last resort.



The *unambiguous* sentence takes *longer* to read, no other differences.

Two-quantifier experiments

What could account for this?

Two-quantifier experiments

What could account for this?

- The Scope Fairy? – why would it be "harder" to have an *unambiguous* item (that won't e.g. move to SpecCP as in Quantifier Raising)?

Two-quantifier experiments

What could account for this?

- The Scope Fairy? – why would it be "harder" to have an *unambiguous* item (that won't e.g. move to SpecCP as in Quantifier Raising)?
- The Pragmatics Fairy? – What is different about the situation described by the definite article that would make it harder to process?

Two-quantifier experiments

What could account for this?

- The Scope Fairy? – why would it be "harder" to have an *unambiguous* item (that won't e.g. move to SpecCP as in Quantifier Raising)?
- The Pragmatics Fairy? – What is different about the situation described by the definite article that would make it harder to process?
- The Statistics Fairy? – Is it really less frequent for a definite determiner to be the narrow scope under a universal?

Two-quantifier experiments

Dwivedi [2013]:

- Scope order **underspecification** filled in by world knowledge.

Two-quantifier experiments

Dwivedi [2013]:

- Scope order **underspecification** filled in by world knowledge.
- **Problem**: cannot detect effect underspecificaton status with only two quantifiers.

Two-quantifier experiments

Dwivedi [2013]:

- Scope order **underspecification** filled in by world knowledge.
- **Problem**: cannot detect effect underspecificaton status with only two quantifiers.
  - At continuation sentence, you can't entirely distinguish reanalysis from violation of world knowledge expectation

Scope ambiguity in processing

Consider the following sentence and continuations [Dotlačil and Brasoveanu 2015]:

(1) A caregiver comforted a child every night.

   a. The caregiver wanted the child to get some rest.
   b. The caregivers wanted the child to get some rest.
   c. The caregiver wanted the children to get some rest.
   d. The caregivers wanted the children to get some rest.

There are four plausible readings of the first sentence, based on the scope of "every night".

Scope ambiguity in processing

Consider the following sentence and continuations [Dotlačil and Brasoveanu 2015]:

(1) A caregiver comforted a child every night.

   a. The caregiver wanted the child to get some rest.
   b. The caregivers wanted the child to get some rest.
   c. The caregiver wanted the children to get some rest.
   d. The caregivers wanted the children to get some rest.

There are four plausible readings of the first sentence, based on the scope of "every night".
With three quantifiers: can investigate whether there is a preferred specified order in incremental context.

Scope ambiguity in processing

(2) A caregiver comforted a child every night.

- The linear order is $\exists = \exists > \forall$, ie, child and caregiver singular.

Scope ambiguity in processing

(2) A caregiver comforted a child every night.

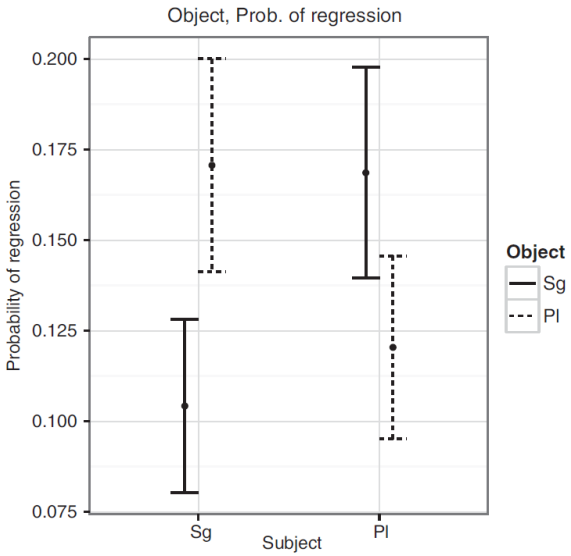
- The linear order is $\exists = \exists > \forall$, ie, child and caregiver singular.
- English speakers can be easily prompted to all other orders.

Scope ambiguity in processing

(2) A caregiver comforted a child every night.

- The linear order is $\exists = \exists > \forall$, ie, child and caregiver singular.
- English speakers can be easily prompted to all other orders.
- Dotlačil and Brasoveanu (2015):

The caregiver wanted the child to get some rest.
The caregivers wanted the child to get some rest.
The caregiver wanted the children to get some rest.
The caregivers wanted the children to get some rest.

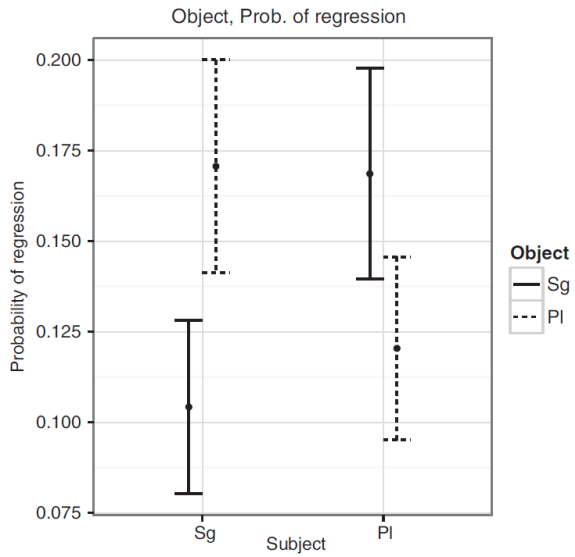


Scope ambiguity in processing

(2) A caregiver comforted a child every night.

- The linear order is $\exists = \exists > \forall$, ie, child and caregiver singular.
- English speakers can be easily prompted to all other orders.
- Dotlačil and Brasoveanu (2015):

The caregiver wanted the child to get some rest.
The caregivers wanted the child to get some rest.
The caregiver wanted the children to get some rest.
The caregivers wanted the children to get some rest.



Evidence for algorithmic processing (as opposed to purely pragmatic considerations).

