Introduction to Sentence Comprehension

Shravan Vasishth

Universität Potsdam, Germany
vasishth@uni-potsdam.de

August 13, 2017

What is sentence processing

Two central goals in this field are to understand

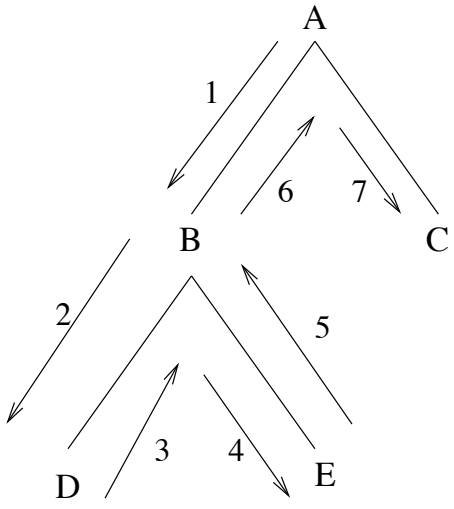
- online **parsing mechanisms** in human sentence comprehension
  - left-corner parsing, top-down, bottom-up? lookahead?
  - probabilistic parsing?
  - serial vs parallel vs ranked parallel?
  - deterministic vs non-deterministic parsing?
  - what kind of information is used to make parsing decisions (syntactic only, syntactic+semantic+...?)
- constraints on **dependency completion**
  - a general preference to attach co-dependents locally
  - the consequences of probabilistic predictive parsing (expectation effects)
  - "good-enough" processing, underspecification, tracking only local n-grams ("local coherence")
  - constraints on retrieval processes

Introduction

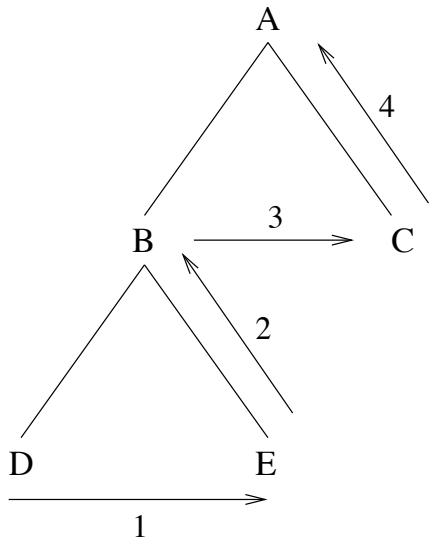
- In this course, we will give a fairly narrow perspective on processing sentences out of context.
- Please consult the references cited in these slides for a fuller picture.

Introduction

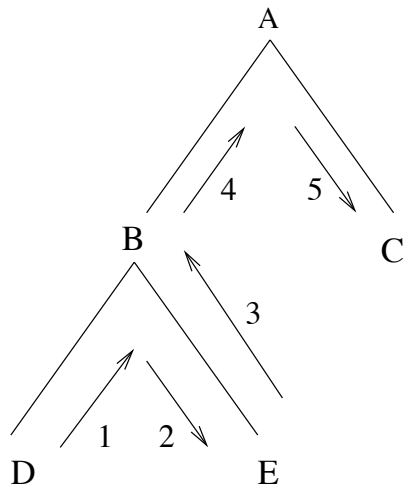
Left-corner parsing, probabilistic parsing



Top–down



Bottom–up



Left–corner

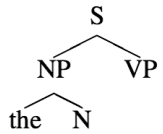
Introduction: parsing mechanisms

Left-corner parsing [1], probabilistic parsing

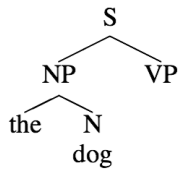
Left-Corner Parsing

$S \rightarrow NP\ VP$ $\quad$ $Det \rightarrow a, the$ $\quad$ $NP \rightarrow Det\ N$
$N \rightarrow man, dog$ $\quad$ $V \rightarrow ran, saw$ $\quad$ $VP \rightarrow V$
$VP \rightarrow V\ NP$

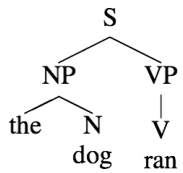
INPUT: *the*
GOAL CATEGORY STACK: [S]
ACTIONS: If *the* is the left corner of any phrase structure rule then replace the stack content with the LHS of that rule. Repeat this left-corner rule until no further steps are possible. Wait for next input word. These actions yield the structure to the right:



INPUT: *dog*
GOAL CATEGORY STACK: [N NP VP S]
ACTIONS: Use the left-corner rule to expand *dog* to N. Since N is predicted in the incremental structure built so far (Step 1), integrate the N built up bottom-up into the tree. Since no further applications of the left-corner rule are possible, wait for the next input.



INPUT: *ran*
GOAL CATEGORY STACK: [VP S]
ACTIONS: Use the left-corner rule to expand *ran* to V, and apply this rule once again to expand to VP. Since a VP is predicted in the structure, integrate this with the tree.



Introduction

Left-corner parsing, probabilistic parsing

Purely top-down or purely bottom-up strategies turn out to be inappropriate models for human parsing [2, 3, 4] since they are unable to capture the observation [5, 468-470] that left-branching and right-branching structures are relatively easy to process compared to center embeddings:

(1) a. Bill's book's cover is dirty.

b. Bill has the book that has the cover that is dirty.

c. The rat the cat the dog chased killed ate the malt.

Introduction

Left-corner parsing, probabilistic parsing

More frequent attachments are preferred over rare attachments [6].

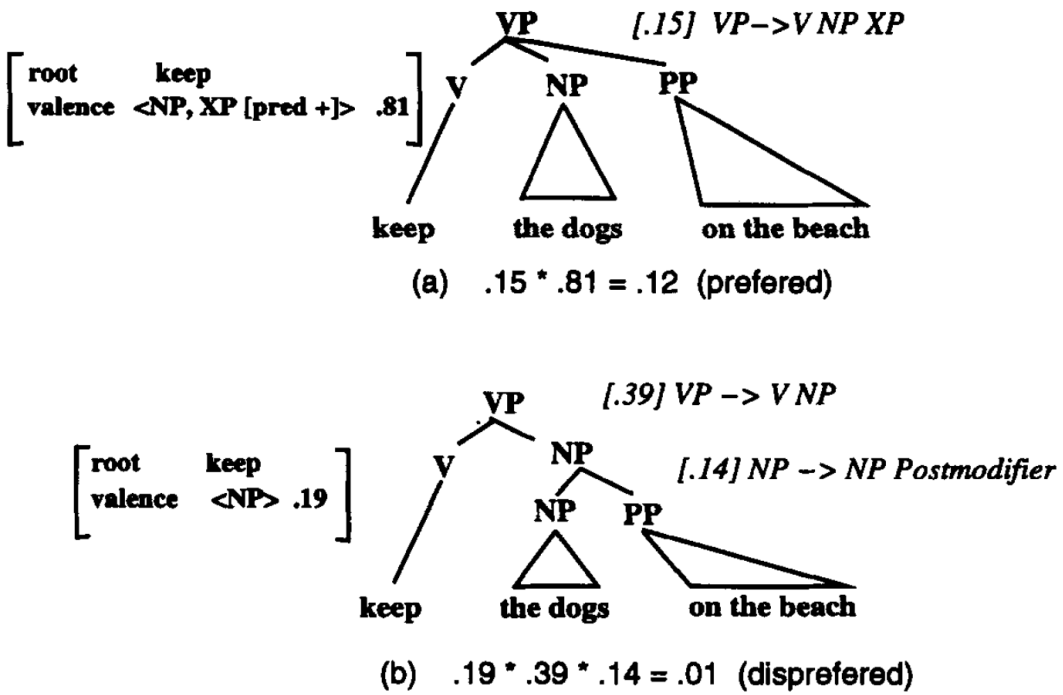
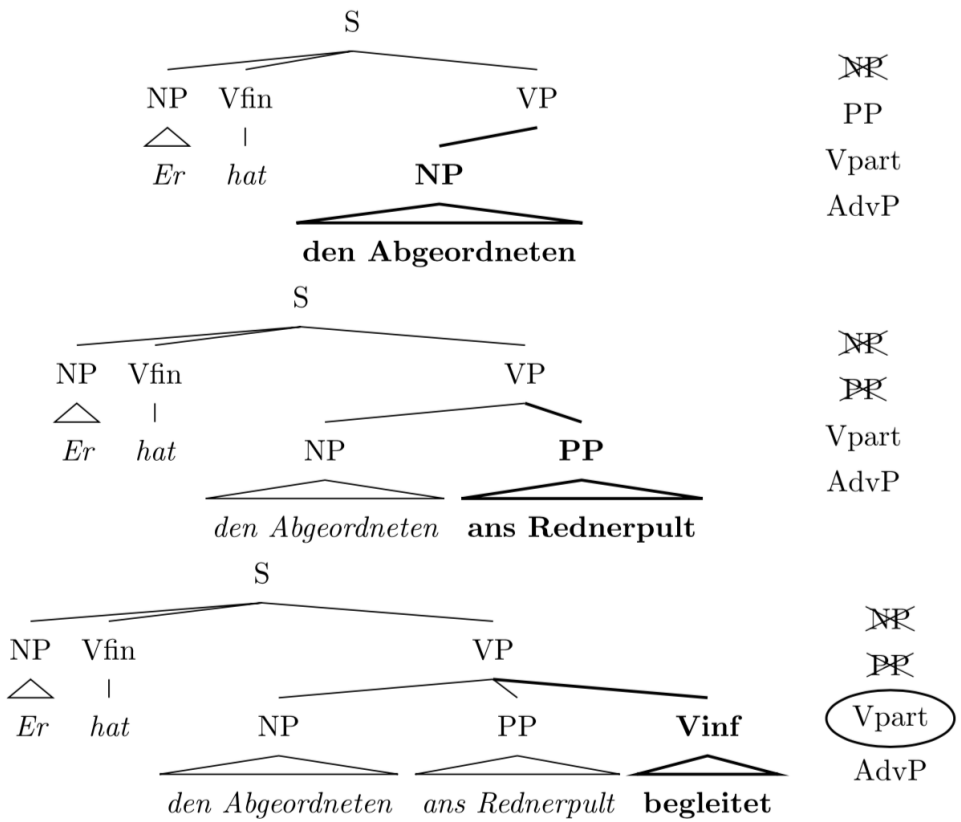


Figure 8. Annotated Parse Trees for Two Interpretations of *keep the dogs on the beach*

Introduction

Left-corner parsing, probabilistic parsing

Expectations for an upcoming verb phrase are sharpened if the verb's appearance is delayed [7].



