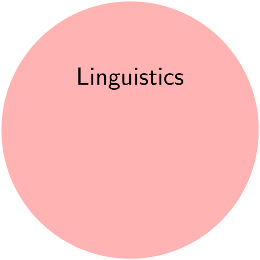
Towards a Computational Linking Hypothesis for Syntactic Theory

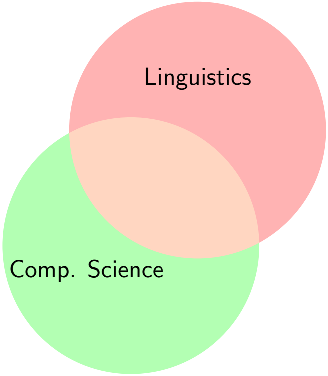
Aniello De Santo
he/him

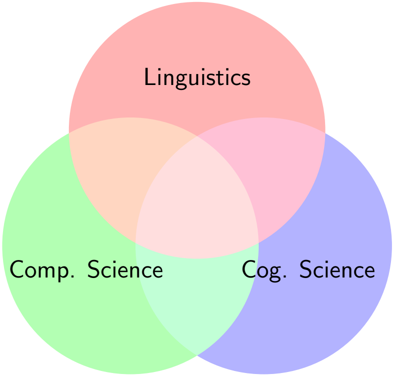
aniellodesanto.github.io/Pitt2024.pdf
aniello.desanto@utah.edu

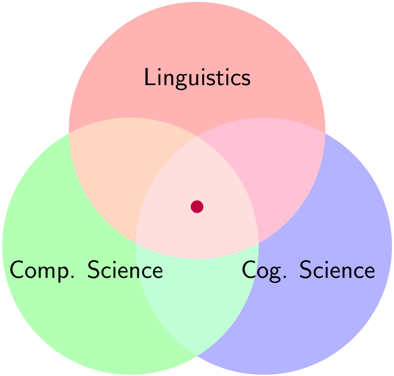
Get the slides!

University of Pittsburgh
January 2024

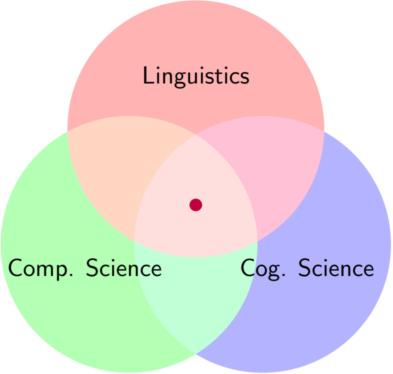

Linguistics


Linguistics
Comp. Science


Linguistics
Comp. Science
Cog. Science



- Modeling processing difficulty (De Santo 2019, 2021, 2022, a.o.)
- Evaluating/Contrasting syntactic analyses
  (De Santo & Shafiei 2019, Lee & De Santo 2022, a.o.)
- Gradience in acceptability judgment (De Santo 2020)
- Locality and Economy Considerations (De Santo & Lee 2022a)
- Online/Offline processing effects
  (De Santo & Lee 2022b, Lee & De Santo in prep., Jacobs, De Santo, Grobol in prep.)
- Memory traces of processing generalized quantifiers (De Santo et al. 2019, De Santo & Drury 2020)
- Theory building (De Santo & Rawski 2022, Baggio, De Santo, Nunez 2024)
- Animal Cognition (De Santo & Rawski, 2021)
- Complexity biases in typology and acquisition
  (De Santo 2018, Graf & De Santo 2020, De Santo & Gutierrez in prep., Johnson and De Santo 2023, in prep.)
- Computational parallels across linguistic modules
  (Aksenova & De Santo 2017, De Santo & Graf 2019, Miller & De Santo 2023, a.o.)
- Mapping syntactic and prosodic constituents (Vu, De Santo, Dolatian 2022)



- Modeling processing difficulty (De Santo 2019, 2021, 2022, a.o.)
- Evaluating/Contrasting syntactic analyses (De Santo & Shafiei 2019, Lee & De Santo 2022, a.o.)
- Gradience in acceptability judgment (De Santo 2020)
- Locality and Economy Considerations (De Santo & Lee 2022a)
- Online/Offline processing effects (De Santo & Lee 2022b, Lee & De Santo in prep., Jacobs, De Santo, Grobol in prep.)
- Memory traces of processing generalized quantifiers (De Santo et al. 2019, De Santo & Drury 2020)
- Theory building (De Santo & Rawski 2022, Baggio, De Santo, Nunez 2024)
- Animal Cognition (De Santo & Rawski, 2021)
- Complexity biases in typology and acquisition (De Santo 2018, Graf & De Santo 2020, De Santo & Gutierrez in prep., Johnson and De Santo 2023, in prep.)
- Computational parallels across linguistic modules (Aksenova & De Santo 2017, De Santo & Graf 2019, Miller & De Santo 2023, a.o.)
- Mapping syntactic and prosodic constituents (Vu, De Santo, Dolatian 2022)

Let's Start with Data!

Asymmetries in Italian Relative Clauses

Italian conforms to the general cross-linguistic preference for SRC over ORC (Adani et al. 2010; Arosio et al. 2018)

(1)

Il	cavallo	che	ha	inseguito	i	leoni
The	horse	that	has	chased	the	lions

"The horse that chased the lions" **SRC**

(2)

Il	cavallo	che	i	leoni	hanno	inseguito
The	horse	that	the	lions	have	chased

"The horse that the lions chased" **ORC**

SRC > ORC

Postverbal Subjects and Ambiguity

Italian allows for postverbal subjects, making some sentences ambiguous (De Vincenzi 1991):

(3) Il cavallo che ha inseguito il leone
The horse that has chased the lion

a. "The horse that chased the lion" **SRC**

b. "The horse that the lion chased" **ORCp**

SRC > ORCp

Agreement can disambiguate:

(4) Il cavallo che hanno inseguito i leoni
The horse that have chased the lions
"The horse that the lions chased" **ORCp**

Postverbal Subjects and Ambiguity

Italian allows for postverbal subjects, making some sentences ambiguous (De Vincenzi 1991):

(3) Il cavallo che ha inseguito il leone
The horse that has chased the lion

a. "The horse that chased the lion" **SRC**

b. "The horse that the lion chased" **ORCp**

SRC > ORCp

Agreement can disambiguate:

(4) Il cavallo che hanno inseguito i leoni
The horse that have chased the lions

"The horse that the lions chased" **ORCp**

Postverbal Subjects and Ambiguity

Italian allows for postverbal subjects, making some sentences ambiguous (De Vincenzi 1991):

(3) Il cavallo che ha inseguito il leone
The horse that has chased the lion

a. "The horse that chased the lion" **SRC**

b. "The horse that the lion chased" **ORCp**

SRC > ORCp

Agreement can disambiguate:

(4) Il cavallo che hanno inseguito i leoni
The horse that have chased the lions

"The horse that the lions chased" **ORCp**

Asymmetries in Italian Relative Clauses

(1) Il cavallo che ha inseguito i leoni
The horse that has chased the lions
"The horse that chased the lions" **SRC**

(2) Il cavallo che i leoni hanno inseguito
The horse that the lions have chased
"The horse that the lions chased" **ORC**

(4) Il cavallo che hanno inseguito i leoni
The horse that have chased the lions
"The horse that the lions chased" **ORCp**

Processing Asymmetry (De Vincenzi 1991, Arosio et al. 2018, a.o.)

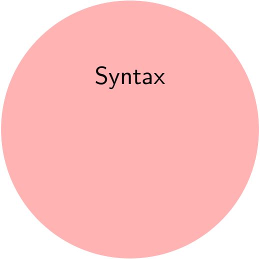
SRC > ORC > ORCp

One Big Question

(How much) does grammatical structure matter in sentence processing?

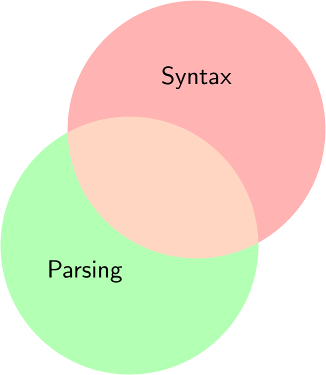
One Big Question

(How much) does grammatical structure matter in sentence processing?



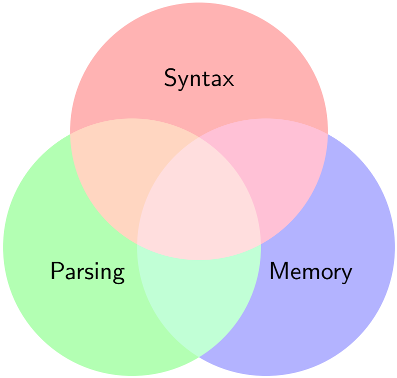
One Big Question

(How much) does grammatical structure matter in sentence processing?



One Big Question

(How much) does grammatical structure matter in sentence processing?



Forward to the Past

- What is the relation between grammatical operations and cognitive processes?

Derivational Theory of Complexity (Miller and Chomsky, 1963)

- Processing complexity $\sim$ length of a derivation (Fodor & Garrett 1967; Berwick & Weinberg 1983)
- Essentially: there is a **cost** to mental computations.

- What is the right notion of syntactic derivation?
- What is costly? And why?

Forward to the Past

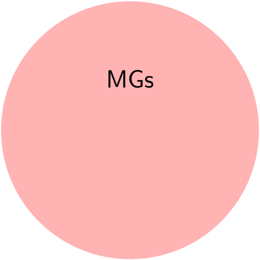
- What is the relation between grammatical operations and cognitive processes?

Derivational Theory of Complexity (Miller and Chomsky, 1963)

- Processing complexity $\sim$ length of a derivation (Fodor & Garrett 1967; Berwick & Weinberg 1983)
- Essentially: there is a **cost** to mental computations.

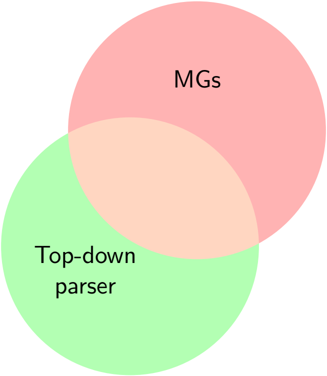
- What is the right notion of syntactic derivation?
- What is costly? And why?

A Formal Model of Sentence Processing



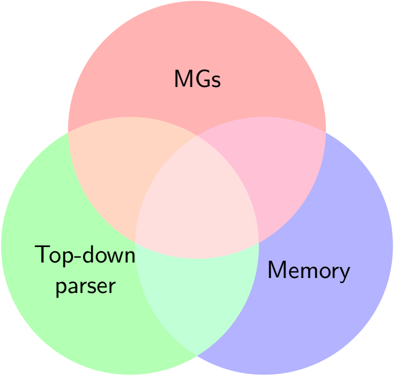
1. An explicit syntactic theory → Minimalist grammars (MGs)

A Formal Model of Sentence Processing



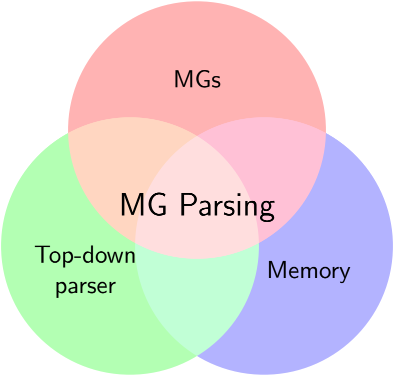
1. An explicit syntactic theory → Minimalist grammars (MGs)
2. A theory of how structures are built → Top-down parser

A Formal Model of Sentence Processing



1. An explicit syntactic theory → Minimalist grammars (MGs)
2. A theory of how structures are built → Top-down parser
3. A psychologically grounded notion of cost → Memory Usage

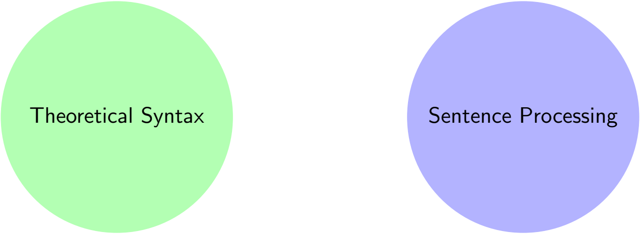
A Formal Model of Sentence Processing



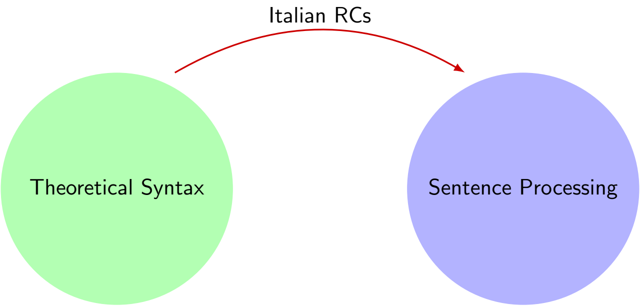
1. An explicit syntactic theory → Minimalist grammars (MGs)
2. A theory of how structures are built → Top-down parser
3. A psychologically grounded notion of cost → Memory Usage

Interpretability for the win!

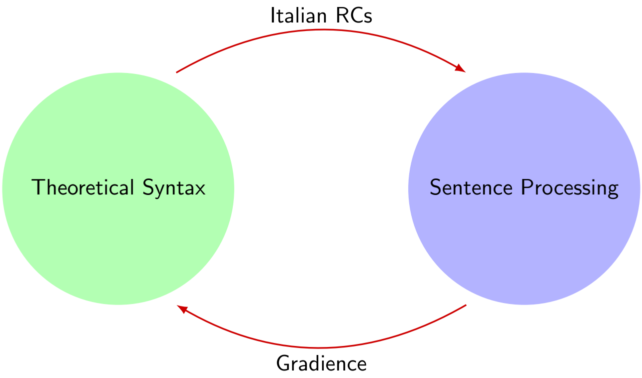
Building Bridges



Building Bridges



Building Bridges



Outline

Minimalist Grammars (MGs)

We need an explicit model of syntactic structures...