Part 2: sources of effort

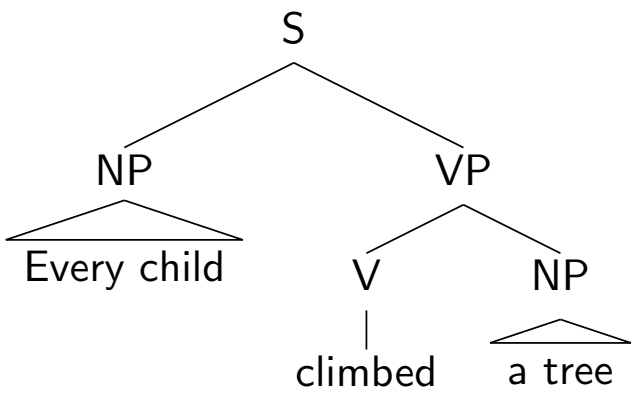
Scope ambiguity as covert syntax

Quantifier raising (QR) approach: use syntactic structure to represent limits of scope structure.

Scope ambiguity as covert syntax

Quantifier raising (QR) approach: use syntactic structure to represent limits of scope structure.

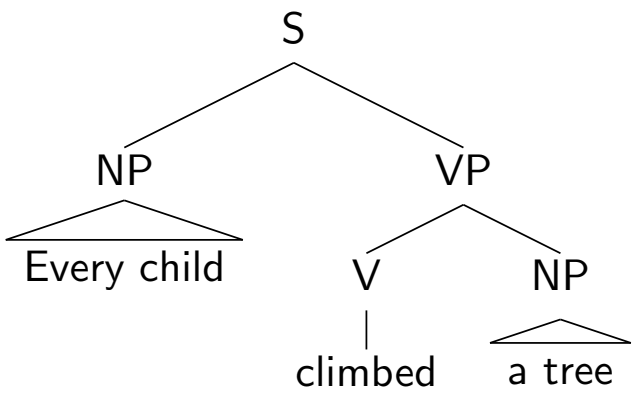
Linear reading:



Scope ambiguity as covert syntax

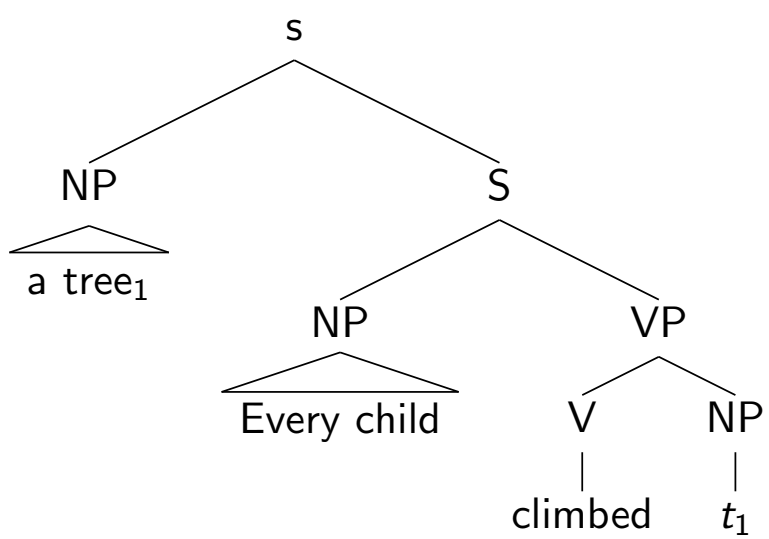
Quantifier raising (QR) approach: use syntactic structure to represent limits of scope structure.

Linear reading:



$\Rightarrow$

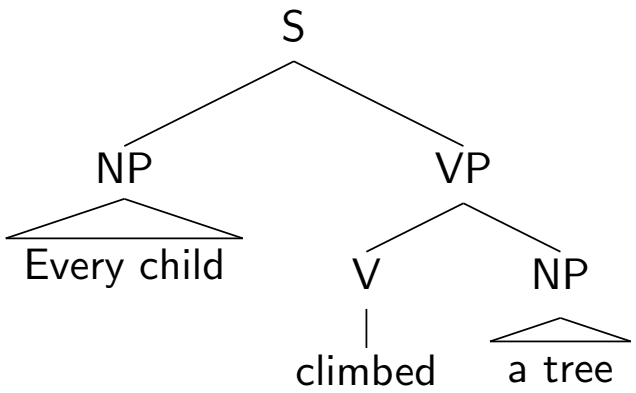
Inverse reading:



Scope ambiguity as covert syntax

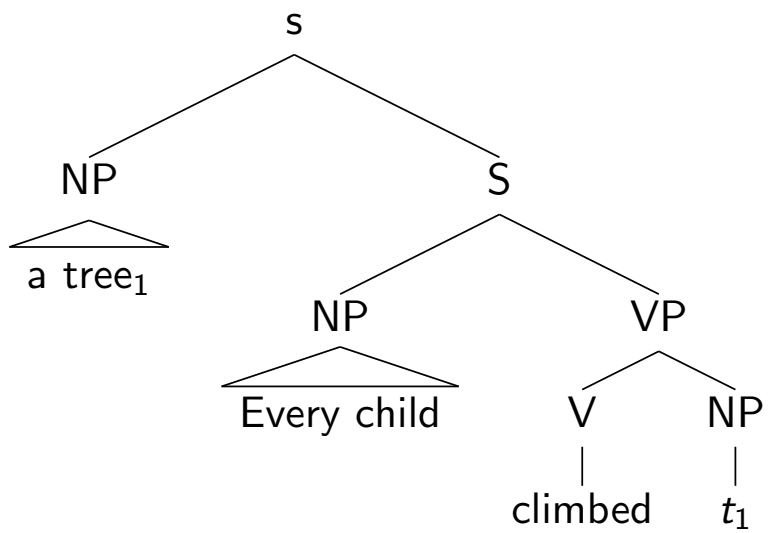
Quantifier raising (QR) approach: use syntactic structure to represent limits of scope structure.

Linear reading:



$\Rightarrow$

Inverse reading:



Captures intuition that inverse reading should be harder – extra step.

Underspecification approaches

Problem: syntactic commitment of QR approach

Underspecification approaches

Problem: syntactic commitment of QR approach

- Not necessarily clear that "covert" semantic ambiguities should be so related to "overt" syntactic structure.

Underspecification approaches

Problem: syntactic commitment of QR approach

- Not necessarily clear that "covert" semantic ambiguities should be so related to "overt" syntactic structure.
- Psycholinguistic controversies (discussed later).