Introduction: parsing mechanisms

serial / parallel / ranked parallel

A general assumption in most work today is that parse choices are strictly serial. But theoretically, other options are possible, and there is some evidence for ranked parallelism [8].

- Serial: compute a single analysis, and if that fails, backtrack and compute new analysis (most classical theories, e.g., [9, 10, 11]).
- Parallel:
  - Ranked: Compute all analyses in parallel, but rank them (e.g. by likelihood).
  - Prune: using, e.g., beam search.
  - Don't prune at all—generate all possible structures and then compute a function over them (e.g. entropy reduction, or surprisal) to find the optimal one [12, 13].

Introduction

deterministic / non-deterministic

- A common early assumption was that parsing was essentially deterministic.
- A heuristic is to always prefer to attach locally [11]. Example:

(2) a. (low attachment)
The car of the **driver** *that had the moustache* was pretty cool.

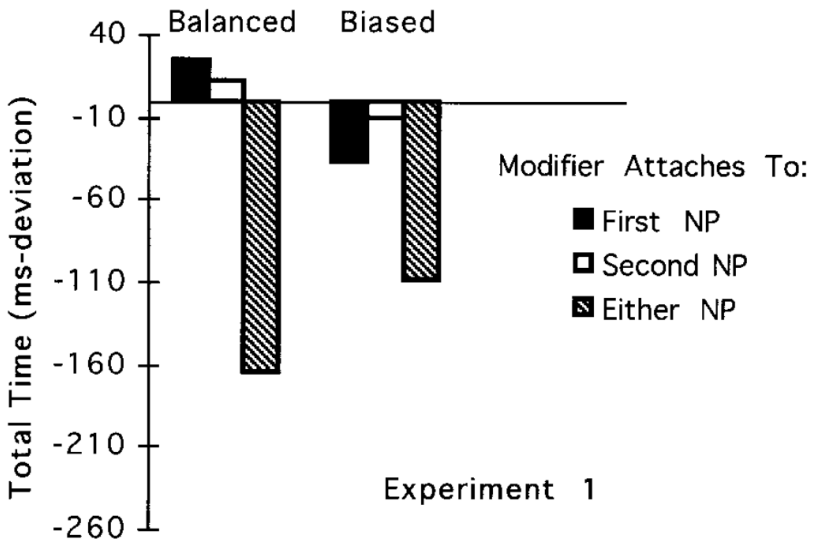
b. (high attachment)
The **driver** of the car *that had the moustache* was pretty cool.

c. (globally ambiguous)
The **son** of the **driver** *that had the moustache* was pretty cool.

- Prediction: 2a,c easier to process than 2b.

Introduction

deterministic / non-deterministic

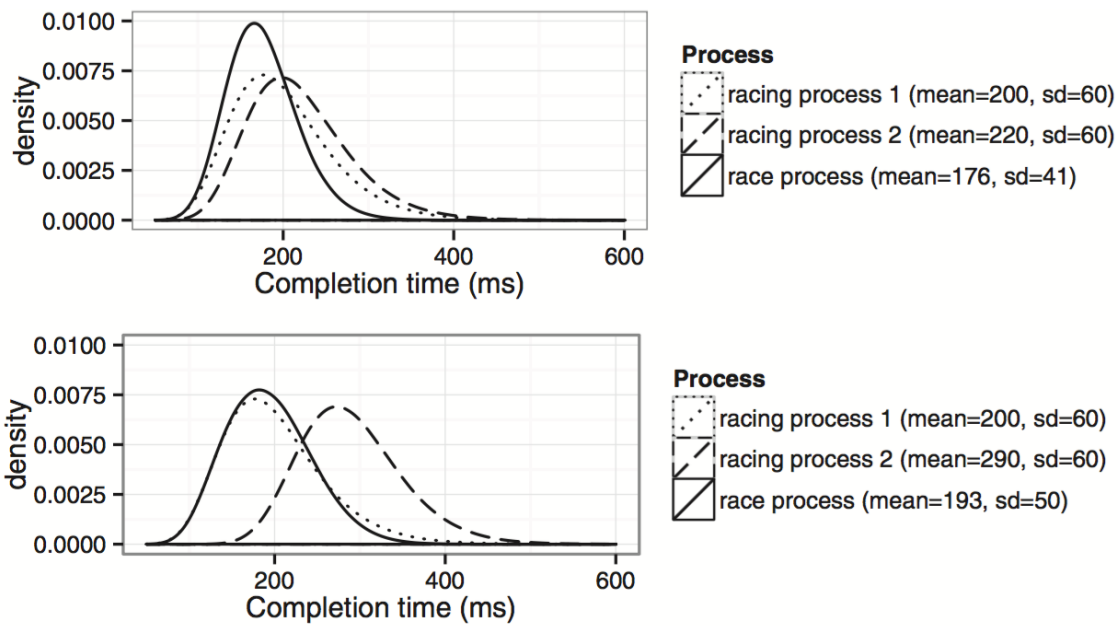


[14] found/claimed that the word *moustache* was read fastest in the globally ambiguous sentence: the **ambiguity advantage**.

Introduction

deterministic / non-deterministic

One explanation [15] for this is to assume a non-deterministic race process (also see [16]):



Introduction: parsing mechanisms

information sources: syntax only / all sources of information

[17] found evidence against syntax-first proposals, but [18] found evidence for syntax-first. (*A too-common example of how prior beliefs of researchers are, uncannily, always magically confirmed*.)

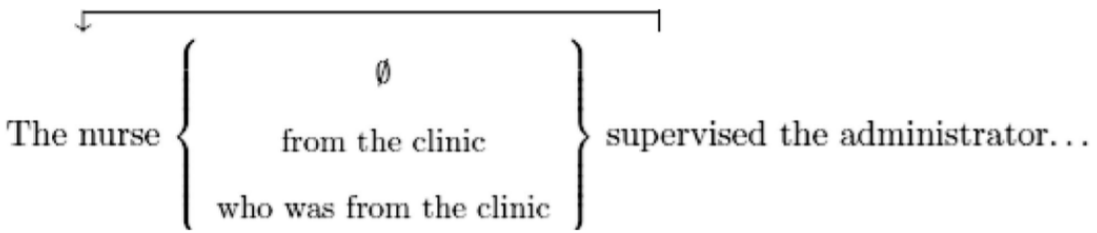
(3) a. The defendant examined by the lawyer turned out to be unreliable.

b. The evidence examined by the lawyer turned out to be unreliable.

Introduction: constraints on dependency completion

A local attachment preference

Non-local dependency completion tends to be more difficult than local dependency completion [19, 20].



Introduction: constraints on dependency completion

A local attachment preference

(4) a. The administrator who the **nurse supervised** scolded the medic while . . .

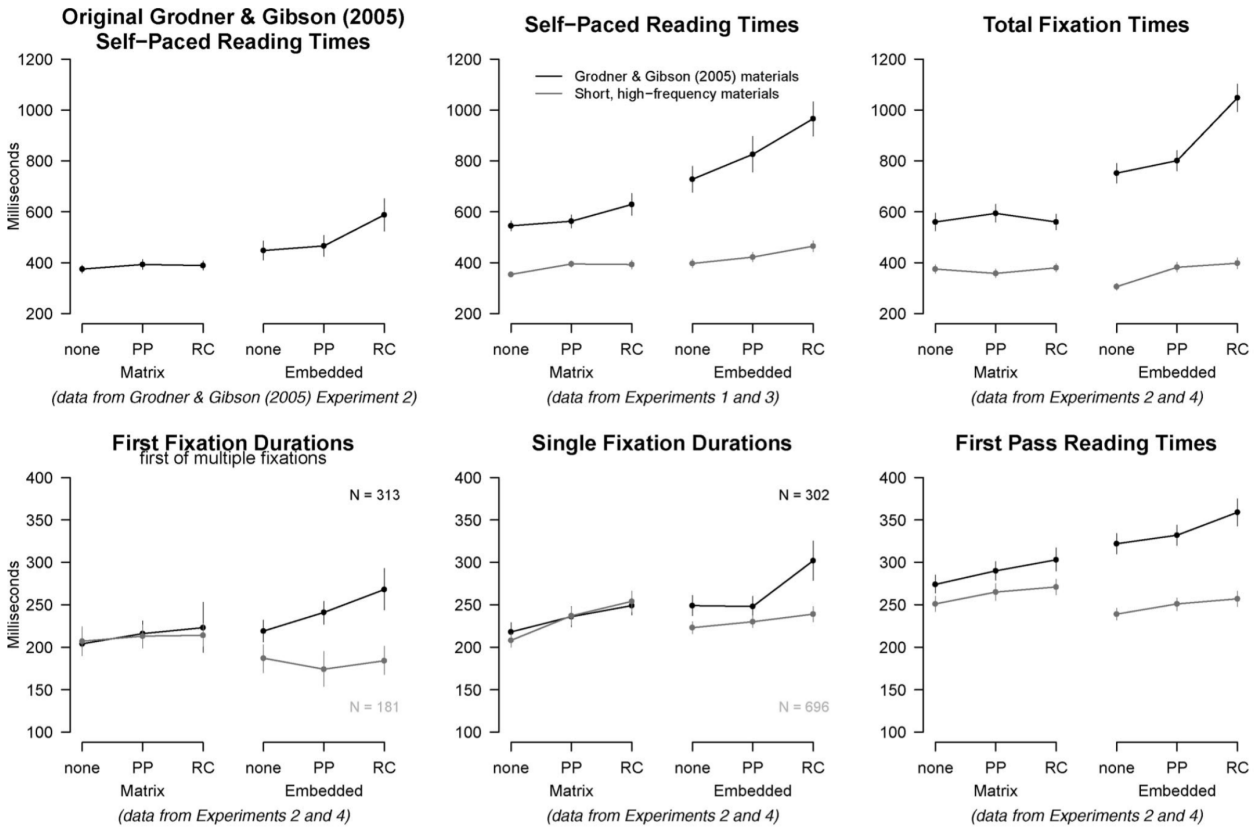
b. The administrator who the **nurse** from the clinic **supervised** scolded the medic while . . .

c. The administrator who the **nurse** who was from the clinic **supervised** scolded the medic while . . .

Introduction: constraints on dependency completion

A local attachment preference

Source: [20].