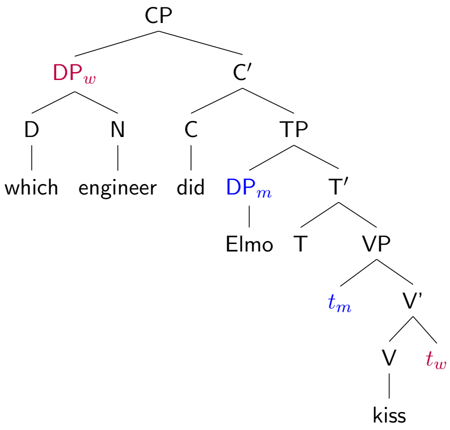
Ed Stabler

- Minimalist grammars (**MGs**): a formalization of Chomskyan syntax (Chomsky 1995; Stabler 1997)

Technical details!

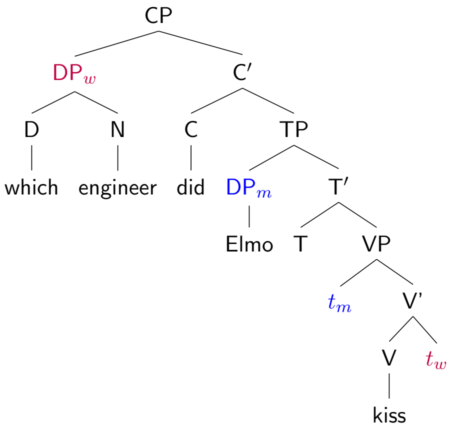
- Weakly equivalent to MCFGs
- Essentially: CFGs with a more complicated mapping from trees to strings

MG Syntax: Derivation Trees

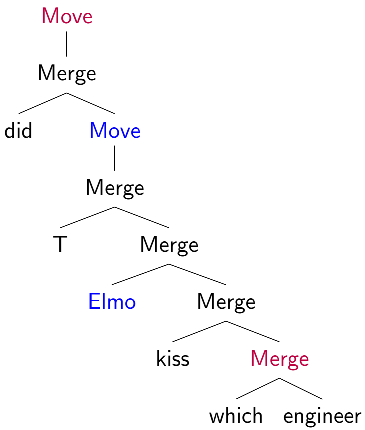


Phrase Structure Tree

MG Syntax: Derivation Trees

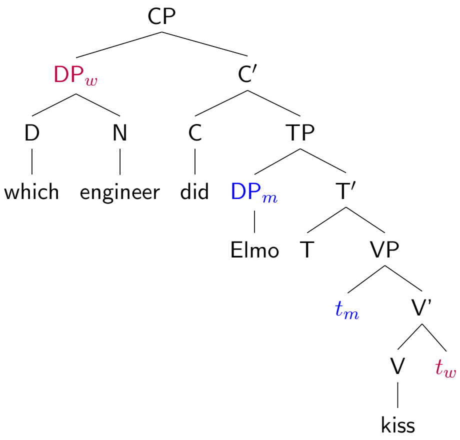


Phrase Structure Tree

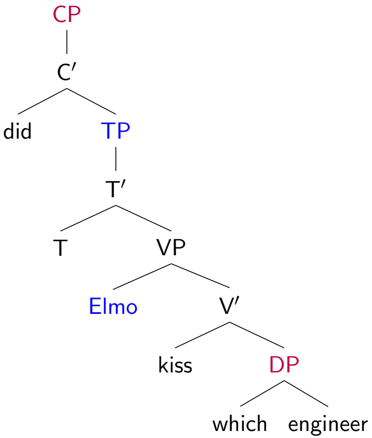


Derivation Tree

MG Syntax: Derivation Trees



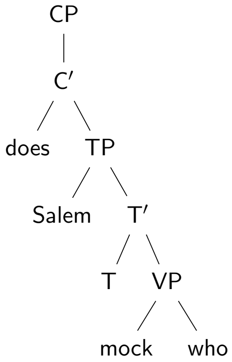
Phrase Structure Tree



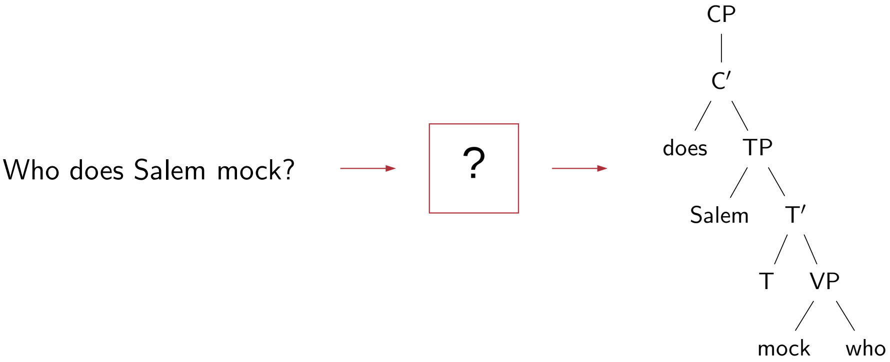
Derivation Tree

The Job of a Parser

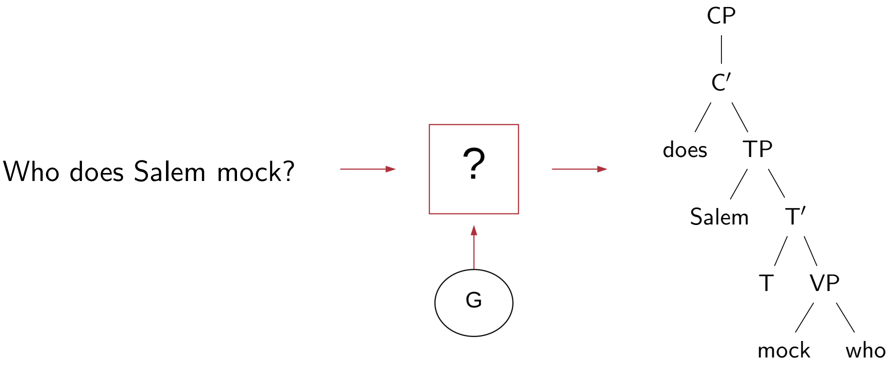
Who does Salem mock?



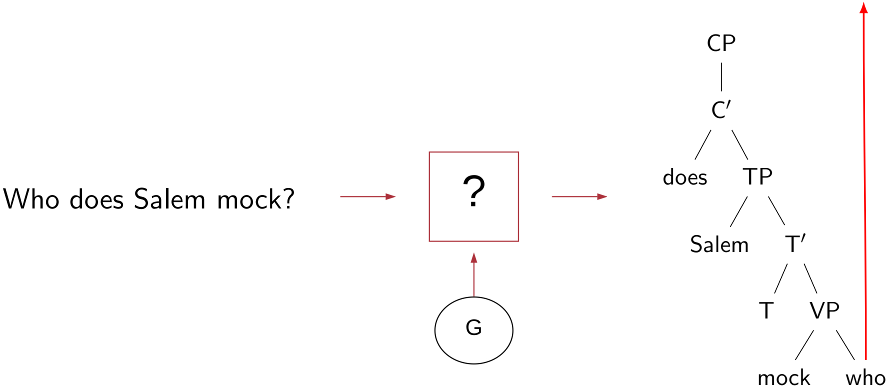
The Job of a Parser



The Job of a Parser

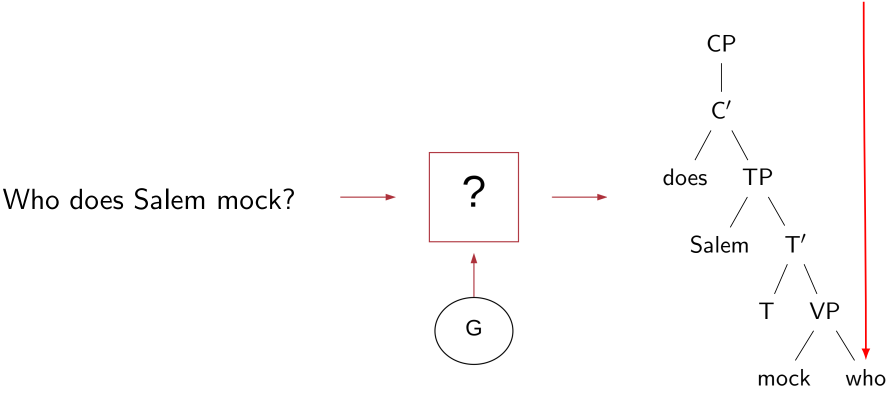


The Job of a Parser



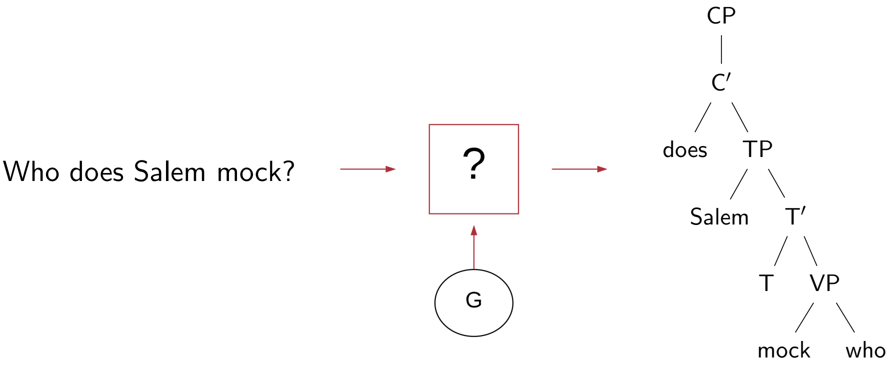
- Bottom-up

The Job of a Parser



- Bottom-up
- Top-down

The Job of a Parser



- Bottom-up
- Top-down
  - Psychologically plausible(-ish)
  - We can build bottom-up grammars top-down!
  - Big idealization: Parser as an oracle!

Top-Down Parsing: The Intuition

Who does Salem mock?

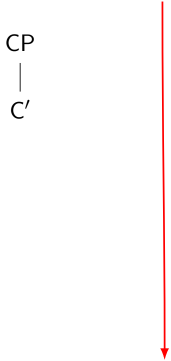
Top-Down Parsing: The Intuition

CP

Who does Salem mock?

- Builds the structure from top to bottom
- Takes elements in an out of memory
- Complexity of the structure $\approx$ how much memory is used!

Top-Down Parsing: The Intuition

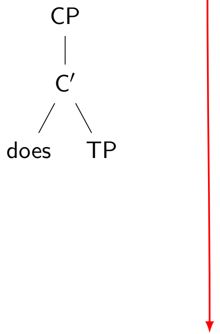


Who does Salem mock?

- Builds the structure from top to bottom
- Takes elements in an out of memory
- Complexity of the structure ≈ how much memory is used!

Top-Down Parsing: The Intuition

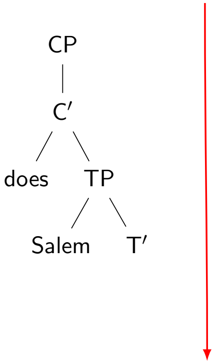
Who does Salem mock?



- Builds the structure from top to bottom
- Takes elements in an out of memory
- Complexity of the structure ≈ how much memory is used!

Top-Down Parsing: The Intuition

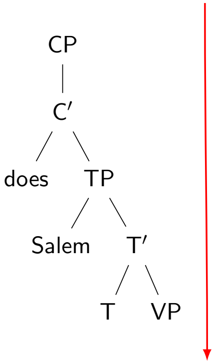
Who does Salem mock?



- Builds the structure from top to bottom
- Takes elements in an out of memory
- Complexity of the structure ≈ how much memory is used!

Top-Down Parsing: The Intuition

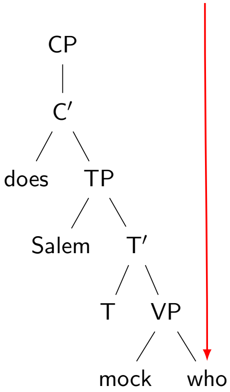
Who does Salem mock?



- Builds the structure from top to bottom
- Takes elements in an out of memory
- Complexity of the structure ≈ how much memory is used!

Top-Down Parsing: The Intuition

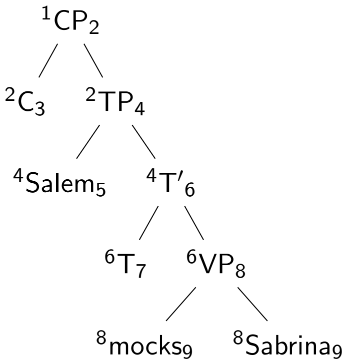
Who does Salem mock?



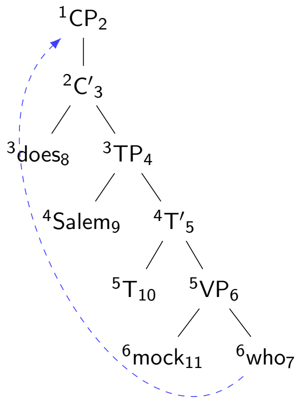
- Builds the structure from top to bottom
- Takes elements in an out of memory
- Complexity of the structure ≈ how much memory is used!

Contrasting Derivations

Memory Usage $= 2$



Memory Usage $= 5$



Summary of the Approach

General Idea

(Kobele et al. 2012; Gerth 2015; Graf et al. 2017; De Santo 2020)

1. Pick two competing derivations for a processing contrast
2. Annotate derivation trees and compute memory usage
3. Evaluate effort over each
   - Lowest score means easiest!
4. Compare parser's prediction to experimental data

Reminder: Asymmetries in Italian Relative Clauses

(1) Il cavallo che ha inseguito i leoni
The horse that has chased the lions
"The horse that chased the lions" **SRC**

(2) Il cavallo che i leoni hanno inseguito
The horse that the lions have chased
"The horse that the lions chased" **ORC**

(4) Il cavallo che hanno inseguito i leoni
The horse that have chased the lions
"The horse that the lions chased" **ORCp**

Processing Asymmetry (De Vincenzi 1991, Arosio et al. 2018, a.o.)