Underspecification approaches

Problem: syntactic commitment of QR approach

- Not necessarily clear that "covert" semantic ambiguities should be so related to "overt" syntactic structure.
- Psycholinguistic controversies (discussed later).

Potential solution: underspecification.

Underspecification approaches

Problem: syntactic commitment of QR approach

- Not necessarily clear that "covert" semantic ambiguities should be so related to "overt" syntactic structure.
- Psycholinguistic controversies (discussed later).

Potential solution: underspecification.

- Represent potential scope relationship without committing to direction until evidence shows up.

Underspecification approaches

Problem: syntactic commitment of QR approach

- Not necessarily clear that "covert" semantic ambiguities should be so related to "overt" syntactic structure.
- Psycholinguistic controversies (discussed later).

Potential solution: underspecification.

- Represent potential scope relationship without committing to direction until evidence shows up.

(3) a. Every child (x_1) climbed a tree (x_2).
b. Scope precedence: $x_1 = x_2$

Underspecification approaches

Problem: syntactic commitment of QR approach

- Not necessarily clear that "covert" semantic ambiguities should be so related to "overt" syntactic structure.
- Psycholinguistic controversies (discussed later).

Potential solution: underspecification.

- Represent potential scope relationship without committing to direction until evidence shows up.

(3) a. Every child (x_1) climbed a tree (x_2).
b. Scope precedence: $x_1 = x_2$

When evidence for precedence shows up, $x_1 > x_2$ or $x_1 < x_2$.

Partial underspecification

- Underspecification approach implemented used with multiple formalisms (e.g. Dynamic Syntax, Minimal Recursion Semantics).

Partial underspecification

- Underspecification approach implemented used with multiple formalisms (e.g. Dynamic Syntax, Minimal Recursion Semantics).
- Potential problems with full underspecification:

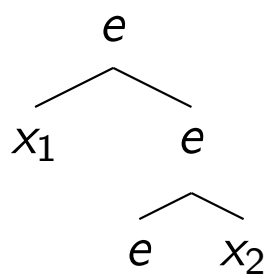
Partial underspecification

- Underspecification approach implemented used with multiple formalisms (e.g. Dynamic Syntax, Minimal Recursion Semantics).
- Potential problems with full underspecification:
  - "throws out the baby with the bathwater" – ignore some insights from syntactic theory.

Partial underspecification

- Underspecification approach implemented used with multiple formalisms (e.g. Dynamic Syntax, Minimal Recursion Semantics).
- Potential problems with full underspecification:
  - "throws out the baby with the bathwater" – ignore some insights from syntactic theory.
  - **Compromise solution**: variable scope trees (VST) [Sayeed, 2016] centered around event variables (e).

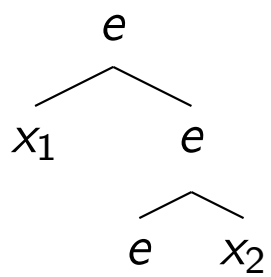
(4) a. Every child (x_1) climbed (e) a tree (x_2).
b. Scope precedence:



Partial underspecification

- Underspecification approach implemented used with multiple formalisms (e.g. Dynamic Syntax, Minimal Recursion Semantics).
- Potential problems with full underspecification:
  - "throws out the baby with the bathwater" – ignore some insights from syntactic theory.
  - **Compromise solution**: variable scope trees (VST) [Sayeed, 2016] centered around event variables (e).

(4) a. Every child (x_1) climbed (e) a tree (x_2).
b. Scope precedence:



  - Small number of constraints allow raising of x_2 without requiring involvement of full syntax.

The event can be deployed as a sort of "ceiling".

But what can we do with VSTs?

Structural ambiguity again

(5) A caregiver comforted a child every night.

- The linear order is $\exists = \exists > \forall$, ie, child and caregiver singular.

Structural ambiguity again

(5) A caregiver comforted a child every night.

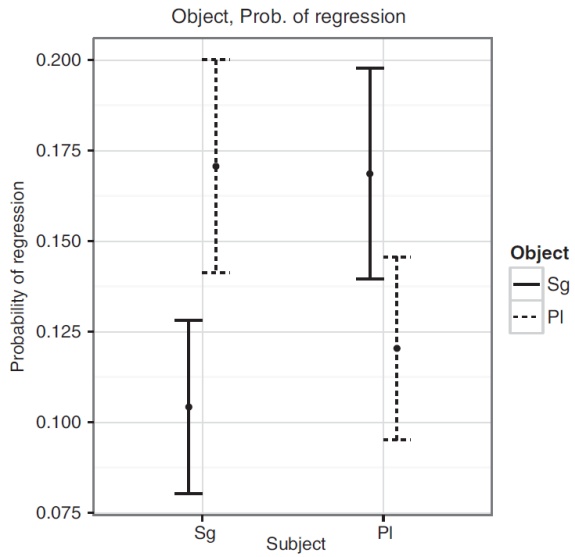
- The linear order is $\exists = \exists > \forall$, ie, child and caregiver singular.
- English speakers can be easily prompted to all other orders.

Structural ambiguity again

(5) A caregiver comforted a child every night.

- The linear order is $\exists = \exists > \forall$, ie, child and caregiver singular.
- English speakers can be easily prompted to all other orders.
- Dotlačil and Brasoveanu (2015):

The caregiver wanted the child to get some rest.
The caregivers wanted the child to get some rest.
The caregiver wanted the children to get some rest.
The caregivers wanted the children to get some rest.



Evidence for algorithmic processing (as opposed to purely pragmatic considerations).

The equivalent expression

"A caregiver comforted a child every night.":

(6) $\text{caregiver}(x^{\exists}) \bullet \text{comfort}(e^{\exists}) \bullet \text{agent}(x, e) \bullet \text{child}(y^{\exists}) \bullet \text{patient}(y, e) \bullet \text{night}(n^{\forall}) \bullet \text{OCCUR}(n, e)$

The equivalent expression

"A caregiver comforted a child every night.":

(6) $\text{caregiver}(x^{\exists}) \bullet \text{comfort}(e^{\exists}) \bullet \text{agent}(x, e) \bullet \text{child}(y^{\exists}) \bullet \text{patient}(y, e) \bullet \text{night}(n^{\forall}) \bullet \text{OCCUR}(n, e)$

- Can be obtained from incremental parse in a strictly rightward fashion (see Sayeed [2016]).