Introduction: constraints on dependency completion

Good-Enough processing / underspecification / local coherence

Source: [21]

(5) a. The coach smiled at the player who was tossed a frisbee

b. The coach smiled at the player ~~who was~~ tossed a frisbee

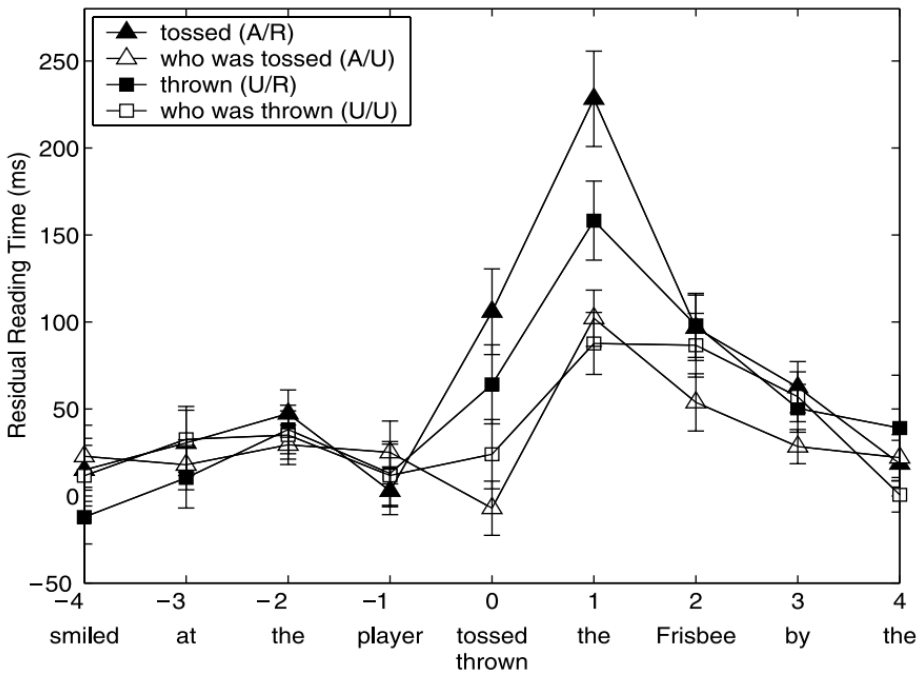
Subjects seem to treat

(6) "the player tossed a frisbee"

as a main clause.

Introduction: constraints on dependency completion

Good-Enough processing / underspecification / local coherence



Dependency completion cost and predictive processing

Safavi, Husain, Vasishth 2016

Expt 1 (SPR), 3 (ET):

(7)
a. *CP-Short*: . . . noun-host PP CP verb . . .
b. *CP-Long*: . . . noun-host RC+PP CP verb . . .
c. *SP-Short*: . . . noun-objPP Simple verb . . .
d. *SP-Long*: . . . noun.obj RC+PP Simple verb . . .

Expt 2 (SPR), 4 (ET):

(8)
a. *CP-Short*: . . . noun-host Short PP CP verb . . .
b. *CP-Long*: . . . noun-host Long PP CP verb . . .
c. *SP-Short*: . . . noun-objShort PP Simple verb . . .
d. *SP-Long*: . . . noun.obj Long PP Simple verb . . .

Dependency completion cost and predictive processing

Safavi, Husain, Vasishth 2016

(9) a. Strong predictability, short distance (PP)

Ali a:rezouyee bara:ye man kard va...
Ali wish-INDEF for 1.S do-PST and...

'Ali made a wish for me and...'

b. Strong predictability, long distance (RC+PP)

Ali a:rezouyee ke besya:r doost-da:sht-am
Ali wish-INDEF that a lot like-1.S-PST
bara:ye man kard va...
for 1.S do-PST and...

'Ali made a wish that I liked a lot for me and...'

Dependency completion cost and predictive processing

Safavi, Husain, Vasishth 2016

(10) a. Weak predictability, short distance (PP)

Ali shokola:ti bara:ye man xarid va
Ali chocolate-INDEF for 1.S buy-PST and...
...

'Ali bought a chocolate for me and ...'

b. Weak predictability, long distance (RC+PP)

Ali shokola:ti ke besya:r doost-da:sht-am
Ali chocolate-INDEF that a lot like-1.S-PST
bara:ye man xarid va...
for 1.S buy-PST and...

'Ali bought a chocolate that I liked a lot for me and....'

Dependency completion cost and predictive processing

Safavi, Husain, Vasishth 2016

- We also computed entropy using sentence completion data, and investigated whether entropy can explain the intervention effects.
- Entropy is an information-theoretic measure that essentially represents how uncertain we are of an outcome (Shannon 2001).
- In the present case, this would translate to our uncertainty about the upcoming verb.
- If there are n possible ways to continue a sentence, and each of the possible ways has probability p_i, where $i = 1, \ldots, n$, then entropy is defined as $-\sum_i p_i \times \log_2(p_i)$.

Dependency completion cost and predictive processing

Safavi, Husain, Vasishth 2016

Example:

- Two verbs predicted with equal probability:
  Entropy= $-[0.5\log_2(0.5) + 0.5\log_2(0.5)] = 1$
- Three verbs predicted with equal probability:
  Entropy=
  $-[0.33\log_2(0.33) + 0.33\log_2(0.33) + 0.33\log_2(0.33)] = 1.6$

Introduction: constraints on dependency completion

Uncertainty increases with argument-verb distance (Safavi et al 2016)

(4) a. Strong predictability, short distance (PP)
Ali a:rezouyee bara:ye man kard va…
Ali wish-INDEF for 1.S do-PST and…
'Ali made a wish for me and…'

b. Strong predictability, long distance (RC+PP)
Ali a:rezouyee ke besya:r doost-da:sht-am
Ali wish-INDEF that a lot like-1.S-PST
bara:ye man kard va…
for 1.S do-PST and…
'Ali made a wish that I liked a lot for me and…'

c. Weak predictability, short distance (PP)
Ali shokola:ti bara:ye man xarid va…
Ali chocolate-INDEF for 1.S buy-PST and…
'Ali bought a chocolate for me and …'

d. Weak predictability, long distance (RC+PP)
Ali shokola:ti ke besya:r doost-da:sht-am
Ali chocolate-INDEF that a lot like-1.S-PST
bara:ye man xarid va…
for 1.S buy-PST and…
'Ali bought a chocolate that I liked a lot for me and…'

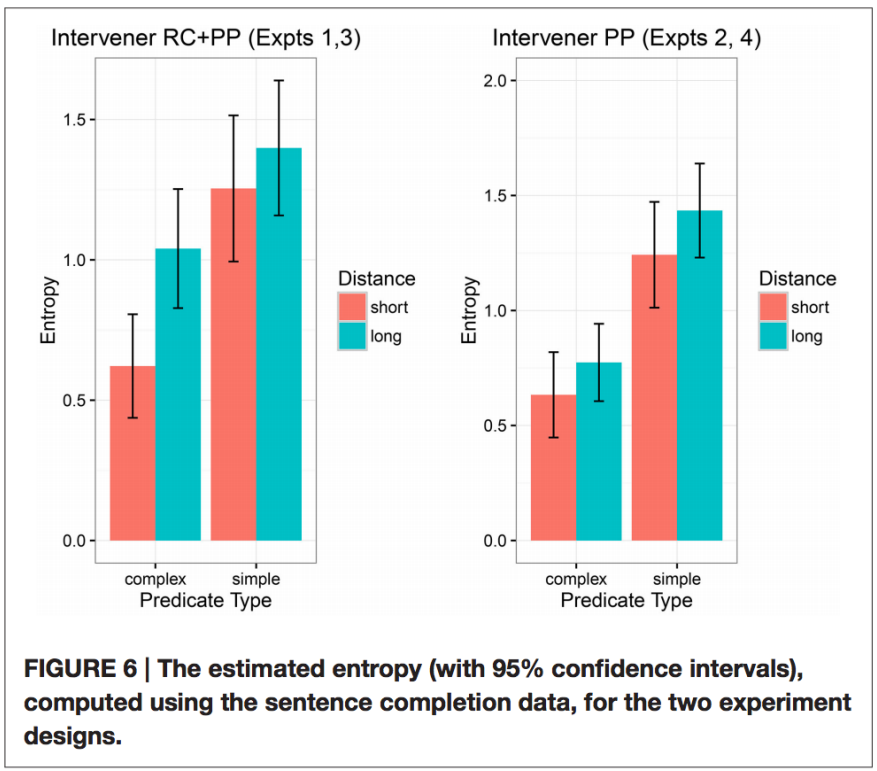


FIGURE 6 | The estimated entropy (with 95% confidence intervals), computed using the sentence completion data, for the two experiment designs.

Inhibition and facilitation due to predictive processing

Safavi, Husain, Vasishth 2016

Why does entropy increase in longer-distance dependencies?

- A possible explanation suggests itself in terms of memory overload.
- In the long-distance complex predicate condition, participants might have forgotten that a noun-verb dependency exists.
- Prediction: participants would tend to produce more ungrammatical continuations in the long-distance condition than the short-distance condition.
- This is some evidence for this in experiment 1: the accuracy in the short condition was 0.97 and 0.92 (log odds -.29 [-0.73, 0.09]).

The idea of forgetting-causing-entropy is worth investigating in the old locality studies involving English, German, Hindi.