SRC > ORC > ORCp

Modeling Assumptions

Reminder:

- Parsing strategy
  $\Rightarrow$ Top-down parser
- Complexity Metrics
  $\Rightarrow$ Memory Usage

Degrees of freedom: Syntactic analyses

1. RC constructions $\rightarrow$ Kayne (1994)
2. Postverbal subjects $\rightarrow$ Belletti & Leonini (2004)

Modeling Assumptions

Reminder:

- Parsing strategy
  ⇒ Top-down parser
- Complexity Metrics
  ⇒ Memory Usage

Degrees of freedom: Syntactic analyses

1. RC constructions → Kayne (1994)
2. Postverbal subjects → Belletti & Leonini (2004)

Modeling Results

(1) Il cavallo che ha inseguito i leoni
The horse that has chased the lions
"The horse that chased the lions" **SRC**

(2) Il cavallo che i leoni hanno inseguito
The horse that the lions have chased
"The horse that the lions chased" **ORC**

(4) Il cavallo che hanno inseguito i leoni
The horse that have chased the lions
"The horse that the lions chased" **ORCp**

SRC $>$ ORC $>$ ORCp

Modeling Results

(1) Il cavallo che ha inseguito i leoni
The horse that has chased the lions
"The horse that chased the lions" **SRC**

(2) Il cavallo che i leoni hanno inseguito
The horse that the lions have chased
"The horse that the lions chased" **ORC**

(4) Il cavallo che hanno inseguito i leoni
The horse that have chased the lions
"The horse that the lions chased" **ORCp**

	SRC	>	ORC	>	ORCp
Memory	8/che		11/ha		16/Foc

Modeling Results

(1) Il cavallo che ha inseguito i leoni
The horse that has chased the lions
"The horse that chased the lions" **SRC**

(2) Il cavallo che i leoni hanno inseguito
The horse that the lions have chased
"The horse that the lions chased" **ORC**

(4) Il cavallo che hanno inseguito i leoni
The horse that have chased the lions
"The horse that the lions chased" **ORCp**

	SRC	>	ORC	>	ORCp	
Memory	8/che		11/ha		16/Foc	✓

Results across Constructions (De Santo 2019)

Clause Type	<Memory>
obj. SRC $>$ ORC	✓
obj. SRC $>$ ORCp	✓
obj. ORC $>$ ORCp	✓
subj. SRC $>$ ORC	✓
subj. SRC $>$ ORCp	✓
subj. ORC $>$ ORCp	✓
matrix SVO $>$ VOS	✓
VS unacc $>$ VS unerg	✓

Table: Predictions of the MG parser by contrast.

Results across Analyses (De Santo 2021)

		SRC < ORC	SRC < ORCp	ORC < ORCp
Postverbal	**RC Type**	MEMORY	MEMORY	MEMORY
Smuggling	Promotion	✓	✓	✓
	Wh-movement	✓	✓	✓
	Extraposition	✓	✓	✓
	DP analysis	✓	✓	✓
Scrambling	Promotion	✓	✓	✓
	Wh-movement	✓	✓	✓
	Extraposition	✓	tie	tie
	DP analysis	✓	tie	tie

Table: Predictions of the MG parser for the RC contrast by analysis.

Interim Summary

- This model gives surprisingly good results!
  - Simplistic model of processing:
    → "just" **(fine-grained) structural differences**!

Interim Summary

- This model gives surprisingly good results!
  - Simplistic model of processing:
    $\rightarrow$ "just" **(fine-grained) structural differences**!

Interim Summary

- This model gives surprisingly good results!
  - Simplistic model of processing:
    → "just" **(fine-grained) structural differences**!
- Asymmetries in Italian postverbal subjects
  - Expand range of syntactic constructions/analyses
    (De Santo 2021, De Santo & Shafiei 2019, in prep.)
  - Cross-linguistic comparison
    (Del Valle & De Santo, 2023;
    Fiorini, Chang, De Santo, 2023)

Nazila Shafiei

Dan Del Valle

Matteo Fiorini

Jillian Chang

Processing Asymmetries All the Way Down

Memory metrics make correct predictions cross-linguistically!

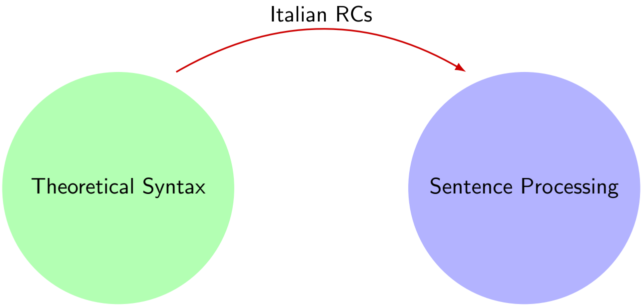
Across Constructions

- Right > center embedding (Kobele et al. 2012)
- Crossing > nested dependencies (Kobele et al. 2012)
- SRC > ORC (Graf et al. 2017, De Santo 2020)
- Postverbal subjects in Romance (De Santo 2019, 20, Del Valle & De Santo 2023)
- Attachment ambiguities (De Santo & Shafiei 2019, Lee & De Santo 2022)
- Structural Priming (De Santo 2020, 2021)

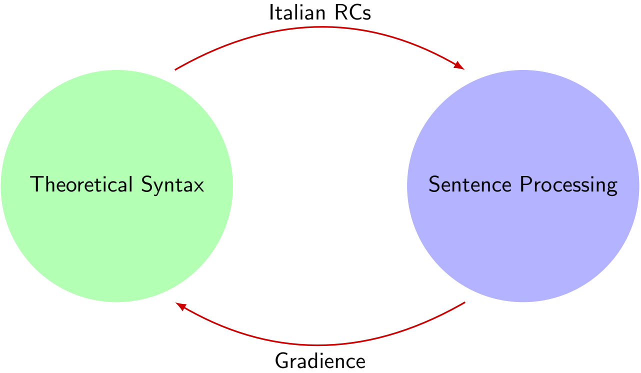
Across Languages

- English, German, Italian, Spanish, French, Korean, Japanese, Mandarin Chinese, Basque, Turkish, Persian, ...

Moving on



Moving on


Italian RCs
Theoretical Syntax
Sentence Processing
Gradience

Acceptability and Grammaticality

1. What do you think that John bought *t*?
2. *What do you wonder whether John bought *t*?

Acceptability and Grammaticality

1. What do you think that John bought *t*?
2. *What do you wonder whether John bought *t*?

> *One way to test the adequacy of a grammar proposed for [language] L is to determine whether or not the sequences that it generates are actually grammatical, i.e., acceptable [...]*
>
> *(Chomsky 1957)*

Acceptability and Grammaticality

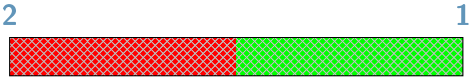
1. What do you think that John bought *t*?
2. *What do you wonder whether John bought *t*?

> *One way to test the adequacy of a grammar proposed for [language] L is to determine whether or not the sequences that it generates are actually grammatical, i.e., acceptable [...]*
>
> *(Chomsky 1957)*

Acceptability judgments $\approx$ Grammaticality judgments

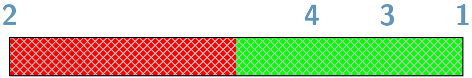
Gradience in Acceptability Judgments

1. What do you think that John bought *t*?
2. *What do you wonder whether John bought *t*?



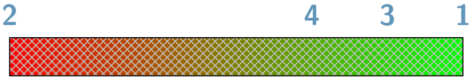
Gradience in Acceptability Judgments

1. What do you think that John bought *t*?
2. *What do you wonder whether John bought *t*?
3. Who *t* thinks that John bought a car?
4. Who *t* wonders whether John bought a car?



Gradience in Acceptability Judgments

1. What do you think that John bought *t*?
2. *What do you wonder whether John bought *t*?
3. Who *t* thinks that John bought a car?
4. Who *t* wonders whether John bought a car?



Gradient Acceptability and Categorical Grammars

Acceptability judgments are not binary but *gradient*:

> *An adequate linguistic theory will have to recognize degrees of grammaticalness [...] there is little doubt that speakers can fairly consistently order new utterances, never previously heard, with respect to their degree of belongingness to the language.*
>
> (Chomsky 1975: 131-132)

But mainstream syntactic theories rely on categorical grammars!

Gradient Acceptability and Categorical Grammars

Acceptability judgments are not binary but *gradient*:

> *An adequate linguistic theory will have to recognize degrees of grammaticalness [...] there is little doubt that speakers can fairly consistently order new utterances, never previously heard, with respect to their degree of belongingness to the language.*
>
> (Chomsky 1975: 131-132)

But mainstream syntactic theories rely on categorical grammars!

(Quantitative) Models of Gradience

Gradient Grammars (Keller 2000; Lau et al. 2014)

- OT-style constraint ranking
- Probabilistic grammars

Extra-grammatical Factors (Chomsky 1975; Schütze 1996)

- Processing effects
  - Plausibility
  - Working memory limitations
  - **But:** few models for quantitative predictions!

Hypothesis

We can use the MG parser to test the relation between categorical grammar, processing difficulty, and gradience!

(Quantitative) Models of Gradience

Gradient Grammars (Keller 2000; Lau et al. 2014)

- OT-style constraint ranking
- Probabilistic grammars

Extra-grammatical Factors (Chomsky 1975; Schütze 1996)

- Processing effects
  - Plausibility
  - Working memory limitations
  - **But:** few models for quantitative predictions!

Hypothesis

We can use the MG parser to test the relation between categorical grammar, processing difficulty, and gradience!

A Proof of Concept: Island Effects

1. What do you think that John bought *t*?
2. What do you wonder whether John bought *t*?
3. Who *t* thinks that John bought a car?
4. Who *t* wonders whether John bought a car?

Results in pairwise comparisons ideal for the MG parser

A Proof of Concept: Island Effects

1. What do you think that John bought *t*?
2. What do you wonder whether John bought *t*?
3. Who *t* thinks that John bought a car?
4. Who *t* wonders whether John bought a car?

Gradience in Islands: Sprouse et al. (2012)

A factorial design for islands effects:

1. Gap Position: Matrix vs. Embedded
2. Structure: Island vs. Non-Island
   (Kluender & Kutas 1993)

Results in pairwise comparisons ideal for the MG parser

A Proof of Concept: Island Effects

1. What do you think that John bought *t*? — Non-Island — Embedded
2. What do you wonder whether John bought *t*? — Island — Embedded
3. Who *t* thinks that John bought a car? — Non-Island — Matrix
4. Who *t* wonders whether John bought a car? — Island — Matrix

Gradience in Islands: Sprouse et al. (2012)

A factorial design for islands effects:

1. GAP POSITION: Matrix vs. Embedded
2. STRUCTURE: Island vs. Non-Island
   (Kluender & Kutas 1993)

Results in pairwise comparisons ideal for the MG parser

A Proof of Concept: Island Effects

1. What do you think that John bought *t*? — Non-Island — Embedded
2. What do you wonder whether John bought *t*? — Island — Embedded
3. Who *t* thinks that John bought a car? — Non-Island — Matrix
4. Who *t* wonders whether John bought a car? — Island — Matrix

Gradience in Islands: Sprouse et al. (2012)

A factorial design for islands effects:

1. Gap Position: Matrix vs. Embedded
2. Structure: Island vs. Non-Island
   (Kluender & Kutas 1993)

Results in pairwise comparisons ideal for the MG parser

Sprouse at al. (2012)

FOUR ISLAND TYPES

Subject islands

- What do you think the speech about *t* interrupted the show about global warming?

Adjunct islands

- What do you laugh if John leaves *t* at the office?

Complex NP islands

- What did you make the claim that John bought *t*?

Whether islands

- What do you wonder whether John bought *t*?

GAP POSITION $\times$ STRUCTURE

1. Matrix vs. Embedded
2. Island vs. Non-Island

Sprouse at al. (2012)

FOUR ISLAND TYPES

Subject islands

- What do you think the speech about *t* interrupted the show about global warming?

Adjunct islands

- What do you laugh if John leaves *t* at the office?

Complex NP islands

- What did you make the claim that John bought *t*?

Whether islands

- What do you wonder whether John bought *t*?

GAP POSITION $\times$ STRUCTURE

1. Matrix vs. Embedded
2. Island vs. Non-Island

Modeling Results (De Santo 2020)

Island Type	Sprouse et al. (2012)			MG Parser
Subj. Island 1	Subj. — Non Isl.	>	Obj. — Non Isl.	✓
	Subj. — Non Isl.	>	Obj. — Isl.	✓
	Subj. — Non Isl.	>	Subj. — Isl.	✓
	Obj. — Non Isl.	>	Obj. — Isl.	✓
	Obj. — Non Isl.	>	Subj. — Isl.	✓
	Obj. — Isl.	>	Subj. — Isl.	×
Subj. Island 2	Matrix — Non Isl.	>	Emb. — Non Isl.	✓
	Matrix — Non Isl.	>	Matrix — Isl.	✓
	Matrix — Non Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Matrix — Isl.	✓
	Emb. — Non Isl.	>	Emb. — Isl.	✓
Adj. Island	Matrix — Non Isl.	>	Emb. — Non Isl.	✓
	Matrix — Non Isl.	>	Matrix — Isl.	✓
	Matrix — Non Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Matrix — Isl.	✓
	Emb. — Non Isl.	>	Emb. — Isl.	✓
CNP Island	Matrix — Non Isl.	>	Emb. — Non Isl.	✓
	Matrix — Non Isl.	=	Matrix — Isl.	✓
	Matrix — Non Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Matrix — Isl.	✓
	Emb. — Non Isl.	>	Emb. — Isl.	✓

Modeling Results (De Santo 2020)

Island Type	Sprouse et al. (2012)			MG Parser
Subj. Island 1	Subj. — Non Isl.	>	Obj. — Non Isl.	✓
	Subj. — Non Isl.	>	Obj. — Isl.	✓
	Subj. — Non Isl.	>	Subj. — Isl.	✓
	Obj. — Non Isl.	>	Obj. — Isl.	✓
	Obj. — Non Isl.	>	Subj. — Isl.	✓
	Obj. — Isl.	>	**Subj. — Isl.**	×
Subj. Island 2	Matrix — Non Isl.	>	Emb. — Non Isl.	✓
	Matrix — Non Isl.	>	Matrix — Isl.	✓
	Matrix — Non Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Matrix — Isl.	✓
	Emb. — Non Isl.	>	Emb. — Isl.	✓
Adj. Island	Matrix — Non Isl.	>	Emb. — Non Isl.	✓
	Matrix — Non Isl.	>	Matrix — Isl.	✓
	Matrix — Non Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Matrix — Isl.	✓
	Emb. — Non Isl.	>	Emb. — Isl.	✓
CNP Island	Matrix — Non Isl.	>	Emb. — Non Isl.	✓
	Matrix — Non Isl.	=	Matrix — Isl.	✓
	Matrix — Non Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Emb. — Isl.	✓
	Matrix — Isl.	>	Matrix — Isl.	✓
	Emb. — Non Isl.	>	Emb. — Isl.	✓

TL;DR

Success in all cases but one!