The equivalent expression

"A caregiver comforted a child every night.":

(6) $\text{caregiver}(x^{\exists}) \bullet \text{comfort}(e^{\exists}) \bullet \text{agent}(x, e) \bullet \text{child}(y^{\exists}) \bullet \text{patient}(y, e) \bullet \text{night}(n^{\forall}) \bullet \text{OCCUR}(n, e)$

- Can be obtained from incremental parse in a strictly rightward fashion (see Sayeed [2016]).
- Obtain VST by constructing a graph of arguments connected to the event.

The equivalent expression

"A caregiver comforted a child every night.":

(6) $\text{caregiver}(x^{\exists}) \bullet \text{comfort}(e^{\exists}) \bullet \text{agent}(x, e) \bullet \text{child}(y^{\exists}) \bullet \text{patient}(y, e) \bullet \text{night}(n^{\forall}) \bullet \text{OCCUR}(n, e)$

- Can be obtained from incremental parse in a strictly rightward fashion (see Sayeed [2016]).
- Obtain VST by constructing a graph of arguments connected to the event.

VST starting point

"A caregiver comforted a child every night.":

(7) $\text{caregiver}(x^{\exists}) \bullet \text{comfort}(e^{\exists}) \bullet \text{agent}(x, e) \bullet \text{child}(y^{\exists}) \bullet \text{patient}(y, e) \bullet \text{night}(n^{\forall}) \bullet \text{OCCUR}(n, e)$

VST starting point

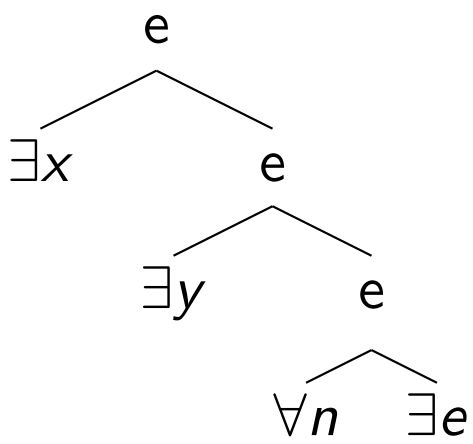
"A caregiver comforted a child every night.":

(7) caregiver($x^{\exists}$) • comfort($e^{\exists}$) • agent(x, e) • child($y^{\exists}$) • patient(y, e) • night($n^{\forall}$) • OCCUR(n, e)

The linear order:

(8) a. The caregiver wanted the child to get some rest. ($\exists x \exists y > \forall n$)

b.

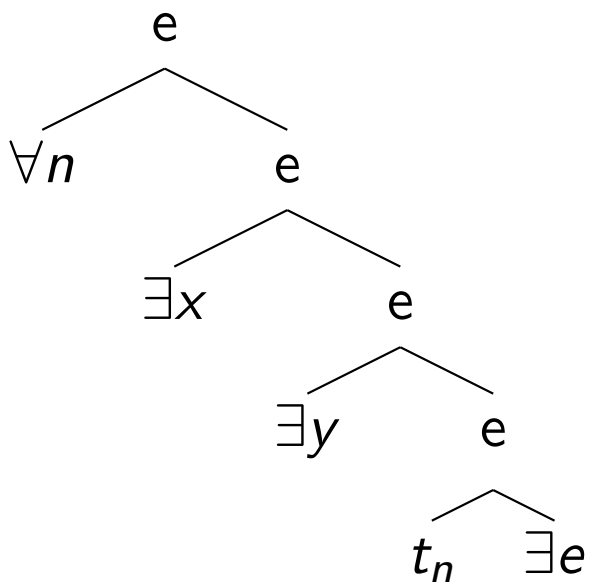


VST-move

"A caregiver comforted a child every night.":
One step:

(9) a. The caregivers wanted the children to get some rest. $(\forall n > \exists x \exists y)$

b.

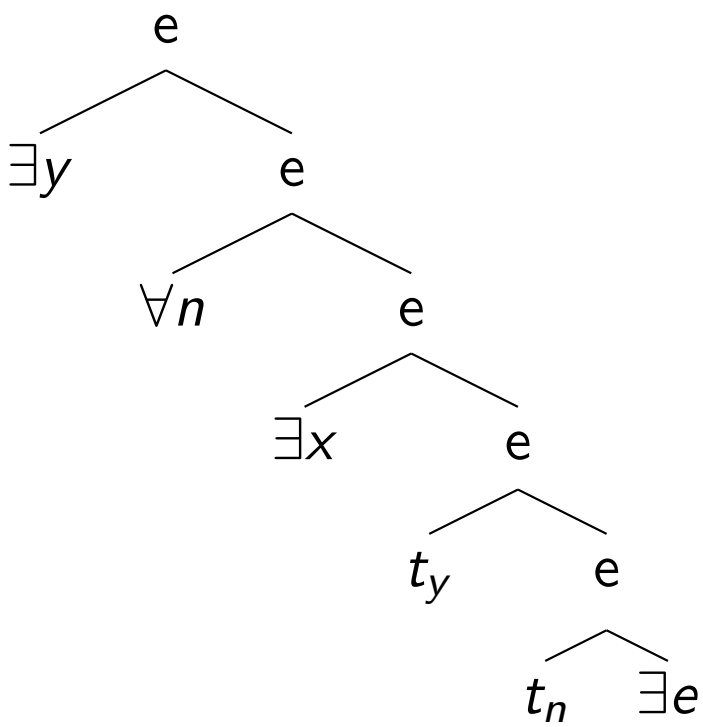


VST-move

"A caregiver comforted a child every night.":
Two steps for number-mismatched reading:

(10) a. The caregivers wanted the child to get some rest. $(\forall n > \exists x)$

b.