Introduction: constraints on dependency completion

Uncertainty increases with argument-verb distance (Safavi et al 2016)

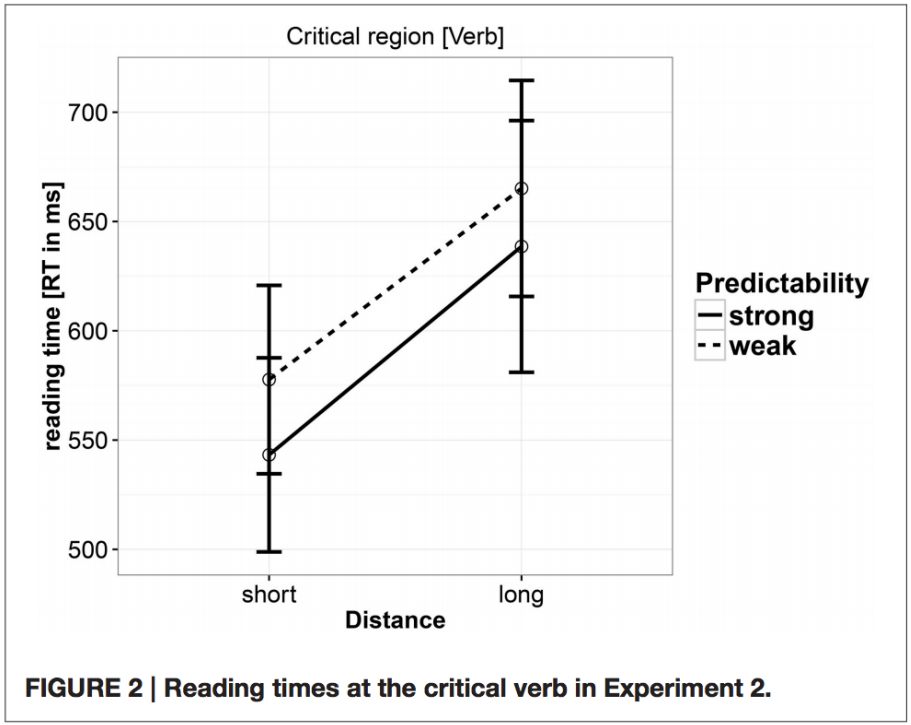
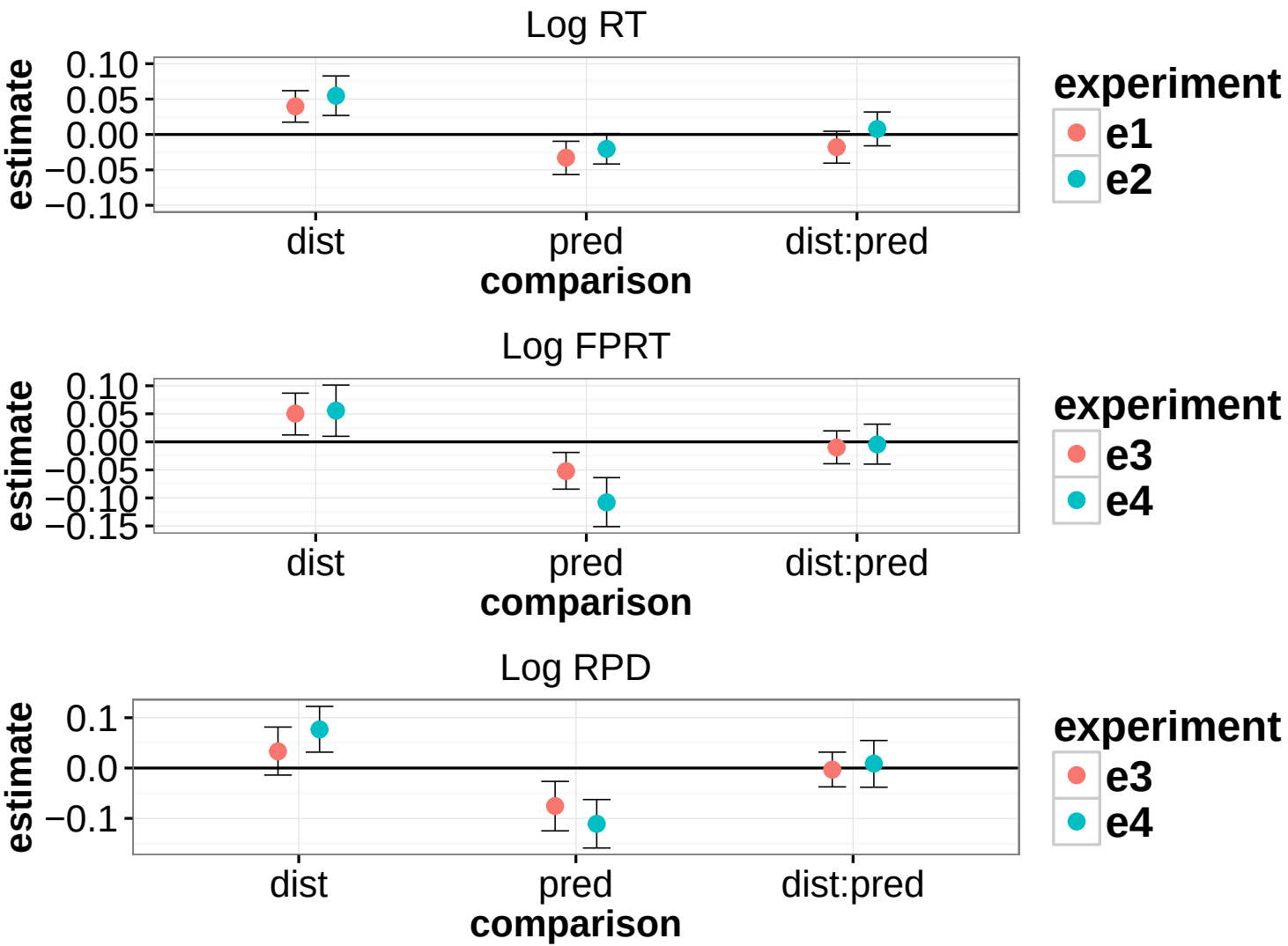


FIGURE 2 | Reading times at the critical verb in Experiment 2.

Introduction: constraints on dependency completion

Uncertainty increases with argument-verb distance (Safavi et al 2016)



Introduction: constraints on dependency completion

Constraints on retrieval

Similarity-based interference has been implicated as a cause for difficulty in completing subject-verb dependencies.
The essential idea is that retrieving an item (e.g., a noun) is harder (e.g., at a verb) if there are other competing items present that are similar on some dimension.

An implementation of this idea is Lewis and Vasishth (2005) (henceforth LV05).

The model assumptions

This is often called "the" cue based model, but there are many cue-based models (Van Dyke's, McElree's conceptions are different from the LV05 model).

1. **Grammatical knowledge and left-corner parsing algorithm**:

   Note that a parser can do nothing without a grammar. So even asking a question like "is it the grammar or the parser?" technically doesn't even mean anything.

   - If-then production rules drive structure building
   - Rules are hand-crafted in toy models, but scaling up has been done (Boston, Hale, Kliegl, Vasishth, Lang Cog Proc 2011).

2. **Constraints on memory processes affecting retrieval**:

   allows us to model individual differences in attention and working memory capacity

Retrieval at any dependency completion point is a key (but not only) determinant of processing difficulty or facilitation.

Introduction and background

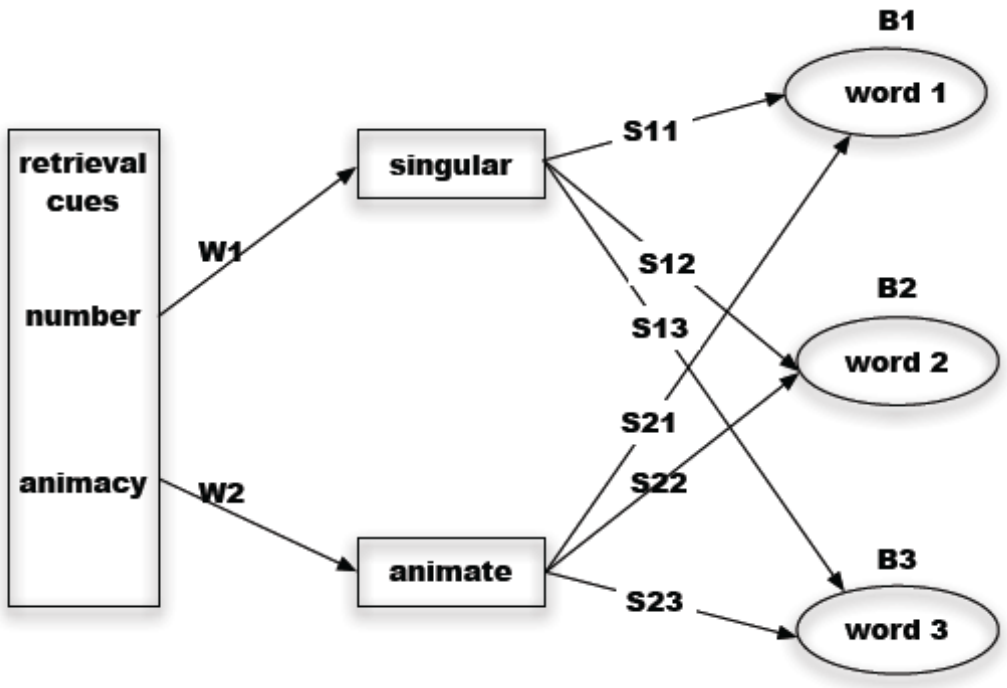
The memory constraints in the model

Code:
https://github.com/felixengelmann/act-r-sentence-parser-em

Latency factor F (:lf) **—> Speed**	$RT = Fe^{-(f*A_i)}$
Decay parameter d (:bll) **—> Speed, forgetting**	$B_i = \ln(\sum_{j=1}^{n} t_j^{-d}) + \beta_i$
Source activation W_k of buffer k (e.g., goalbuffer :ga) This activation is distributed among goal-related chunks. **—> Accuracy (goal-relevant), speed**	$A_i = B_i + S_i + P_i + \varepsilon_i$ $S_i = \sum_k \sum_j W_{kj} S_{ji}$
Mismatch penalty P (:mp) **—> Error sensitivity**	$P_i = \sum_k PM_{ki}$
Similarity M_{ki} between the value k in the retrieval specification and the value in the corresponding slot of chunk i **—> Association between cue and target**	

Introduction and background

The memory constraints in the model: Similarity based interference

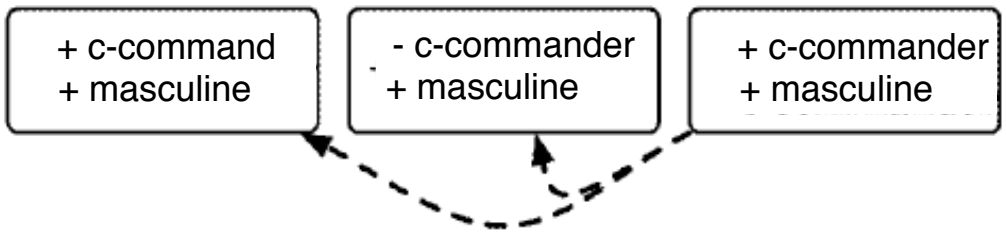


Introduction and background

The memory constraints in the model: Partial Matching

The tough **soldier** who Kathy met killed **himself**.

The tough **soldier** who **Bill** met killed **himself**.



* The tough **girl** who **Kathy** met killed **himself**.