Subject Island: Case 1

(5)
a. What do you think the speech interrupted *t*? — Obj — Non Island
b. What do you think *t* interrupted the show? — Subj — Non Island
c. What do you think the speech about global warming interrupted the show about *t*? — Obj — Island
d. What do you think the speech about *t* interrupted the show about global warming? — Subj — Island

Sprouse et al. (2012)			MG Parser
Subj. — Non Isl.	>	Obj. — Non Isl.	✓
Subj. — Non Isl.	>	Obj. — Isl.	✓
Subj. — Non Isl.	>	Subj. — Isl.	✓
Obj. — Non Isl.	>	Obj. — Isl.	✓
Obj. — Non Isl.	>	Subj. — Isl.	✓
Obj. — Isl.	>	Subj. — Isl.	×

Clause Type	MaxT	SumS
Obj./Non Island	14/*do*	19
Subj./Non Island	11/*do*	14
Obj./Island	23/*T2*	22
Subj./Island	15/*do*	20

Subject Island: Case 1

(5)
a. What do you think the speech interrupted *t*? — Obj — Non Island
b. What do you think *t* interrupted the show? — Subj — Non Island
c. What do you think the speech about global warming interrupted the show about *t*? — Obj — Island
d. * What do you think the speech about *t* interrupted the show about global warming? — Subj — Island

Sprouse et al. (2012)			MG Parser
Subj. — Non Isl.	>	Obj. — Non Isl.	✓
Subj. — Non Isl.	>	Obj. — Isl.	✓
Subj. — Non Isl.	>	Subj. — Isl.	✓
Obj. — Non Isl.	>	Obj. — Isl.	✓
Obj. — Non Isl.	>	Subj. — Isl.	✓
Obj. — Isl.	>	Subj. — Isl.	×

Clause Type	MaxT	SumS
Obj./Non Island	14/*do*	19
Subj./Non Island	11/*do*	14
Obj./Island	23/*T2*	22
Subj./Island	15/*do*	20

Subject Island: Case 2

(6) a. Who *t* thinks the speech interrupted the primetime TV show?
Matrix — Non Island

b. What do you think *t* interrupted the primetime TV show?
Emb. — Non Island

c. Who *t* thinks the speech about global warming interrupted the primetime TV show?
Matrix — Island

d. What do you think the speech about *t* interrupted the primetime TV show?
Emb. — Island

Sprouse et al. (2012)			MG Parser
Matrix — Non Isl.	>	Emb. — Non Isl.	✓
Matrix — Non Isl.	>	Matrix — Isl.	✓
Matrix — Non Isl.	>	Emb. — Isl.	✓
Matrix — Isl.	>	Emb. — Isl.	✓
Matrix — Isl.	>	Matrix — Isl.	✓
Emb. — Non Isl.	>	Emb. — Isl.	✓

Clause Type	MaxT	SumS
Matrix — Non Isl.	$5/C$	9
Emb. — Non Isl.	$11/do$	14
Matrix — Isl.	$11/T_{RC}$	9
Emb. — Isl.	$17/T_{RC}$	20

Summary

Gradience from a categorical MG grammar?

- The **first** (quantitative) model of this kind!
- Overall, a success! ⇒ **just** from structural differences!
- Outlier is expected assuming grammaticalized constraints.

The tip of the iceberg!

- Modulate range of dependencies
- Other examples of gradience
- Cognitive vs. grammatical constraints? (Ferrara-Boston 2012)
- Syntactic constraints ∼ pruning the parsing space (Stabler 2013, Graf & De Santo 2020)
- Economy principles (De Santo & Lee 2022)

Summary

Gradience from a categorical MG grammar?

- The **first** (quantitative) model of this kind!
- Overall, a success! ⇒ **just** from structural differences!
- Outlier is expected assuming grammaticalized constraints.

The tip of the iceberg!

- Modulate range of dependencies
- Other examples of gradience
- Cognitive vs. grammatical constraints? (Ferrara-Boston 2012)
- Syntactic constraints ∼ pruning the parsing space (Stabler 2013, Graf & De Santo 2020)
- Economy principles (De Santo & Lee 2022)

Summary

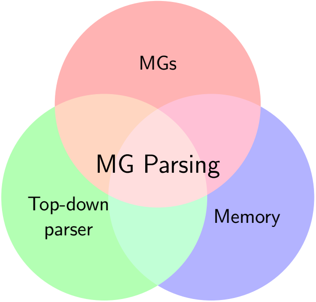
Gradience from a categorical MG grammar?

- The **first** (quantitative) model of this kind!
- Overall, a success! $\Rightarrow$ **just** from structural differences!
- Outlier is expected assuming grammaticalized constraints.

The tip of the iceberg!

- Modulate range of dependencies
- Other examples of gradience
- Cognitive vs. grammatical constraints? (Ferrara-Boston 2012)
- Syntactic constraints $\sim$ pruning the parsing space (Stabler 2013, Graf & De Santo 2020)
- Economy principles (De Santo & Lee 2022)

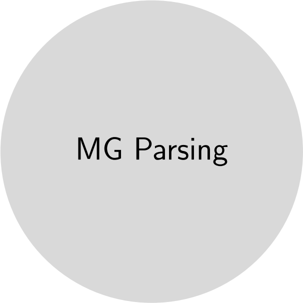
From the Trees (back) to the Forest



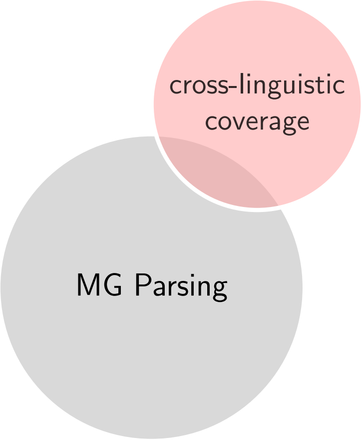
Within the program of research proposed here, joint work by linguists, computer scientists, and psychologists could lead to a deeper scientific understanding of the role of language in cognition.

(Bresnan 1978: pg. 59)

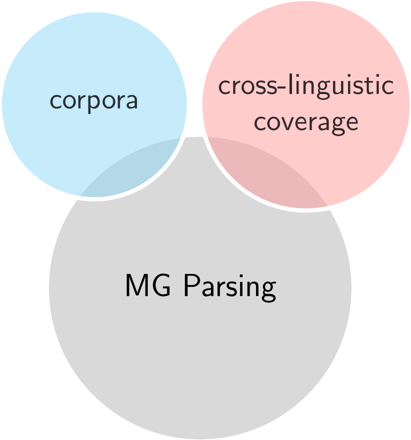
Looking Ahead: A Collaborative Enterprise



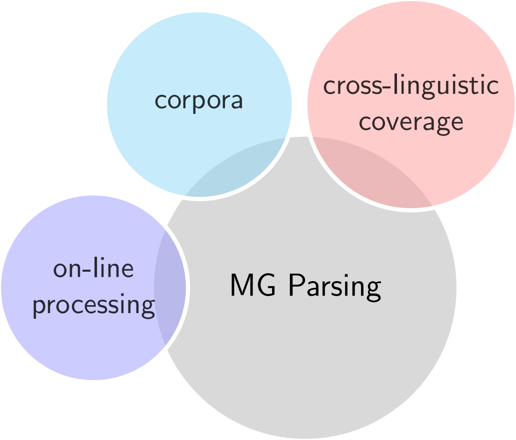
Looking Ahead: A Collaborative Enterprise



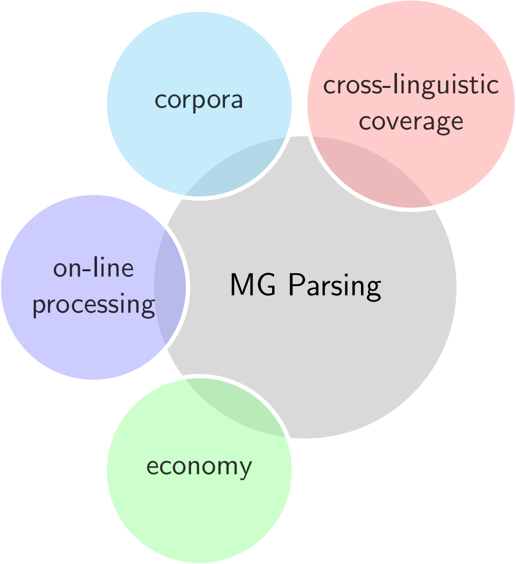
Looking Ahead: A Collaborative Enterprise



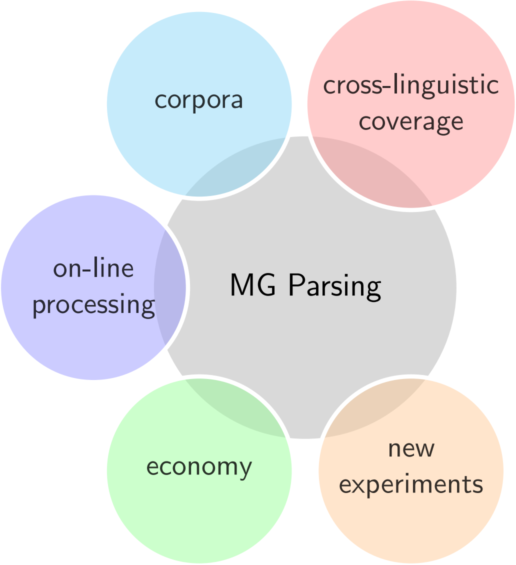
Looking Ahead: A Collaborative Enterprise



Looking Ahead: A Collaborative Enterprise



Looking Ahead: A Collaborative Enterprise



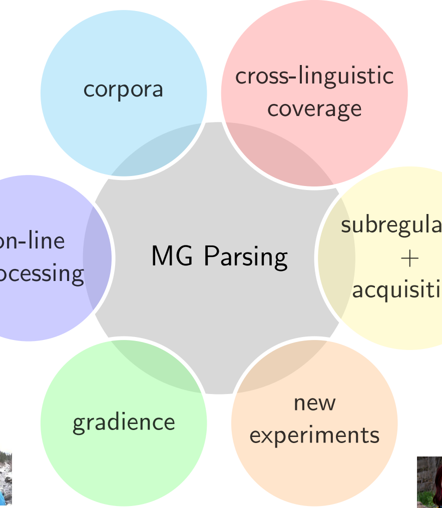
Looking Ahead: A Collaborative Enterprise

Looking Ahead: A Collaborative Enterprise















Looking Ahead: A Collaborative Enterprise

Selected References I

1 **Chomsky, N.** (1995). The minimalist program. Cambridge, Mass.: MIT Press.

2 **De Santo, A.** (2019). Testing a Minimalist gram- mar parser on Italian relative clause asymmetries. In *Proceedings of CMCL 2019*, June 6 2019, Minneapolis, Minnesota.

3 **De Santo, A.** (2020). MG Parsing as a Model of Gradient Acceptability in Syntactic Islands. (To appear) In *Proceedings of SCiL 2020*, Jan 2-5, New Orleans.

4 **De Santo, A.** and Shafiei, N. (2019). On the structure of relative clauses in Persian: Evidence from computational modeling and processing effects. *Talk at the NACIL2*, April 19-21 2019, University of Arizona.

5 **De Santo, A.** and Lee, So Young. (2022a). Evaluating Structural Economy Claims in Relative Clause Attachment. In *Proceedings of SCiL 2022*.

6 **De Santo, A.** and Lee, So Young. (2022b). Pseudo-relative clause effects on the online processing of Italian relative clause attachment. Poster at *HSP 2022*.

7 **Graf, T.** and Monette, J. and Zhang, C. (2017). Relative Clauses as a Benchmark for Minimalist Parsing. *Journal of Language Modelling*.

8 **Grillo, N., & Costa, J.** (2014). A novel argument for the universality of parsing principles. Cognition, 133(1), 156-187.

9 **Kobele, G.M.**, Gerth S., and Hale. J. (2012). Memory resource allocation in top-down minimalist parsing. In *Formal Grammar*, pages 32–51. Springer.

10 **Stabler, E.P.** (2013). Bayesian, minimalist, incremental syntactic analysis. *Topics in Cognitive Science* 5:611–633.

11 **Stabler, E.P.** (1997). Derivational minimalism. In Logical aspects of computational linguistics, ed.

Christian Retore, volume 1328 of Lecture Notes in Computer Science, 68–95. Berlin: Springer.

Appendix

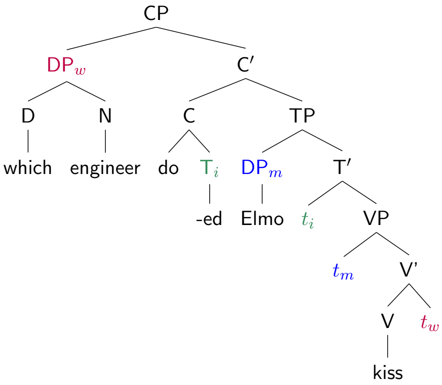
Why MGs?