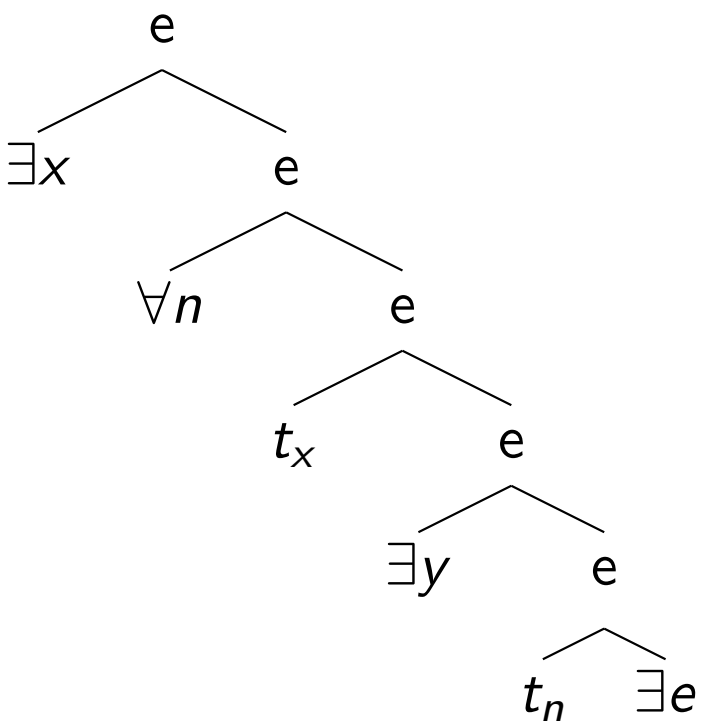


VST-move

"A caregiver comforted a child every night.":
Two steps for alternate number-mismatched reading:

(11) a. The caregiver wanted the children to get some rest. ($\forall n > \exists y$)

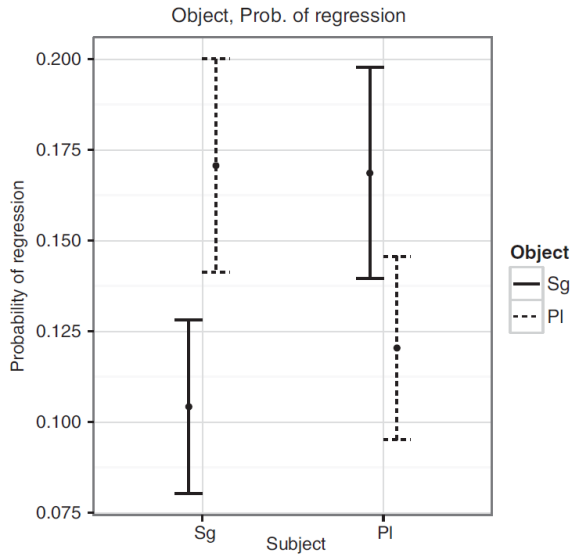
b.



VST-move matches difficulty

The number of VST-move steps matches the difficulty of each Dotlačil and Brasoveanu reading:

The caregiver wanted the child to get some rest.
The caregivers wanted the child to get some rest.
The caregiver wanted the children to get some rest.
The caregivers wanted the children to get some rest.



Pros and. . . cons?

Advantages:

- Rightward/incrementally constructible, only does work when scope ambiguities call for it (rather than constraining the entire processing machinery).
- VST-move is highly constrained, possible psycholinguistic response to Fox and Lappin's critique [2010] of NP-completeness in fully underspecified scope processing mechanisms.

Pros and. . . cons?

Advantages:

- Rightward/incrementally constructible, only does work when scope ambiguities call for it (rather than constraining the entire processing machinery).
- VST-move is highly constrained, possible psycholinguistic response to Fox and Lappin's critique [2010] of NP-completeness in fully underspecified scope processing mechanisms.

Critique:

- Why bother with the neo-Davidsonian event variable/misuse it as a structural element?
- Can represent the same thing in e.g. CCG?
  (But. . . underspecification. . .)

Part 3: event variable speculations

The event as structural ceiling

First role of the event variable: prevent infinite movement.
Radó and Bott [2012]: self-paced reading with picture task in German.

(12) Genau ein Affe ist auf allen/jeder Karte(n) zu finden.
Exactly one monkey is on all/each card(s) to find.

The event as structural ceiling

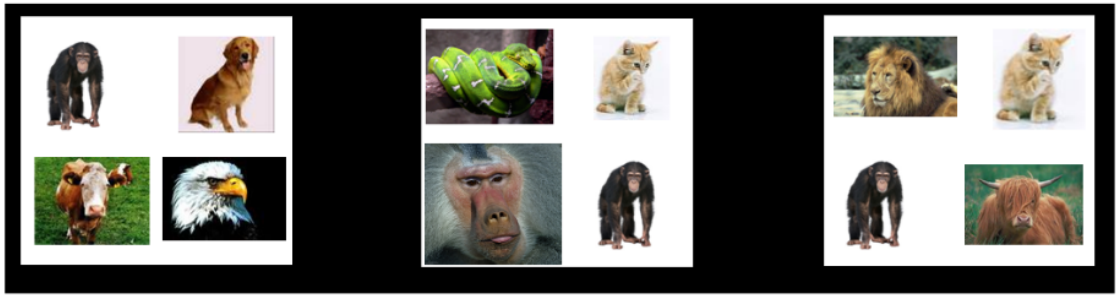
First role of the event variable: prevent infinite movement.
Radó and Bott [2012]: self-paced reading with picture task in German.

(12) Genau ein Affe ist auf allen/jeder Karte(n) zu finden.
Exactly one monkey is on all/each card(s) to find.

- Visual task afterwards: set of cards with pictures of monkeys on it.
  $\Rightarrow$ Participant decides if the statement was true for the cards.
- Long story short: by comparison against control scope-unambiguous sentences, participants slow down during card task to remember the scope order.
- Suggests creation of minimal domain of scope interpretation.