Introduction and background

Possible evidence for partial matching: Processing polarity ([23] cf. [24, 25, 22])

Source: [22]

(11) a. No diplomats that a congressman would trust have ever supported a drone strike.

b. *The diplomats that no congressman could trust have ever supported a drone strike

c. *The diplomats that a congressman would trust have ever supported a drone strike.

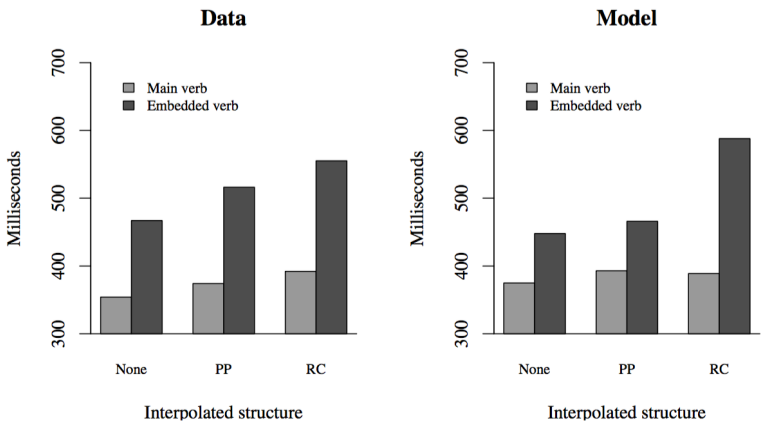
Condition	Data	Model
(11a) Accessible licensor	85	96
(11b) Inaccessible licensor	**70**	**61**
(11c) No licensor	83	86

Introduction: constraints on dependency completion

Constraints on retrieval

Consider again the Grodner and Gibson 05 results and our model [1] results:

The nurse { ∅ / from the clinic / who was from the clinic } supervised the administrator...



Introduction: constraints on dependency completion

Lewis & Vasishth 2005, Jäger, Engelmann, Vasishth 2017

Paper: [26]

(12) a. *Target-match; distractor-mismatch*
The surgeon$^{+masc}_{+\,c\text{-}com}$ who treated Jennifer$^{-masc}_{-\,c\text{-}com}$ had pricked himself$\{^{masc}_{c\text{-}com}\}$...

b. *Target-match; distractor-match*
The surgeon$^{+masc}_{+\,c\text{-}com}$ who treated Jonathan$^{+masc}_{-\,c\text{-}com}$ had pricked himself$\{^{masc}_{c\text{-}com}\}$...

Introduction: constraints on dependency completion

Lewis & Vasishth 2005, Jäger, Engelmann, Vasishth 2017

Another example of target match, from [27]:

(13) a. The worker was surprised that the [$_{\text{Target}}$ resident$^{+anim}_{+locSubj}$] who was living near the dangerous [$_{\text{Distractor}}$ warehouse$^{-anim}_{-locSubj}$] *was complaining*$\{^{anim}_{locSubj}\}$ about the investigation.

b. The worker was surprised that the [$_{\text{Target}}$ resident$^{+anim}_{+locSubj}$] who was living near the dangerous [neighbor$^{+anim}_{-locSubj}$] *was complaining*$\{^{anim}_{locSubj}\}$ about the investigation.

Modeling retrieval processes in sentence comprehension

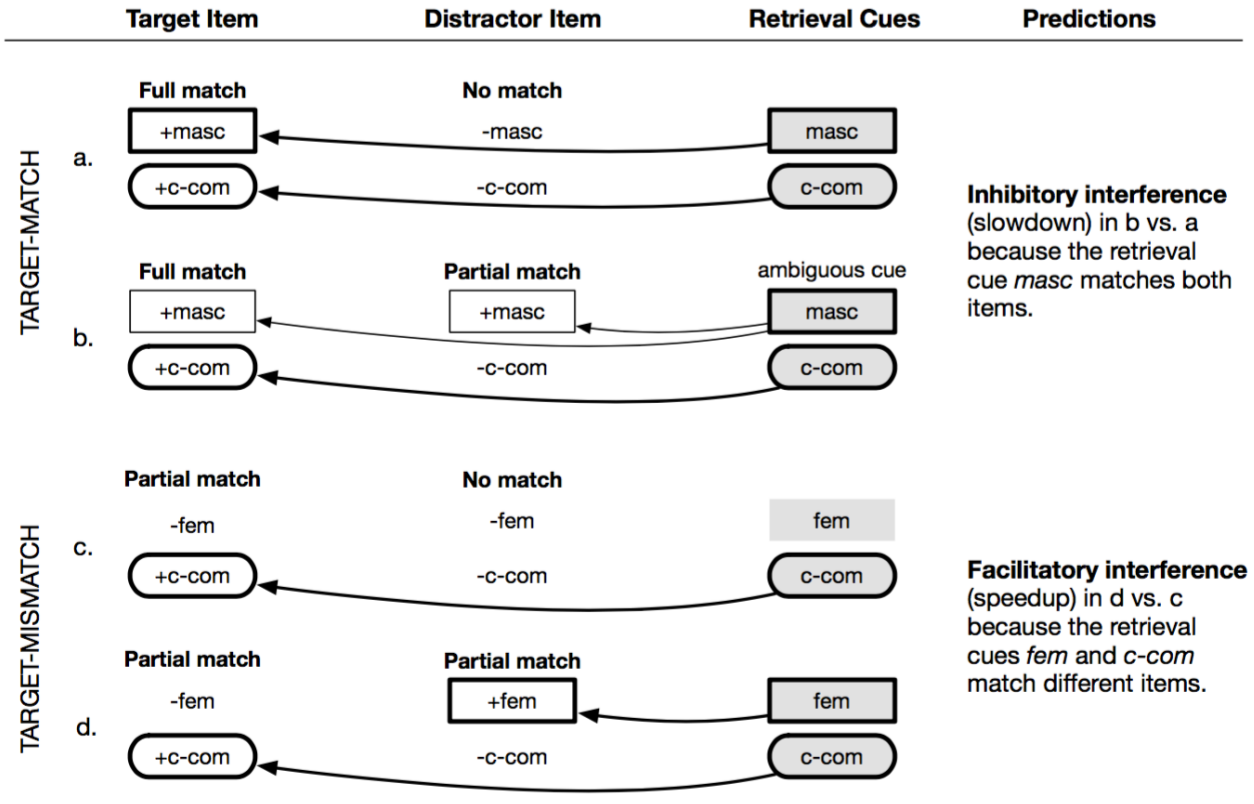
Lewis & Vasishth 2005, Jäger, Engelmann, Vasishth 2017

(14) a. *Target-mismatch; distractor-mismatch*
The surgeon$^{-fem}_{+c\text{-}com}$ who treated Jonathan$^{-fem}_{-c\text{-}com}$ had pricked herself$\{^{fem}_{c\text{-}com}\}$...

b. *Target-mismatch; distractor-match*
The surgeon$^{-fem}_{+c\text{-}com}$ who treated Jennifer$^{+fem}_{-c\text{-}com}$ had pricked herself$\{^{fem}_{c\text{-}com}\}$...

Modeling retrieval processes in parsing

Lewis & Vasishth 2005, Jäger, Engelmann, Vasishth 2017



Modeling retrieval processes in parsing

Lewis & Vasishth 2005, Jäger, Engelmann, Vasishth 2017

Agreement attraction could also be an instance of similarity-based interference:

(15) a. The key$_{+sing}$ to the cabinet$_{+sing}$ is in the box.

b. The key$_{+sing}$ to the cabinets$_{+plur}$ is in the box.

c. * The key$_{+sing}$ to the cabinet$_{+sing}$ are in the box.

d. * The key$_{+sing}$ to the cabinets$_{+plur}$ are in the box.

Modeling retrieval processes in parsing

Jäger, Engelmann, Vasishth 2017

We carried out a random-effects meta-analysis to evaluate the overall effect of interference in the grammatical (target match) and ungrammatical (target mismatch) configurations.

Journal of Memory and Language 94 (2017) 316–339

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis

Lena A. Jäger [a,*], Felix Engelmann [b], Shravan Vasishth [a,c]

[a] University of Potsdam, Germany

Random-effects meta-analysis

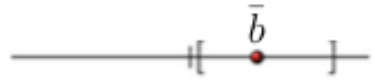
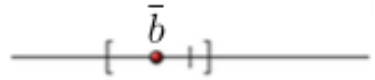
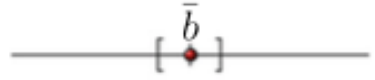
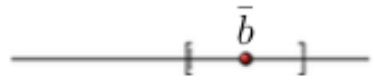
Jäger, Engelmann, Vasishth 2017

$$
\begin{aligned}
y_i \mid \theta_i, \sigma_i^2 &\sim N(\theta_i, \sigma_i^2) \quad i = 1, \ldots, n \\
\theta_i \mid \theta, \tau^2 &\sim N(\theta, \tau^2), \\
\theta &\sim N(0, 100^2), \\
1/\tau^2 &\sim Gamma(0.001, 0.001) \\
\text{OR} : \tau &\sim Uniform(0, 200) \\
\text{OR} : \tau &\sim Normal(0, 200^2) I(0,)
\end{aligned} \tag{1}
$$

1. y_i is the effect size in milliseconds in the i-th study.
2. θ is the true (unknown) effect, to be estimated by the model.
3. σ_i^2 is the true variance of the sampling distribution; each σ_i is estimated from the standard error in study i.
4. The variance parameter τ^2 represents between-study variance.

Random-effects meta-analysis: The effect of interference

Jäger, Engelmann, Vasishth 2017

Dependency	Target	Evidence	LV05 prediction
Subject-verb (non-agreement)	Match		inhibition ✓
Subject-verb agreement	Match		inhibition ✗
	Mismatch		facilitation ✓
Reflexive-/reciprocal-antecedent	Match		inhibition ✗
	Mismatch		facilitation ✗

The effect of number

Jäger, Engelmann, Vasishth 2017

- The effect of the number cue in grammatical agreement attraction configurations does not match the predictions of the ACT-R account.
- ACT-R predicts a slowdown in grammatical (target match) conditions, but people tend to report a speedup.

(16)
a. The key$_{+sing}$ to the cabinet$_{+sing}$ is in the box.
b. The key$_{+sing}$ to the cabinets$_{+plur}$ is in the box.
c. * The key$_{+sing}$ to the cabinet$_{+sing}$ are in the box.
d. * The key$_{+sing}$ to the cabinets$_{+plur}$ are in the box.