1. Vast analytical coverage
   - MGs handle virtually all analyses in the generative literature
2. Centrality of derivation trees
   - MGs can be viewed as CFGs with a more complicated mapping from trees to strings
3. Simple parsing algorithms
   - Variant of a recursive descent parser for CFGs
     ⇒ cf. TAG (Rambow & Joshi, 1995; Demberg, 2008)

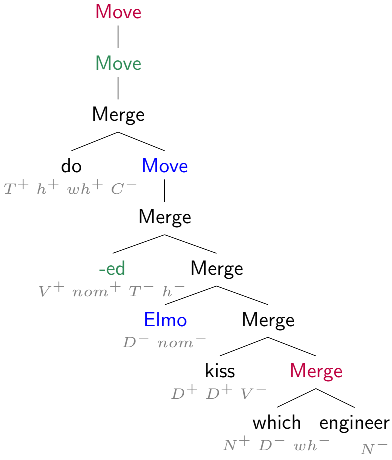
Some Important Properties of MGs

- MGs are weakly equivalent to MCFGs and thus mildly context-sensitive. (Harkema 2001, Michaelis 2001)
- But we can decompose them into two finite-state components: (Michaelis et al. 2001, Kobele et al. 2007, Monnich 2006)
  - a regular language of well-formed derivation trees
  - an MSO-definable mapping from derivations to phrase structure trees
- **Remember:** Every regular tree language can be re-encoded as a CFG (with more fine-grained non-terminal labels). (Thatcher 1967)

Fully Specified Derivation Trees



Phrase Structure Tree



Derivation Tree

Technical Fertility of MGs

MGs can accommodate the full syntactic toolbox:

- sidewards movement (Stabler, 2006; Graf 2013)
- affix hopping (Graf 2012; Graf2013)
- clustering movement (Gartner & Michaelis 2010)
- tucking in (Graf 2013)
- ATB movement (Kobele 2008)
- copy movement (Kobele 2006)
- extraposition (Hunter &Frank 2014)
- Late Merge (Kobele 2010; Graf 2014)
- Agree (Kobele 2011; Graf 2011)
- adjunction (Fowlie 2013; Hunter 2015)
- TAG-style adjunction (Graf 2012)

Why These Metrics?

- These complexity metrics are all related to **storage cost**
  (cf. Gibson, 1998)
- We could implement alternative ones
  (cf. Ferrara-Boston, 2012)
  - number of bounding nodes / phases
  - surprisal
  - feature intervention
  - status of discourse referents
  - integration, retrieval, ...
- We want to keep the model **simple** (but not **trivial**):
  - Tenure and Size only refer to the geometry of the derivation
  - they are sensitive the specifics of tree-traversal
    (cf. node-count; Hale, 2001)

Why These Metrics?

- These complexity metrics are all related to **storage cost** (cf. Gibson, 1998)
- We could implement alternative ones (cf. Ferrara-Boston, 2012)
  - number of bounding nodes / phases
  - surprisal
  - feature intervention
  - status of discourse referents
  - integration, retrieval, ...
- We want to keep the model **simple** (but not **trivial**):
  - Tenure and Size only refer to the geometry of the derivation
  - they are sensitive the specifics of tree-traversal (cf. node-count; Hale, 2001)

Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1	CP is conjectured
step 2	CP expands to C′
step 3	C′ expands to does and TP
step 4	TP expands to Salem and T′
step 5	T′ expands to T and VP
step 6	VP expands to mock and who
step 7	who is found
step 8	does is found
step 9	Salem is found
step 10	T is found
step 11	mock is found

Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

[1]CP

- • Who • does • Salem • T • mock

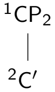
step 1 *CP* is conjectured
step 2 *CP* expands to *C'*
step 3 *C'* expands to *does* and *TP*
step 4 *TP* expands to *Salem* and *T'*
step 5 *T'* expands to *T* and *VP*
step 6 *VP* expands to *mock* and *who*
step 7 *who* is found
step 8 *does* is found
step 9 *Salem* is found
step 10 *T* is found
step 11 *mock* is found

Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock



step 1 *CP* is conjectured
step 2 *CP* expands to *C'*
step 3 *C'* expands to *does* and *TP*
step 4 *TP* expands to *Salem* and *T'*
step 5 *T'* expands to *T* and *VP*
step 6 *VP* expands to *mock* and *who*
step 7 *who* is found
step 8 *does* is found
step 9 *Salem* is found
step 10 *T* is found
step 11 *mock* is found

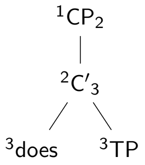
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1 *CP* is conjectured
step 2 *CP* expands to *C'*
step 3 *C'* expands to *does* and *TP*
step 4 *TP* expands to *Salem* and *T'*
step 5 *T'* expands to *T* and *VP*
step 6 *VP* expands to *mock* and *who*
step 7 *who* is found
step 8 *does* is found
step 9 *Salem* is found
step 10 *T* is found
step 11 *mock* is found



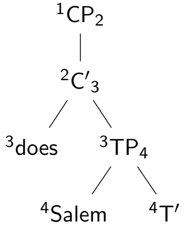
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1 *CP* is conjectured
step 2 *CP* expands to *C'*
step 3 *C'* expands to *does* and *TP*
step 4 *TP* expands to *Salem* and *T'*
step 5 *T'* expands to *T* and *VP*
step 6 *VP* expands to *mock* and *who*
step 7 *who* is found
step 8 *does* is found
step 9 *Salem* is found
step 10 *T* is found
step 11 *mock* is found



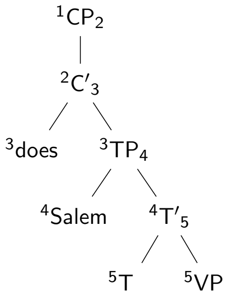
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1 *CP* is conjectured
step 2 *CP* expands to *C'*
step 3 *C'* expands to *does* and *TP*
step 4 *TP* expands to *Salem* and *T'*
step 5 *T'* expands to *T* and *VP*
step 6 *VP* expands to *mock* and *who*
step 7 *who* is found
step 8 *does* is found
step 9 *Salem* is found
step 10 *T* is found
step 11 *mock* is found



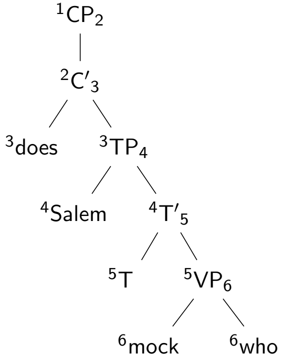
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1 *CP* is conjectured
step 2 *CP* expands to *C'*
step 3 *C'* expands to *does* and *TP*
step 4 *TP* expands to *Salem* and *T'*
step 5 *T'* expands to *T* and *VP*
step 6 *VP* expands to *mock* and *who*
step 7 *who* is found
step 8 *does* is found
step 9 *Salem* is found
step 10 *T* is found
step 11 *mock* is found



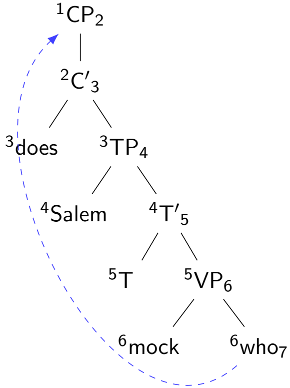
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1	*CP* is conjectured
step 2	*CP* expands to *C'*
step 3	*C'* expands to *does* and *TP*
step 4	*TP* expands to *Salem* and *T'*
step 5	*T'* expands to *T* and *VP*
step 6	*VP* expands to *mock* and *who*
step 7	*who* is found
step 8	*does* is found
step 9	*Salem* is found
step 10	*T* is found
step 11	*mock* is found



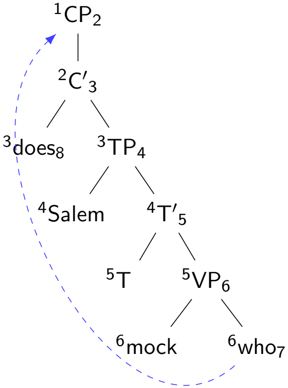
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1 *CP* is conjectured
step 2 *CP* expands to *C'*
step 3 *C'* expands to *does* and *TP*
step 4 *TP* expands to *Salem* and *T'*
step 5 *T'* expands to *T* and *VP*
step 6 *VP* expands to *mock* and *who*
step 7 *who* is found
step 8 *does* is found
step 9 *Salem* is found
step 10 *T* is found
step 11 *mock* is found



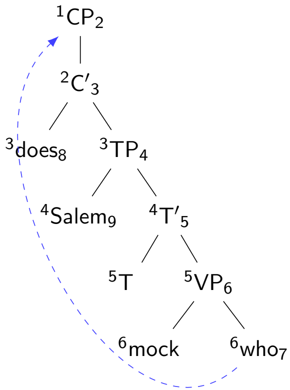
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1 *CP* is conjectured
step 2 *CP* expands to *C'*
step 3 *C'* expands to *does* and *TP*
step 4 *TP* expands to *Salem* and *T'*
step 5 *T'* expands to *T* and *VP*
step 6 *VP* expands to *mock* and *who*
step 7 *who* is found
step 8 *does* is found
step 9 *Salem* is found
step 10 *T* is found
step 11 *mock* is found



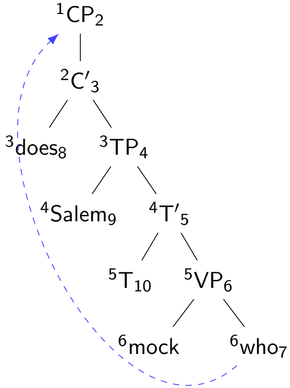
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1	*CP* is conjectured
step 2	*CP* expands to *C'*
step 3	*C'* expands to *does* and *TP*
step 4	*TP* expands to *Salem* and *T'*
step 5	*T'* expands to *T* and *VP*
step 6	*VP* expands to *mock* and *who*
step 7	*who* is found
step 8	*does* is found
step 9	*Salem* is found
step 10	*T* is found
step 11	*mock* is found



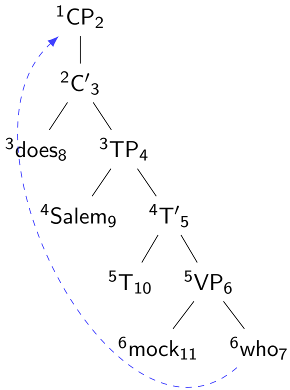
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1	*CP* is conjectured
step 2	*CP* expands to *C'*
step 3	*C'* expands to *does* and *TP*
step 4	*TP* expands to *Salem* and *T'*
step 5	*T'* expands to *T* and *VP*
step 6	*VP* expands to *mock* and *who*
step 7	*who* is found
step 8	*does* is found
step 9	*Salem* is found
step 10	*T* is found
step 11	*mock* is found



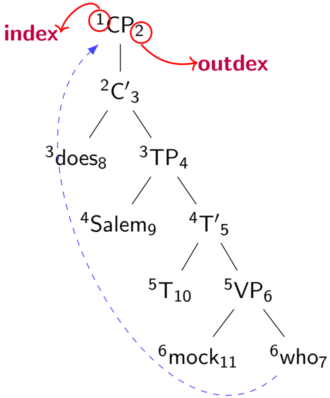
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1	*CP* is conjectured
step 2	*CP* expands to *C'*
step 3	*C'* expands to *does* and *TP*
step 4	*TP* expands to *Salem* and *T'*
step 5	*T'* expands to *T* and *VP*
step 6	*VP* expands to *mock* and *who*
step 7	*who* is found
step 8	*does* is found
step 9	*Salem* is found
step 10	*T* is found
step 11	*mock* is found



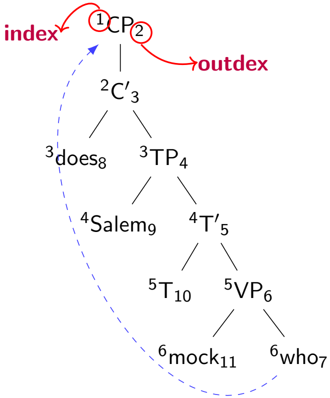
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

- • Who • does • Salem • T • mock

step 1	*CP* is conjectured
step 2	*CP* expands to *C'*
step 3	*C'* expands to *does* and *TP*
step 4	*TP* expands to *Salem* and *T'*
step 5	*T'* expands to *T* and *VP*
step 6	*VP* expands to *mock* and *who*
step 7	*who* is found
step 8	*does* is found
step 9	*Salem* is found
step 10	*T* is found
step 11	*mock* is found



Index and Outdex are our connection to memory!

Memory-Based Complexity Metrics

- **Memory usage**
  (Gibson 1998, Kobele et al. 2012):

Tenure: How long a node is kept in memory

Size: How much information is stored in a node
$\Rightarrow$ Intuitively, the length of its movement dependency!

- These can be formalized into **complexity metrics**

MaxTenure: $max(\{\text{tenure-of}(n) | n \text{ a node of the tree}\})$

SumSize: $\sum_{m \in M} size(m)$

Ranked: $\langle MaxTenure, SumSize \rangle$

Greg Kobele

Sabrina Gerth

John Hale

Memory-Based Complexity Metrics

- **Memory usage** (Gibson 1998, Kobele et al. 2012):

Tenure: How long a node is kept in memory
Size: How much information is stored in a node
$\Rightarrow$ Intuitively, the length of its movement dependency!

- These can be formalized into **complexity metrics**

MaxTenure: $max(\{\text{tenure-of}(n) | n \text{ a node of the tree}\})$
SumSize: $\sum_{m \in M} size(m)$
Ranked: $\langle MaxTenure, SumSize \rangle$

Greg Kobele

Sabrina Gerth

John Hale

Memory-Based Complexity Metrics

- **Memory usage** (Gibson 1998, Kobele et al. 2012):

Tenure: How long a node is kept in memory
Size: How much information is stored in a node
$\Rightarrow$ Intuitively, the length of its movement dependency!