The event as structural ceiling

Radó and Bott [2012]:

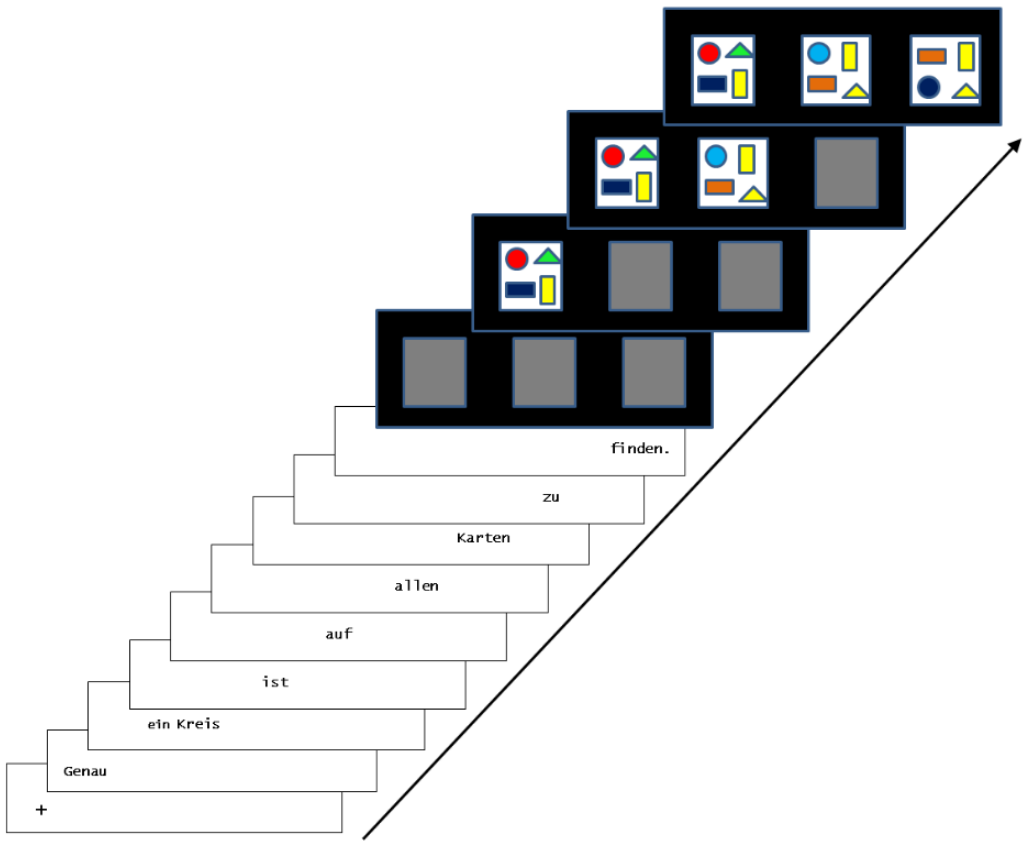


(a) ∃!∀ card display

(b) ∀∃! card display

The event as structural ceiling

Radó and Bott [2012]:

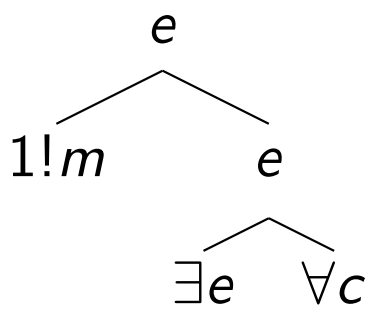


The event as structural ceiling

(13) Genau ein Affe$_m$ ist auf allen/jeder Karte(n)$_c$ zu finden.
Exactly one monkey is on all/each card(s) to find.

As VST in linear order:

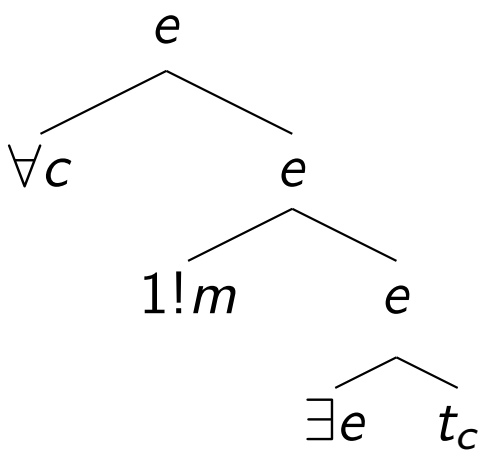
(14)



The event as structural ceiling

However, the pragmatics suggest the inverse reading which the processor immediately assumes:

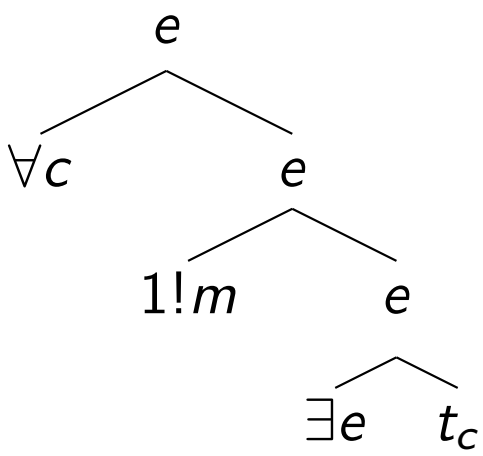
(15)



The event as structural ceiling

However, the pragmatics suggest the inverse reading which the processor immediately assumes:

(15)

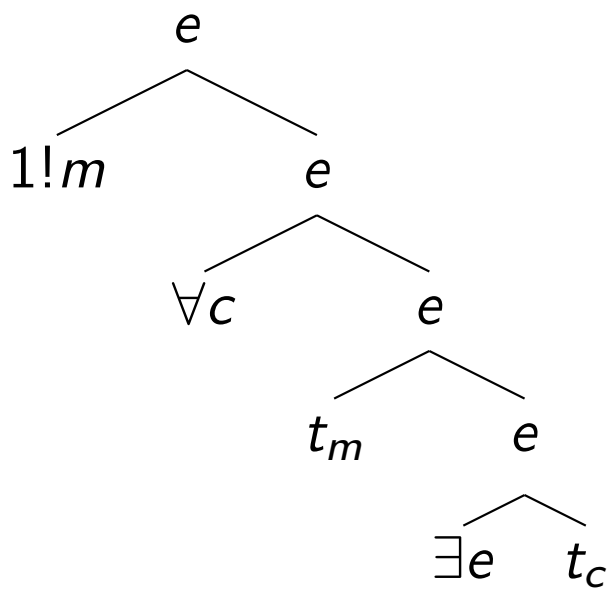


(The "trace" records that this is not the linear state, formally your mileage may vary.)