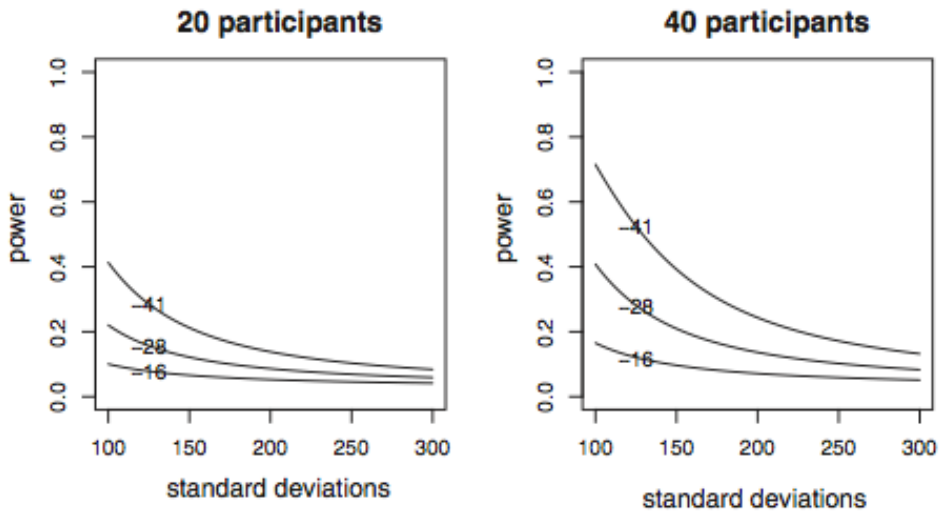
The effect of number

Jäger, Engelmann, Vasishth 2017

- The effects of the meta-analysis for target-match do not match theoretical predictions.
- The Lewis & Vasishth (2005) model predicts a slowdown. But we see that when number is a cue, the interference effect is reduced in magnitude.
- We speculated that this observed reduction in interference might be a Type S error (and possibly also a result of publication bias).

The effect of number: low power?

Jäger, Engelmann, Vasishth 2017



Number interference in German

Nicenboim et al 2017

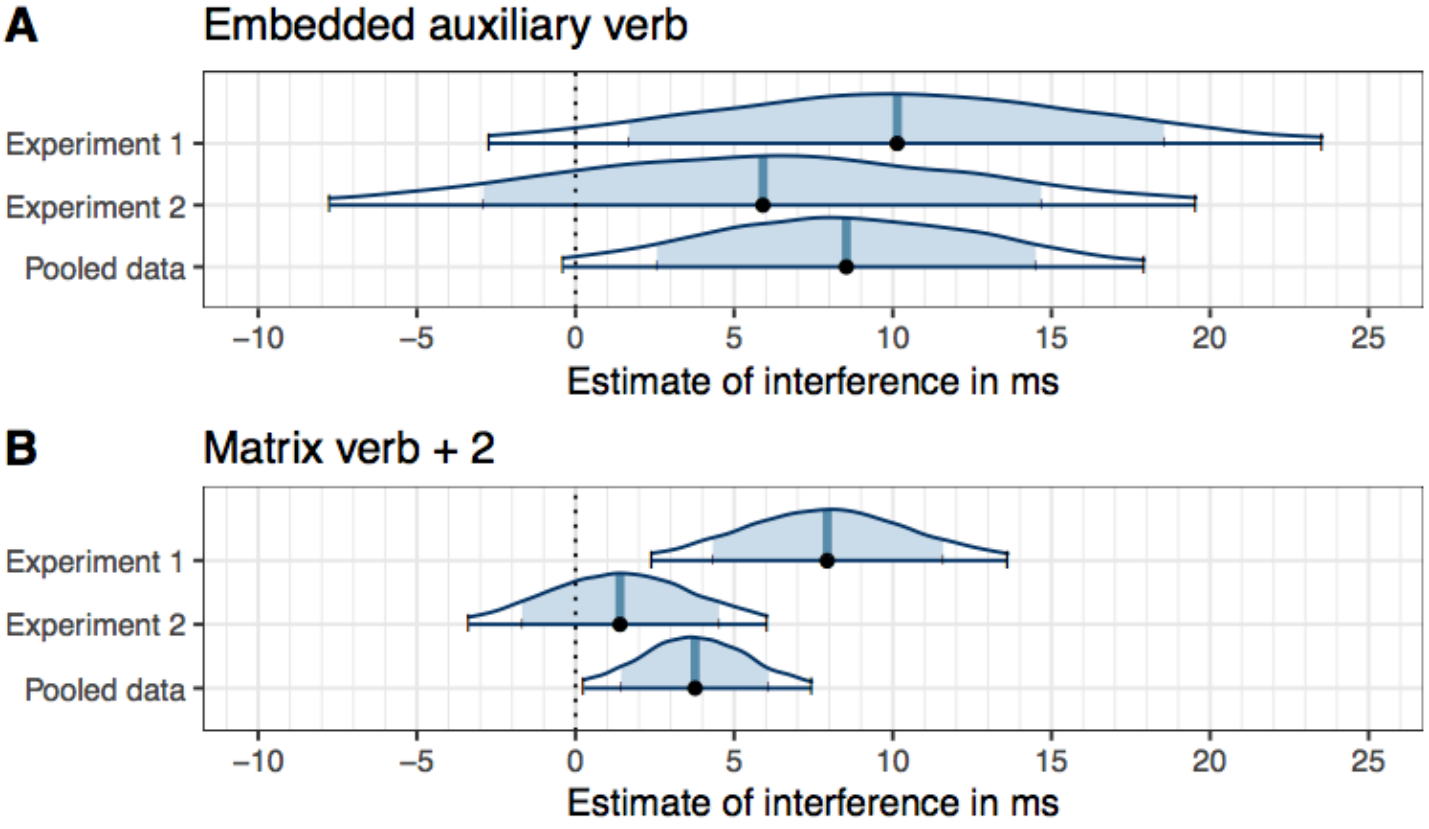
Two studies with 60 items: Exploratory (84 subjects), confirmatory (100). Paper: [28]

(17) a. High Interference
Der Wohltäter, der den Assistenten des Direktors **begrüsst hatte**, sass später im Spendenausschuss.

b. Low Interference
Der Wohltäter, der die Assistenten der Direktoren **begrüsst hatte**, sass später im Spendenausschuss.

Number interference in German: Results

Nicenboim et al 2017



Number interference in German: Discussion

Nicenboim et al 2017

- There might be a small effect (9 ms, 0-18 ms credible interval) of number interference.
- The effect in our studies is consistent with theory.
- However, the Nicenboim has a stronger interference manipulation (two intruders rather than one).
- What is needed next is cross-linguistically replication of the effect.

Intervention does not have a unitary effect on language comprehension

1. Intervention can lead to *greater difficulty* in completing dependencies due to
   - backward looking processes: interference+decay constraints
   - forward looking processes: storage cost (Gibson 2000), predictions *increasing* entropy (Safavi et al 2016)
2. Intervention can also lead to *reductions in difficulty*:
   - through misretrievals of the intervener (NPIs, agreement attraction)
   - through richer encoding of arguments (Hofmeister 2007)
   - by sharpening expectations (surprisal effects, Levy 2008).
   - by *lowering* entropy (Linzen and Jaeger 2015)

Some key factors in play when interveners are present

So, the effect of intervention seems to depend at least on:

1. what **retrieval processes** are triggered after intervention region
2. what **linguistic retrieval cues** are deployed during retrieval (morphosyntactic cues play a role here)
3. **individual differences** in working memory capacity and cognitive control
4. **systematic underspecification**, which could lead to subject not completing a dependency
5. what is being **predicted** (could vary by individual-level ability to maintain predictions)
6. **what information** is contained in the intervener (could sharpen or weak predictions)

Intervention as a cause of complexity

Explanations for speedups and slowdowns in dependency completion have three major classes of explanation:

1. Inhibition and facilitation arising from constraints on retrieval
   The computational model of Lewis and Vasishth 2005, and recent extensions by Engelmann et al.
2. Inhibition and facilitation arising from capacity-induced retrieval failures
   The computational model of Nicenboim et al. 2016
3. Inhibition and facilitation arising from predictive-processing (storage, surprisal, entropy)
   The probabilistic parsing computational models of Hale 2001 and Levy 2008

I will discuss these three classes of explanation today.

Interference arising from constraints on retrieval

Example 1a: Slowdowns due to proactive SBI in reflexives

Similarity-based interference (retrieval cues matching multiple nouns) leads to increased difficulty in completing noun-reflexive dependency:

Badecker and Straub 2002:

[Distractor Jane$^{-masc}_{-\underline{c\text{-}com}}$/John$^{+masc}_{-\underline{c\text{-}com}}$] thought that [Target Bill$^{+masc}_{+c\text{-}com}$] owed *himself* $\{^{masc}_{c\text{-}com}\}$ another opportunity to solve the problem.

Interference arising from constraints on retrieval

Example 1b: Slowdowns due to retroactive SBI in subject-verb dependencies

Van Dyke 2007: Another example of SBI leading to increased difficulty in completing argument-verb dependency:

The worker was surprised that the [$_{\text{Target}}$ resident$^{+anim}_{+locSubj}$] who was living near the dangerous
[$_{\text{Distractor}}$ warehouse$^{-anim}_{-locSubj}$/neighbor$^{+anim}_{-locSubj}$]
was complaining$\{^{anim}_{locSubj}\}$ about the investigation.

Interference arising from constraints on retrieval

Example 1b: Slowdowns due to retroactive SBI in subject-verb dependencies

Van Dyke 2007: Another example of SBI leading to increased difficulty in completing argument-verb dependency:

The worker was surprised that the [$_{\text{Target}}$ resident$^{+anim}_{+locSubj}$] who was living near the dangerous
[$_{\text{Distractor}}$ warehouse$^{-anim}_{-locSubj}$/neighbor$^{+anim}_{-locSubj}$]
was complaining$\{^{anim}_{locSubj}\}$ about the investigation.