- These can be formalized into **complexity metrics**

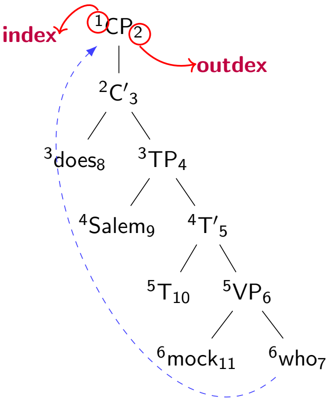
MaxTenure: $max(\{\text{tenure-of}(n) | n \text{ a node of the tree}\})$
SumSize: $\sum_{m \in M} size(m)$
Ranked: $\langle MaxTenure, SumSize \rangle$

Greg Kobele

Sabrina Gerth

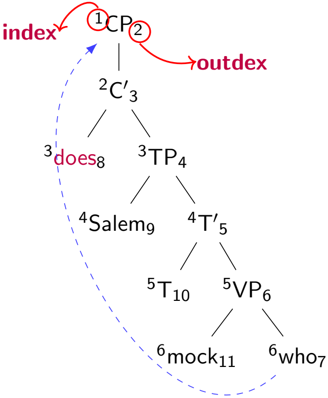
John Hale

Computing Metrics: An Example



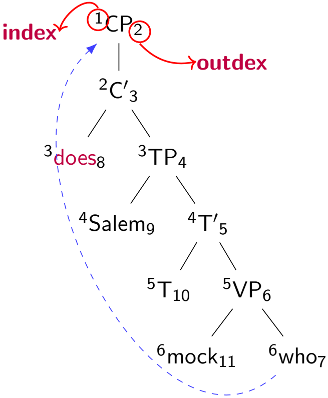
Tenure how long a node is kept in memory

Computing Metrics: An Example



Tenure how long a node is kept in memory
Tenure(*does*) $= 8 - 3 = 5$

Computing Metrics: An Example



Tenure how long a node is kept in memory

Tenure($does$**)** $= 8 - 3 = 5$

MaxTenure $= max\{$**Tenure**$(does),$ **Tenure**$(Salem), \dots\} = 5$

Automatizing Helps!

mgproc: A Python Package for MG Processing Research

This is a collection of Python3 scripts to facilitate the investigation of human processing from the perspective of Minimalist grammars (MGs).

Background

MGs were developed in Stabler (1997) as a formalization of Chomsky's Minimalist program. A top-down parser for MGs is defined in Stabler (2013) and has been implemented in a number of languages. A number of subsequent works have successfully used this parser to make predictions about relative difficulty in sentence processing. Good starting points with a review of the previous literature are Gerth (2015) and Graf et al. (to appear).

- Gerth, Sabrina: Memory Limitations in Sentence Comprehension
- Graf, Thomas, James Monette, and Chong Zhang (to appear): Relative Clauses as a Benchmark for Minimalist Parsing (link to be added soon)
- Stabler, Edward (1997): Derivational Minimalism
- Stabler, Edward (2013): Two Models of Minimalist, Incremental Syntactic Analysis

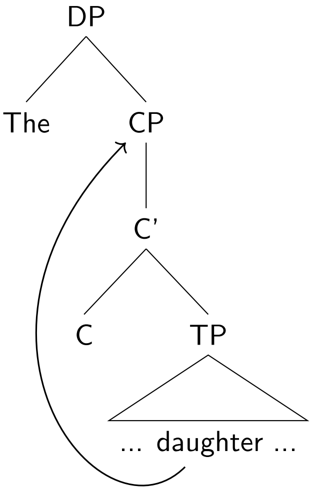
Quick Start Guide

With *mgproc* you can easily compare MG derivation trees with respect to thousands of complexity m processing. The scripts integrate well with a LaTeX-centric workflow, following the ideal of OpenScie publication form a cohesive unit. Usually a parsed derivation tree is specified by four files. Assuming `foo`, we have:

- Open source $\Rightarrow$ in prep. for *Journal of Open Source Software*
- User-friendly!
- Easy to modify!

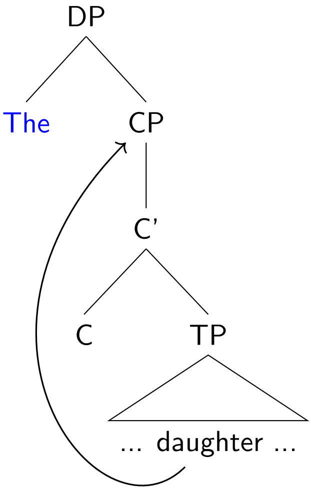
Kayne's Promotion Analysis (Kayne 1994)

- RC is selected by an external D^0
- the RC head is a nominal constituent
- the RC head raises from its base position to [Spec, CP]



$[_{DP}$ The $[_{CP}$ daughter$_i$ [that t_i was on the balcony]]]

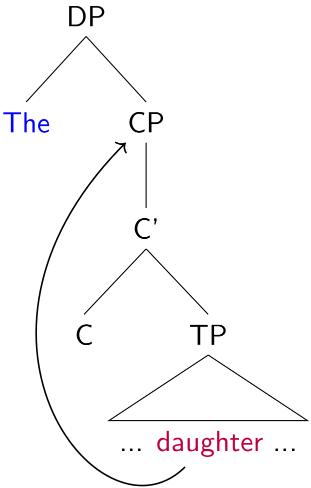
Kayne's Promotion Analysis (Kayne 1994)



- RC is selected by an external D^0
- the RC head is a nominal constituent
- the RC head raises from its base position to [Spec, CP]

[$_{DP}$ The [$_{CP}$ daughter$_i$ [that t_i was on the balcony]]]

Kayne's Promotion Analysis (Kayne 1994)

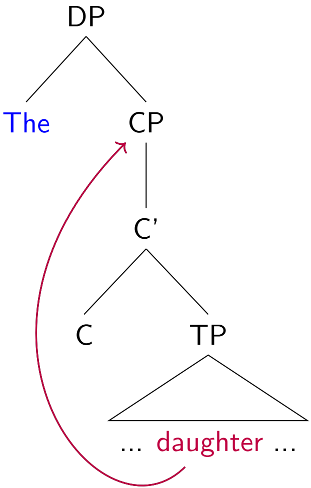


- RC is selected by an external D^0
- the RC head is a nominal constituent
- the RC head raises from its base position to [Spec, CP]

$[_{DP}$ The $[_{CP}$ daughter$_i$ [that t_i was on the balcony]]]

Kayne's Promotion Analysis (Kayne 1994)

- RC is selected by an external D^0
- the RC head is a nominal constituent
- the RC head raises from its base position to [Spec, CP]



$[_{DP}$ The $[_{CP}$ daughter$_i$ [that t_i was on the balcony]]]

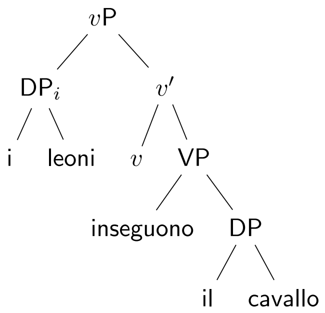
Postverbal Subjects (Belletti & Leonini 2004)

(7) Inseguono il cavallo i leoni
Chase the horse the lions
"The lions chase the horse"

- the subject DP raises to Spec, FocP
- The whole vP raises to Spec,TopP

Technical details!

- an expletive *pro* is base generated in Spec,TP



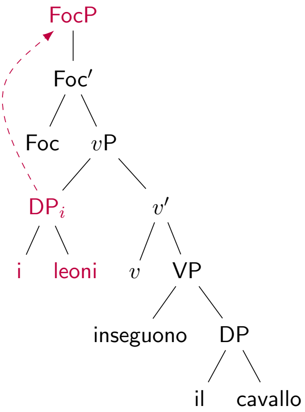
Postverbal Subjects (Belletti & Leonini 2004)

(8) Inseguono il cavallo i leoni
Chase the horse the lions
"The lions chase the horse"

- the subject DP raises to Spec, FocP
- The whole vP raises to Spec,TopP

Technical details!

- an expletive *pro* is base generated in Spec,TP



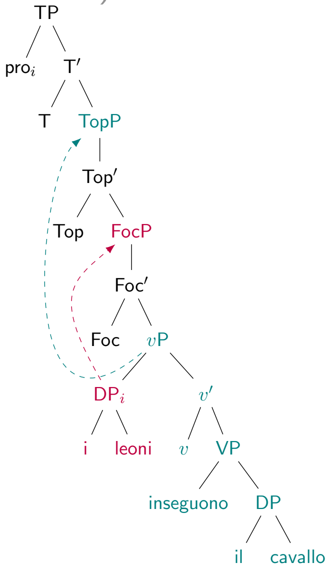
Postverbal Subjects (Belletti & Leonini 2004)

(9) Inseguono il cavallo i leoni
Chase the horse the lions
"The lions chase the horse"

- the subject DP raises to Spec, FocP
- The whole vP raises to Spec,TopP

Technical details!

- an expletive *pro* is base generated in Spec,TP



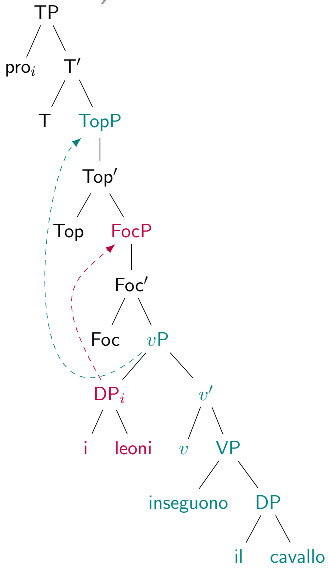
Postverbal Subjects (Belletti & Leonini 2004)

(9) Inseguono il cavallo i leoni
Chase the horse the lions
"The lions chase the horse"

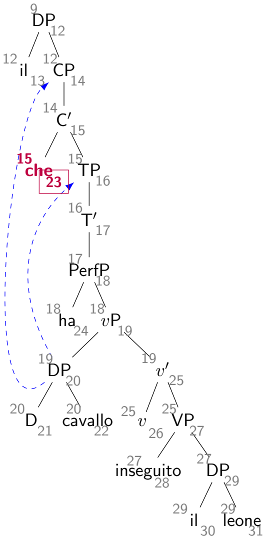
- the subject DP raises to Spec, FocP
- The whole vP raises to Spec,TopP

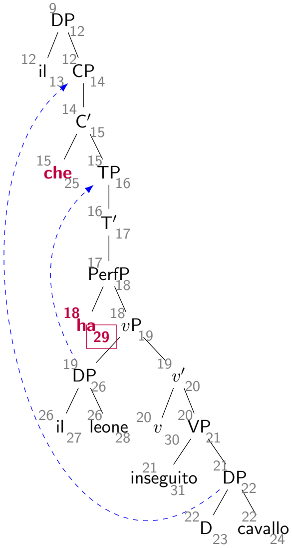
Technical details!