The event as structural ceiling

However, but the visual task can force the linear reading, which is a step harder:

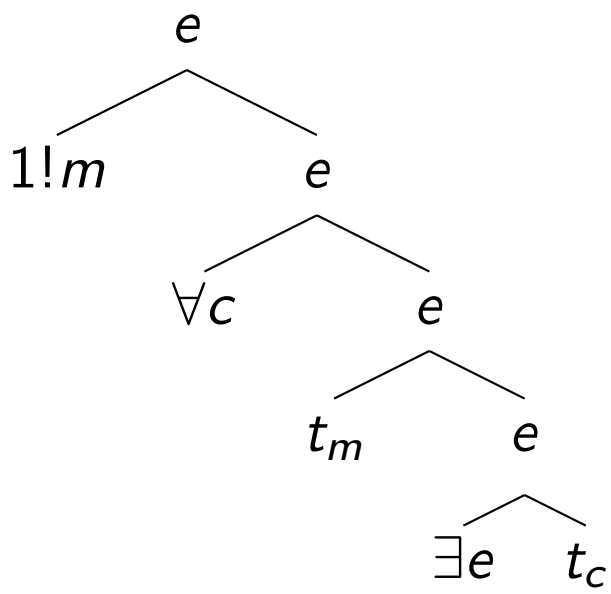
(16)



The event as structural ceiling

However, but the visual task can force the linear reading, which is a step harder:

(16)



This is strongly disfavoured by the parser, because it contains an image of the initial state, since the "traces" contain no semantic content.

Part 4: beyond the clause

Antecedent-contained deletion (ACD)

Consider:

(17) Bob demands that he ride every car that Bill did.

Antecedent-contained deletion (ACD)

Consider:

(17) Bob demands that he ride every car that Bill did.

- Lower scope: Bill rode every car that Bob demands to ride.
- Higher scope: Bill demanded to ride every car that Bob demands to ride. *(rejected by most adults)*

Antecedent-contained deletion (ACD)

Consider:

(17) Bob demands that he ride every car that Bill did.

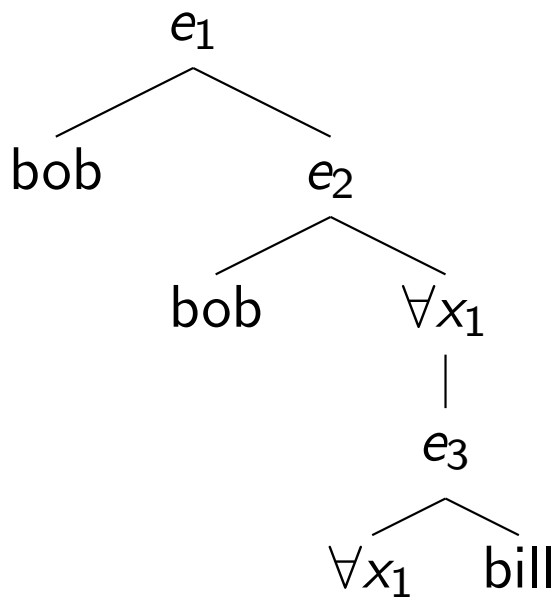
- Lower scope: Bill rode every car that Bob demands to ride.
- Higher scope: Bill demanded to ride every car that Bob demands to ride. *(rejected by most adults)*

$\Rightarrow$ "Tensed clause barrier" – higher scope cannot be obtained

Antecedent-contained deletion (ACD)

(18) Bob demands$_1$ that he ride$_2$ every car that Bill did$_3$.

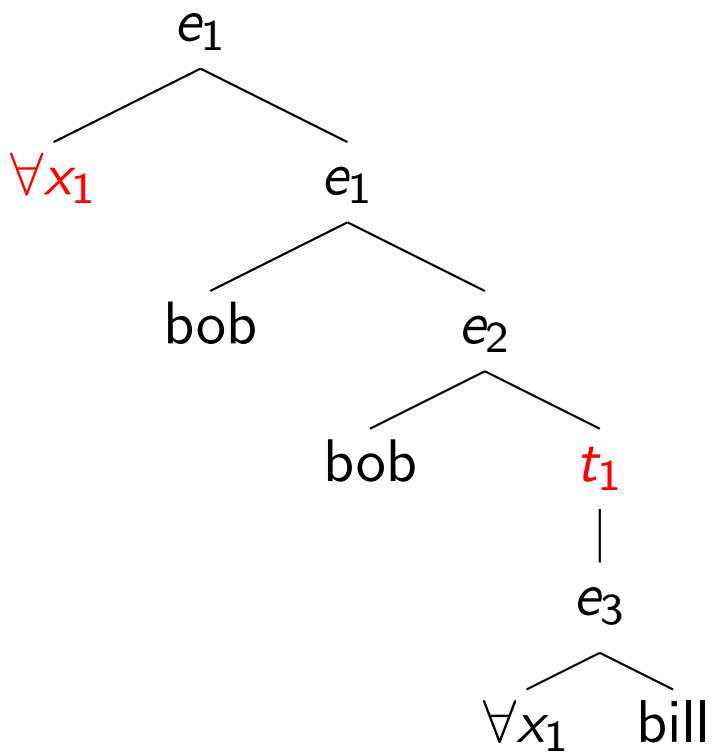
Lower scope reading (correct):



Antecedent-contained deletion (ACD)

(19) Bob demands$_1$ that he ride$_2$ every car that Bill did$_3$.