Cf: Rizzi 2013:

(19) a. ?[Which problem$^{+Q}_{+N}$] do you wonder how$^{+Q}_{-N}$ to solve GAP$^{+Q}_{+N}$

b. *How$^{+Q}_{-N}$ do you wonder which problem$^{+Q}_{+N}$ to solve GAP$^{+Q}_{+N}$

Interference arising from constraints on retrieval

Example 2a: Speedups due to misretrievals in subject-verb dependencies (agreement attraction)

*The [$_{\text{Target}}$ key$^{-plur}_{+locSubj}$ to the [$_{\text{Distractor}}$ cabinet$^{-plur}_{-locSubj}$/cabinets$^{+plur}_{-locSubj}$] *were*$\{^{plur}_{locSubj}\}$ rusty from many years of disuse.

This facilitation due to misretrieval is predicted for reflexives, but there is only limited evidence for this at the moment (King et al 2012):

The [$_{\text{Target}}$ *mechanic*$^{-fem}_{+c\text{-}com}$] who spoke to [$_{\text{Distractor}}$ John$^{-fem}_{-c\text{-}com}$/Mary$^{+fem}_{-c\text{-}com}$] sent a package to *herself*$\{^{fem}_{c\text{-}com}\}$ ...

Inhibition and facilitation due to capacity-induced retrieval failure

- We usually assume that **longer** reading times index **increased** processing difficulty and **shorter** reading times index **reduction** in processing difficulty
- What if **shorter RTs** sometimes indexed **increased** processing difficulty?
- We explored this question in:
  Bruno Nicenboim, Pavel Logačev, Carolina Gattei, and Shravan Vasishth. When high-capacity readers slow down and low-capacity readers speed up: Working memory differences in unbounded dependencies. 2016, *Frontiers in Psychology, Special Issue on Encoding and Navigating Linguistic Representations in Memory*.

Inhibition and facilitation due to capacity-induced retrieval failure

Nicenboim et al 2016

We investigated short vs long wh-dependencies in Spanish and German (both SPR).

- SHORT Someone asked **what** the man **did** last summer.
- LONG Someone asked **what** the man [INTERVENER] **did** last summer.

Inhibition and facilitation due to capacity-induced retrieval failure

Nicenboim et al 2016

(20) a. SHORT - UNBOUNDED DEPENDENCY

La hermana menor de Sof?a pregunt? **a qui?n**
The younger sister of Sofia asked **who.ACC**
fue que Mar?a **hab?a saludado** ...
was that Mar?a **had greeted** ...

b. LONG - UNBOUNDED DEPENDENCY

Sof?a pregunt? **a qui?n** fue que
Sofia asked **who.ACC** was that
la hermana menor de Mar?a **hab?a saludado** ...
the younger sister of Mar?a **had greeted** ...

Inhibition and facilitation due to capacity-induced retrieval failure

Nicenboim et al 2016

(21) a. SHORT - BASELINE

La hermana menor de Sof?a pregunt? si Mar?a
The younger sister of Sofia asked if Maria
hab?a saludado ...
had greeted ...

b. LONG - BASELINE

Sof?a pregunt? si la hermana menor de Mar?a
Sofia asked if the younger sister of Maria
hab?a saludado ...
had greeted ...

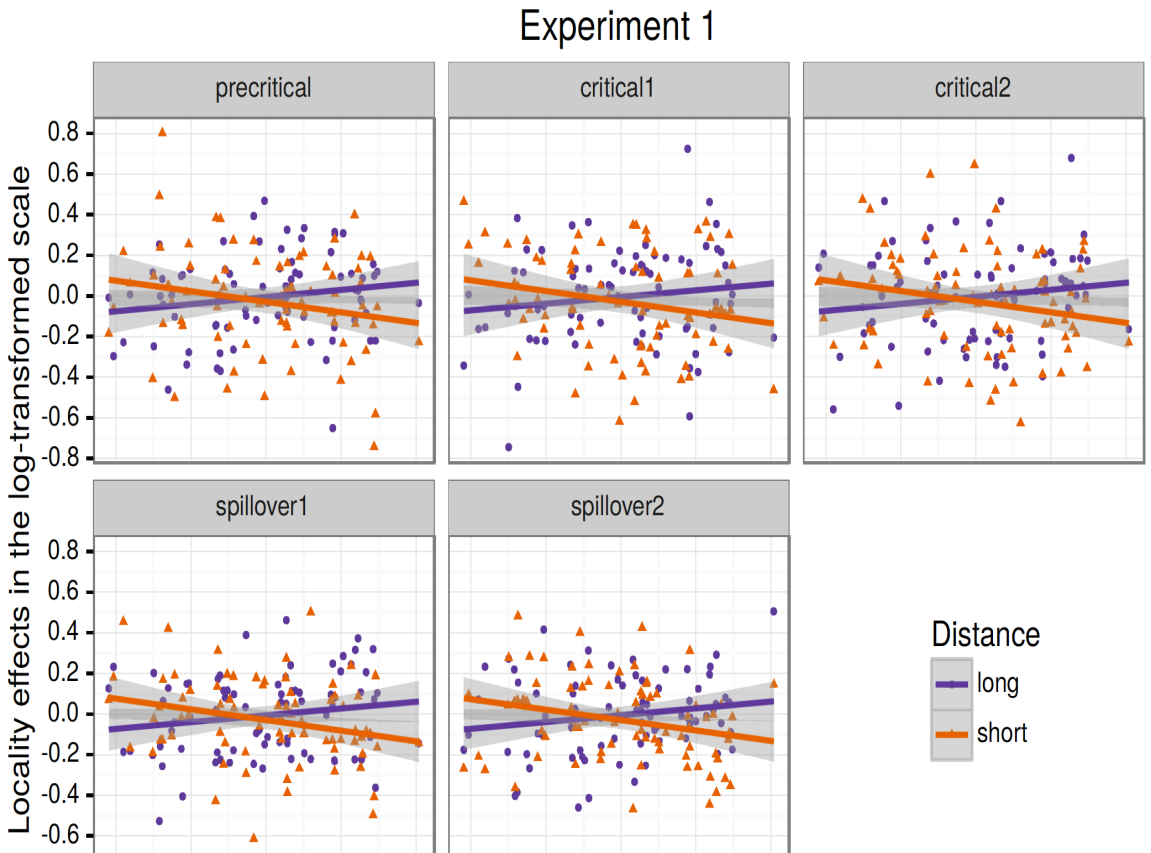
Inhibition and facilitation due to capacity-induced retrieval failure

Nicenboim et al 2016

- One question was: does working memory capacity affect dependency resolution difficulty?
- The Lewis & Vasishth ACT-R model predicts stronger intervention effects for low capacities compared to high capacities.
- This prediction turns out to be wrong.

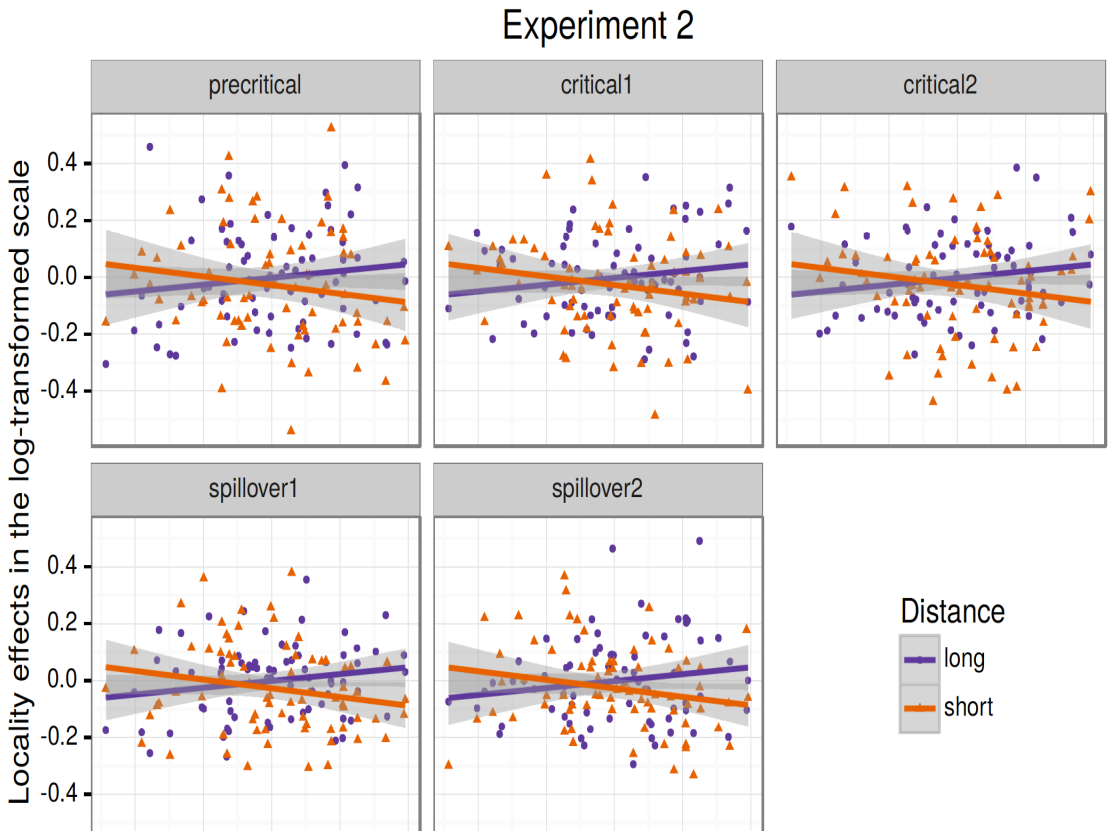
Inhibition and facilitation due to capacity-induced retrieval failure

Nicenboim et al 2016



Inhibition and facilitation due to capacity-induced retrieval failure

Nicenboim et al 2016



Inhibition and facilitation due to capacity-induced retrieval failure

Nicenboim et al 2016

- High capacity readers showed slowdowns due to intervention, but low capacities showed a speed-up (we see a graded crossover in the figures).
- Increase in processing difficulty may have different effects as a function of working memory capacity.
- Intervention may lead to slowdowns in high-capacity readers and speedups in low-capacity ones.
- A computational implementation suggests one possible explanation:
  - low capacities experience retrieval failure more frequently
  - retrieval failures end faster on average, because items in memory with an activation below a certain threshold may show shorter latencies due to an early aborting of the retrieval process.