- an expletive *pro* is base generated in Spec,TP

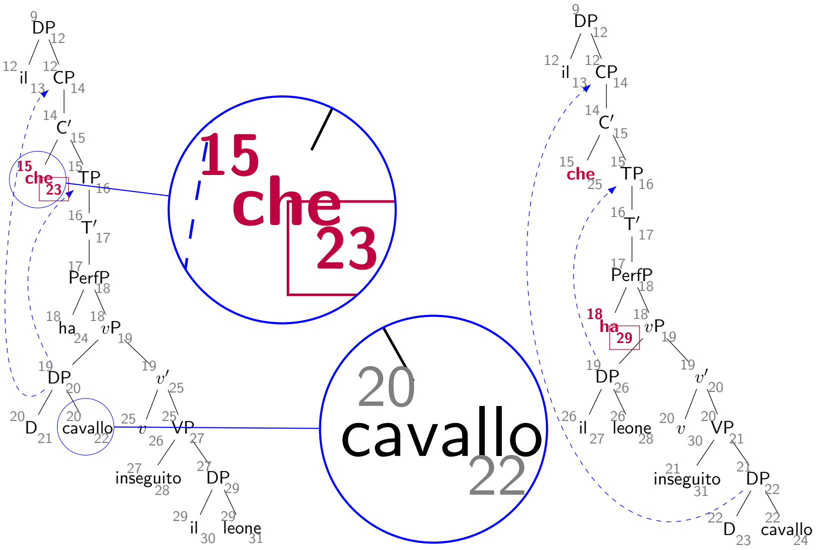


Results: SRC > ORC

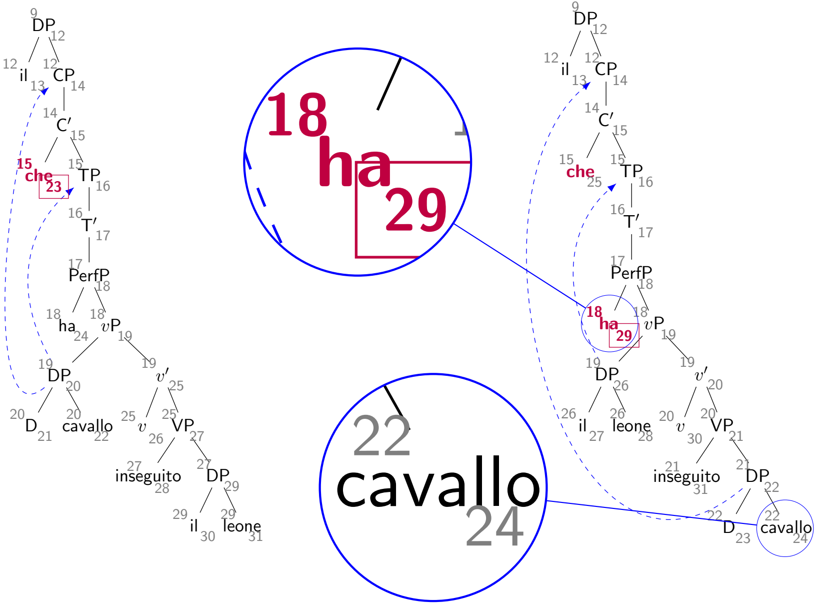




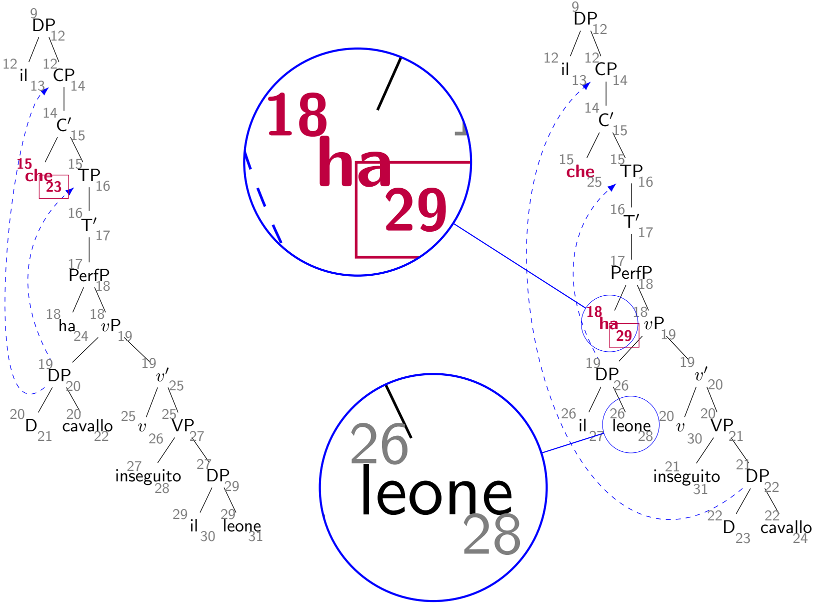
Results: SRC > ORC



Results: SRC > ORC



Results: SRC > ORC



Italian Subjects: Probing the Results

Clause Type	**MaxT**	**SumS**
obj. SRC	8/*che*	18
obj. ORC	11/*ha*	24
obj. ORCp	16/*Foc*	31
subj. SRC	21/v'	37
subj. ORC	21/v'	44
subj. ORCp	28/v'	56
matrix SVO	3/*ha*/v'	7
matrix VOS	7/*Top*/*Foc*	11
VS unacc	2/vP	3
VS unerg	7/*Top*/*Foc*	11

Table: Summary of MaxT (*value*/*node*) and SumS by construction. Obj. and subj. indicate the landing site of the RC head in the matrix clause.

Postverbal Asymmetries: Possible Accounts?

SRC > ORC

- DLT, active-filler strategy, Competition model, ...

ORC > ORCp

- more problematic (e.g., for DLT)
- can be explained by
  1. economy of gap prediction + structural re-analysis;
  2. intervention effects + featural Relativized Minimality

Can we give a purely structural account?

Postverbal Asymmetries: Possible Accounts?

SRC > ORC

- DLT, active-filler strategy, Competition model, ...

ORC > ORCp

- more problematic (e.g., for DLT)
- can be explained by
  1. economy of gap prediction + structural re-analysis;
  2. intervention effects + featural Relativized Minimality

Can we give a purely structural account?

Postverbal Asymmetries: Possible Accounts?

SRC > ORC

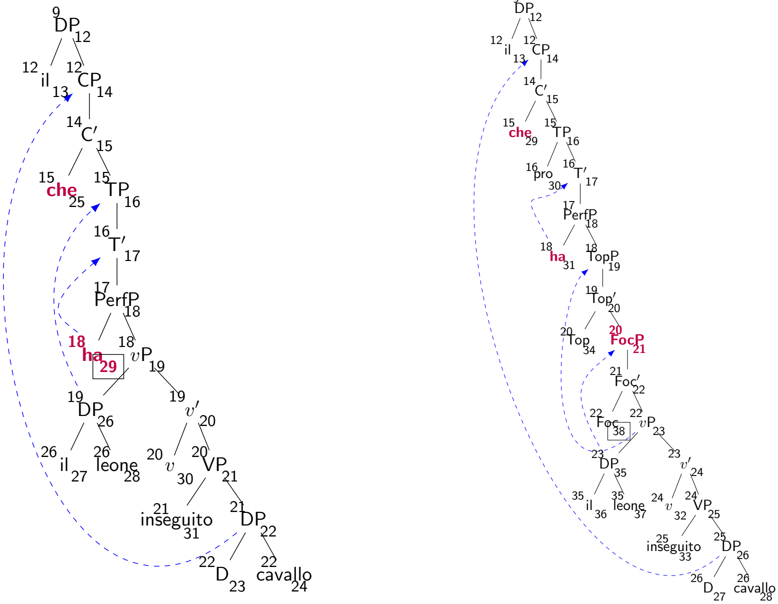
- DLT, active-filler strategy, Competition model, ...

ORC > ORCp

- more problematic (e.g., for DLT)
- can be explained by
  1. economy of gap prediction + structural re-analysis;
  2. intervention effects + featural Relativized Minimality

Can we give a purely structural account?

Results: ORC > ORCp



Additional Constructions

- Ambiguity in Matrix Clauses

(10) Ha chiamato Gio
Has called Giovanni
a. "He/she/it called Gio" **SVO**
b. "Gio called" **VS**

- Unaccusatives vs. Unergatives

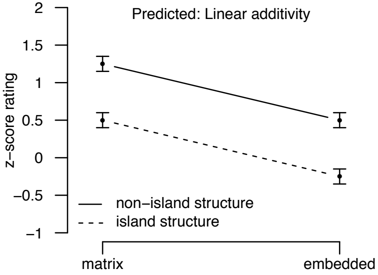
(11) È arrivato Gio
Is arrived Gio
"Gio arrived" **Unaccusative**

(12) Ha corso Gio
Has ran Gio
"Gio ran" **Unergative**

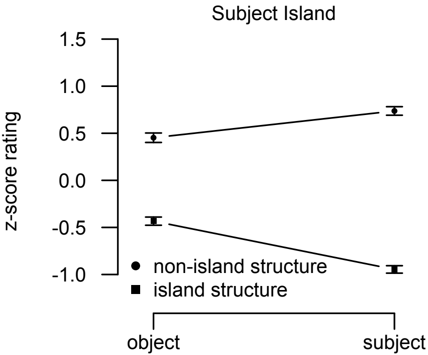
Gradience in Islands

A factorial design for islands effect:

- GAP POSITION × STRUCTURE



Deriving Pairwise Comparisons



- Subj — Non Island > Obj — Non Island
- Subj — Non Island > Obj — Island
- Subj — Non Island > Subj — Island
- etc.

A Caveat on Island Effects

The Goal

Can gradience in acceptability judgments arise from a categorical grammar due to processing factors?

- Sprouse et al.'s (2012) design is ideal for the MG model.

But I am not interested in island effects per se:

- Islands: grammatical or processing effects? (Hofmeister et al., 2012a; Sprouse et al., 2012a,b)
  - hence, not modeling super-additivity
  - spoilers: maybe we get some insights?
- Islands: syntax or semantics? (Truswell, 2011; Kush et al., 2018; Matchin et al., 2018)

A Caveat on Island Effects

The Goal

Can gradience in acceptability judgments arise from a categorical grammar due to processing factors?

- Sprouse et al.'s (2012) design is ideal for the MG model.

But I am not interested in island effects per se:

- Islands: grammatical or processing effects?
  (Hofmeister et al., 2012a; Sprouse et al., 2012a,b)
  - hence, not modeling super-additivity
  - spoilers: maybe we get some insights?
- Islands: syntax or semantics?
  (Truswell, 2011; Kush et al., 2018; Matchin et al., 2018)

A Caveat on Island Effects

The Goal

Can gradience in acceptability judgments arise from a categorical grammar due to processing factors?

- Sprouse et al.'s (2012) design is ideal for the MG model.

But I am not interested in island effects per se:

- Islands: grammatical or processing effects? (Hofmeister et al., 2012a; Sprouse et al., 2012a,b)
  - hence, not modeling super-additivity
  - spoilers: maybe we get some insights?
- Islands: syntax or semantics? (Truswell, 2011; Kush et al., 2018; Matchin et al., 2018)

A Caveat on Island Effects

The Goal

Can gradience in acceptability judgments arise from a categorical grammar due to processing factors?

- Sprouse et al.'s (2012) design is ideal for the MG model.

But I am not interested in island effects per se:

- Islands: grammatical or processing effects? (Hofmeister et al., 2012a; Sprouse et al., 2012a,b)
  - hence, not modeling super-additivity
  - spoilers: maybe we get some insights?
- Islands: syntax or semantics? (Truswell, 2011; Kush et al., 2018; Matchin et al., 2018)

Models of Gradience

(At least two) theories of gradience:

- Gradience incorporated in the grammar
  (Keller 2000; Featherston 2005; Lau et al. 2014)
- Gradience due to extra-grammatical factors
  (Chomsky 1975; Schütze 1996)

The contribution of formal models?

Quantify what each approach needs to account for the data:

- Additional syntactic assumptions
- Additional complexity in acquisition, processing strategies, etc.

Models of Gradience

(At least two) theories of gradience:

- Gradience incorporated in the grammar (Keller 2000; Featherston 2005; Lau et al. 2014)
- Gradience due to extra-grammatical factors (Chomsky 1975; Schütze 1996)

The contribution of formal models?

Quantify what each approach needs to account for the data:

- Additional syntactic assumptions
- Additional complexity in acquisition, processing strategies, etc.

Subject Islands

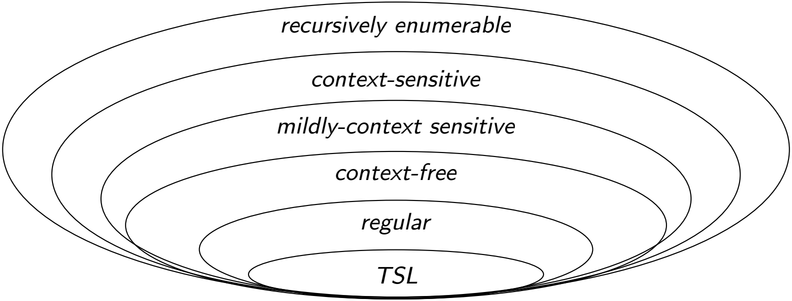
Case 1:

(13)
a. What do you think the speech interrupted *t*? — Obj — Non Island
b. What do you think *t* interrupted the show? — Subj — Non Island
c. What do you think the speech about global warming interrupted the show about *t*? — Obj — Island
d. What do you think the speech about *t* interrupted the show about global warming? — Subj — Island

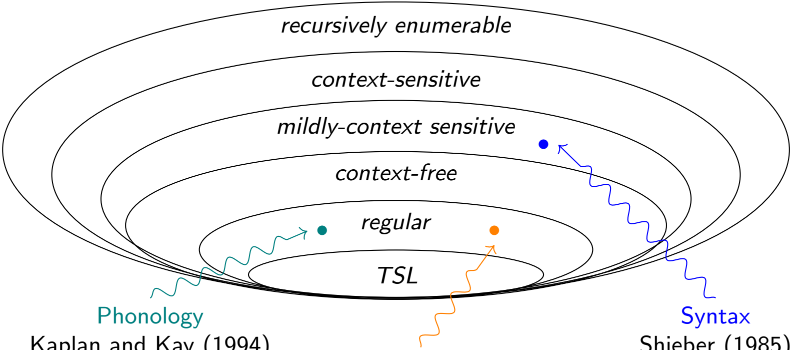
Case 2:

(14)
a. Who *t* thinks the speech interrupted the primetime TV show? — Matrix — Non Island
b. What do you think *t* interrupted the primetime TV show? — Emb. — Non Island
c. Who *t* thinks the speech about global warming interrupted the primetime TV show? — Matrix — Island
d. What do you think the speech about *t* interrupted the primetime TV show? — Emb. — Island

Subregular Complexity



Subregular Complexity

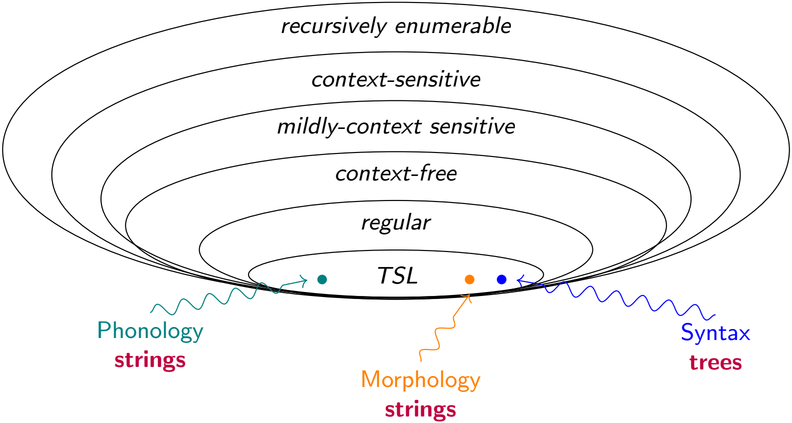


Phonology
Kaplan and Kay (1994)
strings

Morphology
Karttunen et al. (1992)
strings

Syntax
Shieber (1985)
strings

Subregular Complexity



Cognitive Parallelism

Strong Cognitive Parallelism Hypothesis

Phonology, (morphology), and syntax have the **same subregular complexity** over their respective **structural representations**.

We gain a unified perspective on:

- typology
- learnability
- cognition

Cognitive Parallelism

Strong Cognitive Parallelism Hypothesis

Phonology, (morphology), and syntax have the **same subregular complexity** over their respective **structural representations**.

We gain a unified perspective on:

- typology
  - × Intervocalic Voicing iff applied an even times in the string
  - × Have a CP iff it dominates ≥ 3 TPs
- learnability
- cognition

Cognitive Parallelism

Strong Cognitive Parallelism Hypothesis

Phonology, (morphology), and syntax have the **same subregular complexity** over their respective **structural representations**.