Higher scope reading:

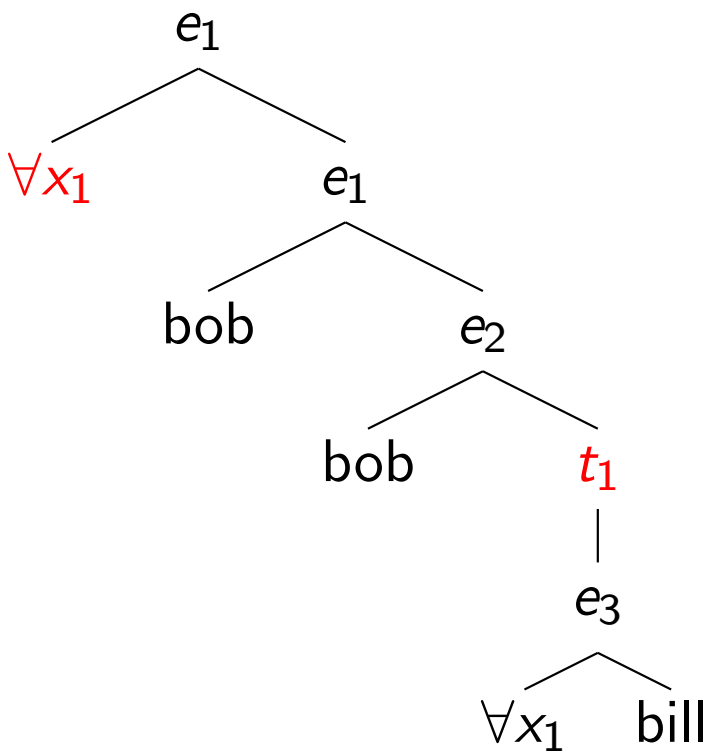


(left out $\exists$es for space)

Antecedent-contained deletion (ACD)

(19) Bob demands$_1$ that he ride$_2$ every car that Bill did$_3$.

Higher scope reading:



(left out $\exists$es for space)Normally banned: drastic change to e_2 if "every car" is part of "demand" event.

Antecedent-contained deletion (ACD)

- Syrett and Lidz (2010): children (and many adults) do not respect tensed clause barrier.
- Why do the majority of adults find tensed clause blocks higher reading? Syrett and Lidz suggest processing constraints.

Antecedent-contained deletion (ACD)

- Syrett and Lidz (2010): children (and many adults) do not respect tensed clause barrier.
- Why do the majority of adults find tensed clause blocks higher reading? Syrett and Lidz suggest processing constraints.

How could this be allowed?

Antecedent-contained deletion (ACD)

- Syrett and Lidz (2010): children (and many adults) do not respect tensed clause barrier.
- Why do the majority of adults find tensed clause blocks higher reading? Syrett and Lidz suggest processing constraints.

How could this be allowed?
Hypothesis: Memory management of propositions (events) less developed in children. $e_1 = e_2$, so VST-move restriction **circumvented**.

Antecedent-contained deletion (ACD)

- Syrett and Lidz (2010): children (and many adults) do not respect tensed clause barrier.
- Why do the majority of adults find tensed clause blocks higher reading? Syrett and Lidz suggest processing constraints.

How could this be allowed?
Hypothesis: Memory management of propositions (events) less developed in children. $e_1 = e_2$, so VST-move restriction **circumvented**.
Experimental evidence: Role of working memory in propositional content, Caplan and Waters (1999). Role of memory decay in sentence processing, Lewis et al. (2006).

Each and every

Events and quantifiers

"Each" and "every" are $\forall$ but still not created the same.
Beghelli and Stowell [1996].

(20) a. It took all the boys to lift the piano.
b. It took every boy to lift the piano.
c.*It took each boy to lift the piano.
d. One boy ate almost every apple.
e.*One boy ate almost each apple.

Events and quantifiers

"Each" and "every" are $\forall$ but still not created the same.
Beghelli and Stowell [1996].

(20) a. It took all the boys to lift the piano.
b. It took every boy to lift the piano.
c.*It took each boy to lift the piano.
d. One boy ate almost every apple.
e.*One boy ate almost each apple.

It's as though "each" is quantifying events of eating/lifting, which are infelicitous here.

“Each” and “together”

Patson and Warren (2010): self-paced reading/timed judgement study.

(21) a. Each of the men carried a box/some boxes.
b. Together the men carried a box/some boxes.

The distributed “each” reading has slower judgement time in the singular condition than the collective “together” reading.

“Each” and “together”

Patson and Warren (2010): self-paced reading/timed judgement study.

(21) a. Each of the men carried a box/some boxes.
b. Together the men carried a box/some boxes.

The distributed “each” reading has slower judgement time in the singular condition than the collective “together” reading.
$\Rightarrow$ distinction is distribution over events.

Back to "every"

Dwivedi and Gibson [2017]:

- Attempt to replicate Patson and Warren, but with different stimuli and in an ERP setting.
- Conclude (similar to Dwivedi [2013]) that computation of distribution dominated by heuristic/pragmatic factors.

Back to "every"

Dwivedi and Gibson [2017]:

- Attempt to replicate Patson and Warren, but with different stimuli and in an ERP setting.
- Conclude (similar to Dwivedi [2013]) that computation of distribution dominated by heuristic/pragmatic factors.

Stimuli

(22) a. Every kid climbed a/the tree(s).
b. The kid climbed a/the tree(s).

Back to "every"

Dwivedi and Gibson's result may simply be due to the distributivity of the event.

- Patson and Warren constrast ("together" vs. "each") involve an ambiguous scope over the event variable in the "each" case.
- Dwivedi and Gibson's result comes from computation strictly over entity variables, so replicates Dwivedi [2013].

Some last discussion questions

No matter what, some portion of the representation must remain *under-specified*:

Some last discussion questions

No matter what, some portion of the representation must remain *under-specified*:

1. How can we most easily acquire the sources of world knowledge needed to represent processor decisions?

Some last discussion questions

No matter what, some portion of the representation must remain *under-specified*:

1. How can we most easily acquire the sources of world knowledge needed to represent processor decisions?
2. Is there a role for *generalized event knowledge* and thematic role representation?

Some last discussion questions

No matter what, some portion of the representation must remain *under-specified*:

1. How can we most easily acquire the sources of world knowledge needed to represent processor decisions?
2. Is there a role for *generalized event knowledge* and thematic role representation?
3. Theoretical syntax has spent a lot of effort in identifying constraints on "covert" phenomena – how many of these can be accounted for purely information-theoretically?

Thanks and enjoy ESSLLI!
http://bit.ly/esslli19scope