Inhibition and facilitation due to predictive processing (surprisal, entropy)

Husain, Narayanan, Vasishth 2014

Surprisal claims that intervention sharpens expectations, leading to faclitation (Levy 2008).
But facilitation depends on something we call "expectation strength":

Definition of expectation strength

- Strong expectation: Prediction of exact verb (or verb class)
- Weak expectation: Prediction of some unknown verb

Inhibition and facilitation due to predictive processing (surprisal, entropy)

Husain, Narayanan, Vasishth 2014

- 2×2 design.
  - Predicate type: Complex vs. Simple
  - Distance between noun and verb: Long vs. Short
- A sentence completion study ensured that in CP cases, the nominal host predicts the 'exact' verb with high probability:
  - Mean percentage of exact verb predicted in Complex Predicate conditions: 86%.
  - Mean percentage of exact verb predicted in Simple Predicate conditions: 17%.

Inhibition and facilitation due to predictive processing (surprisal, entropy)

Husain, Narayanan, Vasishth 2014

Schematic of design:

(22) a. *CP-Long*: . . . noun-host 2-3-adjuncts verb . . .

b. *CP-Short*: . . . noun-host 1-2-adjuncts verb . . .

c. *SP-Long*: . . . noun.obj 2-3-adjuncts verb . . .

d. *SP-Short*: . . . noun-obj 1-2-adjuncts verb . . .

Inhibition and facilitation due to predictive processing (surprisal, entropy)

Husain, Narayanan, Vasishth 2014

(23) a. **Complex Predicate Conditions**

maa ne / bachche ko / skUla / CoRaa / Ora /
mother ERG child ACC school dropped and
usse / kahaa / ki / vah / apnaa / khayaal /
to her said that she her care
binaa kisi laaparvaahi ke / achCe se / rakhe, / phir /
without any carelessness properly keep, then
vah / apne / daftar ki ora / chal paRii
she her towards office proceeded

'The mother dropped the child off at the school and asked her to take care of herself properly without any carelessness, she then proceeded towards her office.'

Inhibition and facilitation due to predictive processing (surprisal, entropy)

Husain, Narayanan, Vasishth 2014

(24) a. **Simple Predicate Conditions**

maa ne / bachche ko / skUla / CoRaa / Ora /
mother ERG child ACC school dropped and
usse / kahaa / ki / vah / apnaa / gitaar /
to her said that she her guitar
binaa kisi laaparvaahi ke / achCe se / rakhe, / phir /
without any carelessness properly keep, then
vah / apne / daftar ki ora / chal paRii
she her towards office proceeded

'The mother dropped the child off at the school and asked her to keep her guitar properly without any carelessness, she then proceeded towards her office.'

Inhibition and facilitation due to predictive processing (surprisal, entropy)

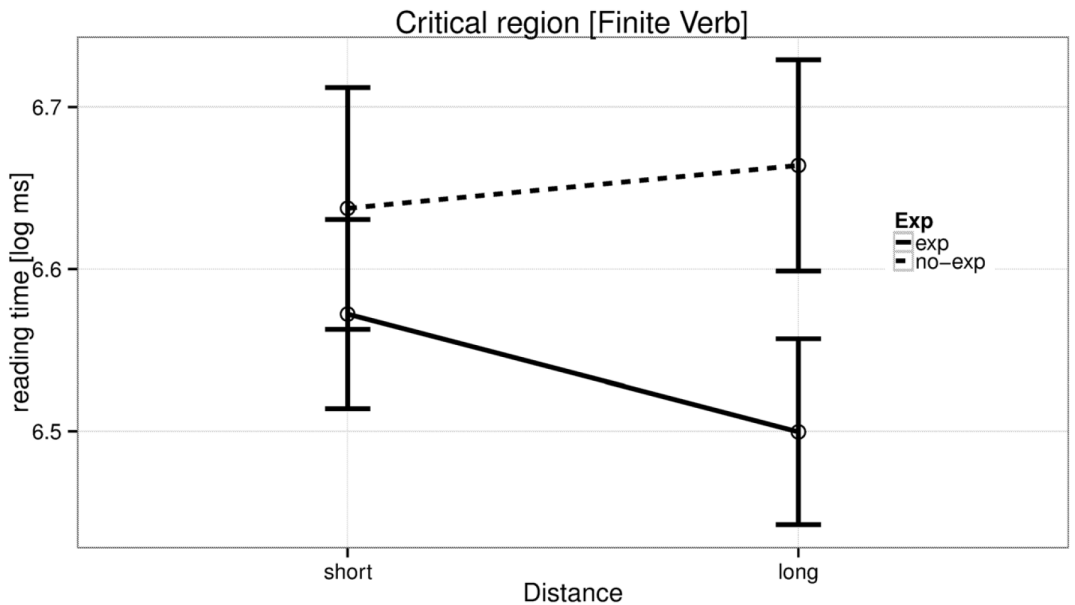
Husain, Narayanan, Vasishth 2014

Three alternative predictions:

- (Our version of) Expectation-based processing:
  - Reading time at the verb in the long conditions (SP-Long and CP-Long) should be faster than the short conditions.
  - (New:) In the CP condition, a strong prediction of verb identity is generated. So, the speed-up in the CP condition should be greater than in SP condition.
- Classical interference accounts (Gibson 2005, LV05):
  In both CP and SP, Long conditions should be slower than Short.
- Interference and Expectation together:
  If strong prediction strength modulates the effects of locality, speed-up in CP-Long vs CP-Short, but slowdown in SP-Long vs SP-short.

Inhibition and facilitation due to predictive processing (surprisal, entropy)

Husain, Narayanan, Vasishth 2014



[1] Richard L. Lewis and Shravan Vasishth.
An activation-based model of sentence processing as skilled memory retrieval.
Cognitive Science, 29:1–45, 2005.

[2] P. N. Johnson-Laird.
Mental Models: Towards a cognitive science of language, inference, and consciousness.
Harvard University Press, Cambridge, MA, 1983.

[3] Philip Resnik.
Left–corner parsing and psychological plausibility.
In Proceedings of COLING, pages 191–197, address not known, 1992.

[4] S. Abney and M. Johnson.
Memory requirements and local ambiguities of parsing strategies.
Journal of Psycholinguistic Research, 20(3):233–250, 1991.

[5] George Miller and Noam Chomsky.
Finitary models of language users.
In R. Duncan Luce, Robert R. Bush, and Eugene Galanter, editors, Handbook of Mathematical Psychology, Volume II, pages 419–492. John Wiley, 1963.

[6] Daniel Jurafsky.
A probabilistic model of lexical and syntactic access and disambiguation.
Cognition, 20:137–194, 1996.

[7] Roger Levy.
Expectation-based syntactic comprehension.
Cognition, 106:1126–1177, 2008.

[8] Marisa F. Boston, John T. Hale, Shravan Vasishth, and Reinhold Kliegl.
Parallel processing and sentence comprehension difficulty.
Language and Cognitive Processes, 26(3):301–349, 2011.

[9] Lyn Frazier.
On Comprehending Sentences: Syntactic Parsing Strategies.
PhD thesis, University of Massachusetts, Amherst, 1979.

[10] Lyn Frazier and Keith Rayner.
Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences.
Cognitive Psychology, 14:178–210, 1982.

[11] L. Frazier.
Syntactic complexity.
In David R. Dowty, Lauri Kartunnen, and Arnold M. Zwicky, editors, Natural language parsing: Psychological, Computational, and Theoretical perspectives, pages 129–189. Cambridge University Press, 1985.

[12] John T. Hale.
A probabilistic Earley parser as a psycholinguistic model.
In Proceedings of the Second Meeting of the North American Chapter of the Association for Computational Linguistics, Pittsburgh, PA, 2001.

[13] John Hale.
The information conveyed by words in sentences.
Journal of Psycholinguistic Research, 32(2):101–123, March 2003.

[14] Matthew J Traxler, Martin J Pickering, and Charles Clifton.
Adjunct attachment is not a form of lexical ambiguity resolution.
Journal of Memory and Language, 39(4):558–592, 1998.

[15] Roger P.G. van Gompel, Martin J. Pickering, and Matthew J. Traxler.
Unrestricted race: A new model of syntactic ambiguity resolution.
In A Kennedy, R Radach, D Heller, and J Pynte, editors, Reading as a perceptual process. Elsevier, Oxford, 2000.

[16] Pavel Logačev and Shravan Vasishth.
A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension.
Cognitive Science, 2015.

[17] J. C. Trueswell, M. Tanenhaus, and S. M. Garnsey.
Semantic influences on parsing: Use of thematic role information in syntactic ambiguity resolution.
Journal of Memory and Language, 33:285–318, 1994.

[18] C. Clifton, B. Juhasz, J. Ashby, M. J. Traxler, M. T. Mohamed, R. S. Williams, R. K. Morris, and K. Rayner.
The use of thematic role information in parsing: Syntactic processing autonomy revisited.
Journal of Memory and Language, 49:317–334, 2003.

[19] Daniel Grodner and Edward Gibson.
Consequences of the serial nature of linguistic input.
Cognitive Science, 29:261–290, 2005.

[20] Brian Bartek, Richard L. Lewis, Shravan Vasishth, and Mason Smith.
In search of on-line locality effects in sentence comprehension.
Journal of Experimental Psychology: Learning, Memory and Cognition, 37(5):1178–1198, 2011.

[21] Whitney Tabor, Bruno Galantucci, and Daniel Richardson.
Effects of merely local syntactic coherence on sentence processing.
Journal of Memory and Language, 50:355–370, 2004.

[22] Dan Parker and Colin Phillips.
Negative polarity illusions and the format of hierarchical encodings in memory.
Under review with Cognition, 2016.

[23] Shravan Vasishth, Sven Bruessow, Richard L. Lewis, and Heiner Drenhaus.
Processing polarity: How the ungrammatical intrudes on the grammatical.
Cognitive Science, 32(4):685–712, 2008.

[24] M. Xiang, B. Dillon, and C. Phillips.
Illusory licensing effects across dependency types: ERP evidence.
Brain and Language, 108(1):40–55, 2009.

[25] Brian Dillon, Alan Mishler, Shayne Sloggett, and Colin Phillips.
A computational cognitive model of syntactic priming.
Journal of Memory and Language, 69(4):85–103, 2011.

[26] Lena A. Jäger, Felix Engelmann, and Shravan Vasishth.
Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis.
Journal of Memory and Language, 94:316–339, 2017.

[27] JA Van Dyke.
Interference effects from grammatically unavailable constituents during sentence processing.
Journal of Experimental Psychology: Learning Memory and Cognition, 33(2):407–30, 2007.

[28] Bruno Nicenboim, Felix Engelmann, Katja Suckow, and Shravan Vasishth.
Number interference in German: Evidence for cue-based retrieval.
submitted to Cognitive Science, 2017.