We gain a unified perspective on:

- typology
  - × Intervocalic Voicing iff applied an even times in the string
  - × Have a CP iff it dominates ≥ 3 TPs
- learnability

  Learnable from positive examples of strings/trees.
- cognition

Cognitive Parallelism

Strong Cognitive Parallelism Hypothesis

Phonology, (morphology), and syntax have the **same subregular complexity** over their respective **structural representations**.

We gain a unified perspective on:

- typology
  - × Intervocalic Voicing iff applied an even times in the string
  - × Have a CP iff it dominates ≥ 3 TPs
- learnability
  Learnable from positive examples of strings/trees.
- cognition
  Finite, flat memory

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \to b$; $1(b) \to b$
- $1(a) \to a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \to b$; $1(b) \to b$
- $1(a) \to a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- Some island constrains arise naturally from this perspective (e.g., Adjunct Island Constraint, SpIC, ATB movement)

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

- Some island constrains arise naturally from this perspective (e.g., Adjunct Island Constraint, SpIC, ATB movement)
- Constraints improve parsing performance by **exponentially reducing** the search space (Stabler 2013)

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

Sensing Tree Automata (Martens 2006) as a subregular bound on the complexity of syntactic dependencies.

- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

- Some island constrains arise naturally from this perspective (e.g., Adjunct Island Constraint, SpIC, ATB movement)
- Constraints improve parsing performance by **exponentially reducing** the search space (Stabler 2013)
- Can be pre-compiled in the MG parse schema as a deterministic **top-down filter** (De Santo & Graf, in prep.)

Stacked RCs and Parallelism Effects

English Stacked RCs (Zhang, 2017)

(15) **The horse** $[_{RC_1}$ that t chased the wolf] $[_{RC_2}$ that t kicked the elephant] . . . **ss**

(16) **The horse** $[_{RC_1}$ that the wolf chased t] $[_{RC_2}$ that t kicked the elephant] . . . **os**

(17) **The horse** $[_{RC_1}$ that the wolf chased t] $[_{RC_2}$ that the elephant kicked t] . . . **oo**

(18) **The horse** $[_{RC_1}$ that t chased the wolf] $[_{RC_2}$ that the elephant kicked t] . . . **so**

- Zhang (2017) found **parallelism effects** in stacked RC processing:
  SS << OS, OO << SO.
- But she also showed that no combination of metrics can account for these effects.
- Proposal: metric encoding memory reactivation

Stacked RCs and Parallelism Effects

English Stacked RCs (Zhang, 2017)

(15) **The horse** [$_{RC_1}$ that t chased the wolf] [$_{RC_2}$ that t kicked the elephant] . . . **ss**

(16) **The horse** [$_{RC_1}$ that the wolf chased t] [$_{RC_2}$ that t kicked the elephant] . . . **os**

(17) **The horse** [$_{RC_1}$ that the wolf chased t] [$_{RC_2}$ that the elephant kicked t] . . . **oo**

(18) **The horse** [$_{RC_1}$ that t chased the wolf] [$_{RC_2}$ that the elephant kicked t] . . . **so**

- Zhang (2017) found **parallelism effects** in stacked RC processing:
  SS << OS, OO << SO.
- But she also showed that no combination of metrics can account for these effects.
- Proposal: metric encoding memory reactivation

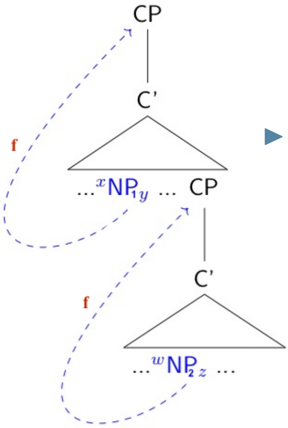
Feature Reactivation

REACTIVATION For each node m_i associated to a movement feature f^-, its reactivation is $i(m_i) - o(m_{i-1})$; the index of m_i minus the outdex of the closest preceding node also associated to f^-, if it exists.

- Assume the NPs are associated to the same movement feature f^-

Feature Reactivation

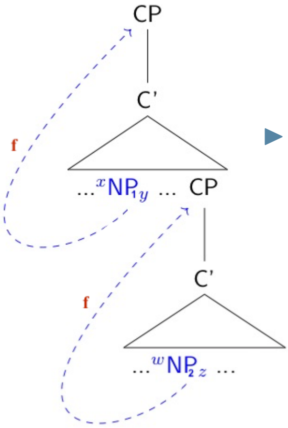
REACTIVATION For each node m_i associated to a movement feature f^-, its reactivation is $i(m_i) - o(m_{i-1})$; the index of m_i minus the outdex of the closest preceding node also associated to f^-, if it exists.



- Assume the NPs are associated to the same movement feature f^-

Feature Reactivation

REACTIVATION For each node m_i associated to a movement feature f^-, its reactivation is $i(m_i) - o(m_{i-1})$; the index of m_i minus the outdex of the closest preceding node also associated to f^-, if it exists.



- Assume the NPs are associated to the same movement feature f^-

TENURE (NP$_1$) $y - x$
TENURE (NP$_2$) $z - w$
REACTIVATION(NP$_2$) $w - y$

Feature Reactivation: Base Metrics

- feature-associated metrics

$\textsc{SumR}^f$ $\sum_{m_i \in M^f} i(m_i) - o(m_{i-1})$

$\textsc{MaxR}^f$ $max(\{i(m_i) - o(m_{i-1}) | m_i \in M^f\})$

$\textsc{AvgR}^f$ $\frac{\textsc{SumR}}{|M^f|}$

- comprehensive metrics

$\textsc{SumR}$ $\sum_{f \in \mathcal{M}} \textsc{SumR}^f$

$\textsc{MaxR}$ $max(\{\textsc{SumR}^f | f \in \mathcal{M}\})$

$\textsc{AvgR}$ $\frac{\textsc{SumR}}{|\mathcal{M}|}$

Priming Effects

(19) I saw

a. $[_{RC_1}$ the horse that chased the lions $]$ **SRC**

b. and $[_{RC_2}$ the mouse that kissed the chicken $]$ **SRC**

(20) I saw

a. $[_{RC_1}$ The horse that chased the lions$]$ **SRC**

b. and $[_{RC_2}$ the mouse that the chicken kissed $]$ **ORC**

(21) I saw

a. $[_{RC_1}$ the horse that the lions chased $]$ **ORC**

b. and $[_{RC_2}$ the mouse that kissed the chicken $]$ **SRC**

(22) I saw

a. $[_{RC_1}$ the horse that the lions chased$]$ **ORC**

b. and $[_{RC_2}$ the mouse that the chicken kissed$]$ **ORC**

The Role of Economy

- Economy considerations ubiquitous in Generative syntax (Chomsky 1995, Collins 2001, Boskovic and Messick 2017, a.o.)

But:

- What is the relevant notion of cost?
- What does simplicity mean in practice?
- Do fine-grained syntactic details matter?

What's to come

- Implemented economy principles might diverge from general intuitions
- **A Test Case**:
  → The PR-First Hypothesis for Italian
  → MG model as a testing framework!

The Role of Economy

- Economy considerations ubiquitous in Generative syntax (Chomsky 1995, Collins 2001, Boskovic and Messick 2017, a.o.)

But:

- What is the relevant notion of cost?
- What does simplicity mean in practice?
- Do fine-grained syntactic details matter?

What's to come

- Implemented economy principles might diverge from general intuitions
- **A Test Case**:
  → The PR-First Hypothesis for Italian
  → MG model as a testing framework!

So Young Lee

Attachment and Relative Clauses (RC)

- They saw the daughter of the actress that was on the balcony

NP_1	**The daughter** was on the balcony	HA
NP_2	**The actress** was on the balcony	LA

English: **LA** interpretation

- Late Closure (Frazier 1978),
  Recency (Gibson 1991, Gibson et al. 1996), ...

Universal locality principles?

- Spanish: **HA** interpretation
  - Tuning Hypothesis
    (Cuetos & Mitchell 1988, Mitchell & Cuetos 1991)
    Construal (Frazier & Clifton 1996), ...

Attachment and Relative Clauses (RC)

- They saw the daughter of the actress that was on the balcony

NP_1	**The daughter** was on the balcony	HA
NP_2	**The actress** was on the balcony	LA

English: **LA** interpretation

- Late Closure (Frazier 1978),
  Recency (Gibson 1991, Gibson et al. 1996), ...

Universal locality principles?

- Spanish: **HA** interpretation
  - Tuning Hypothesis
    (Cuetos & Mitchell 1988, Mitchell & Cuetos 1991)
    Construal (Frazier & Clifton 1996), ...

A Complex Cross-Linguistic Scenario

HA vs LA languages?
RC preferences cross-linguistically affected by a variety of factors

- Syntactic environment
  (Fernandez 2003, Gibson et al. 1996, De Vincenzi and Job 1993)
- Prosodic effects (Teira and Igoa 2007, Hemforth et al. 2015)
- Lexical-semantic properties of the DPs
  (MacDonald et al. 1994, Gilboy et al. 1995)
- Online vs. Offline Differences
  (Fernandez 2003, Wager et al. 2009, Lourenco-Gomes et al. 2011)
- Individual WM effects (Swets et al. 2007)

None of these fully accounts for the LA vs HA variation

A Complex Cross-Linguistic Scenario

HA vs LA languages?
RC preferences cross-linguistically affected by a variety of factors

- Syntactic environment
  (Fernandez 2003, Gibson et al. 1996, De Vincenzi and Job 1993)
- Prosodic effects (Teira and Igoa 2007, Hemforth et al. 2015)
- Lexical-semantic properties of the DPs
  (MacDonald et al. 1994, Gilboy et al. 1995)
- Online vs. Offline Differences
  (Fernandez 2003, Wager et al. 2009, Lourenco-Gomes et al. 2011)
- Individual WM effects (Swets et al. 2007)

None of these fully accounts for the LA vs HA variation

Grillo & Costa: Pseudo-RCs in Italian

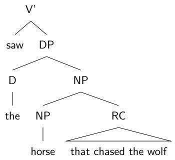
(23) (Io) Ho visto **[la nonna della ragazza che gridava]**
(I) have seen the grandma of the girl that screaming
'I saw **[the grandma of the girl that was screaming]**"

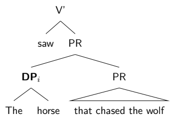
- RC: HA
- RC: LA

Grillo & Costa: Pseudo-RCs in Italian

(23) (Io) Ho visto **[la nonna della ragazza che gridava]**
(I) have seen the grandma of the girl that screaming
'I saw **[the grandma of the girl that was screaming]**"

- RC: HA
- RC: LA
- PR





Grillo & Costa: Pseudo-RCs in Italian

(23) (Io) Ho visto [la nonna della ragazza che gridava]
(I) have seen the grandma of the girl that screaming
'I saw **[the grandma of the girl that was screaming]**'

- RC: HA
- RC: LA
- PR: ∼ HA

- RCs are NP-modifiers and denote properties of entities
- PRs are complements of VPs and denote events/situations
  - **Only** compatible with a **HA** reading!

So What? PRs and Attachment Preferences

- The grandma of the girl that was screaming
  - RC: HA
  - RC: LA
  - PR: HA

The Pseudo-Relative First Hypothesis (Grillo & Costa 2014)

All else being equal:

- When available: PR **preferred over** RC parse (so: $\sim$ HA)
- Otherwise: LA RC **preferred over** HA RC parse

So What? PRs and Attachment Preferences

- The grandma of the girl that was screaming
  - RC: HA
  - RC: LA
  - PR: HA

The Pseudo-Relative First Hypothesis (Grillo & Costa 2014)

All else being equal:

- When available: PR **preferred over** RC parse (so: $\sim$ HA)
- Otherwise: LA RC **preferred over** HA RC parse

The PR First Hypothesis

(24)

(Io)	Ho	visto	[la	nonna	della	ragazza	che	gridava]
(I)	have	seen	the	grandma	of the	girl	that	screaming

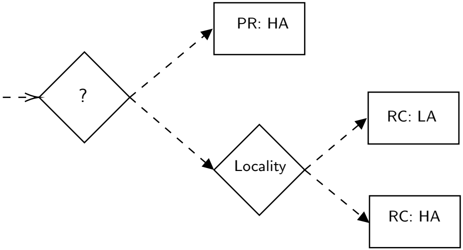
'I saw **[the grandma of the girl that was screaming]**'

Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013; a.o.)

- Appear freely with proper names, no relative pronouns, ...
- Verb type restrictions
- Tense/aspect restrictions

The PR First Hypothesis

(24) (Io) Ho visto [la nonna della ragazza che gridava]
(I) have seen the grandma of the girl that screaming
'I saw [the grandma of the girl that was screaming]'

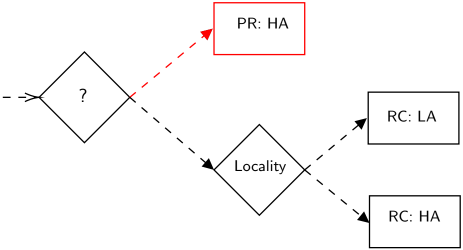


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

- Appear freely with proper names, no relative pronouns, ...
- Verb type restrictions
- Tense/aspect restrictions

The PR First Hypothesis

(24) (Io) Ho visto [la nonna della ragazza che gridava]
(I) have seen the grandma of the girl that screaming
'I saw **[the grandma of the girl that was screaming]**'

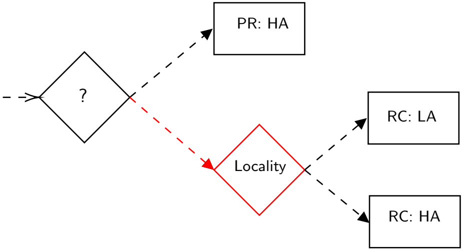


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

- Appear freely with proper names, no relative pronouns, ...
- Verb type restrictions
- Tense/aspect restrictions

The PR First Hypothesis

(24) (Io) Ho visto [la nonna della ragazza che gridava]
(I) have seen the grandma of the girl that screaming
'I saw [the grandma of the girl that was screaming]'

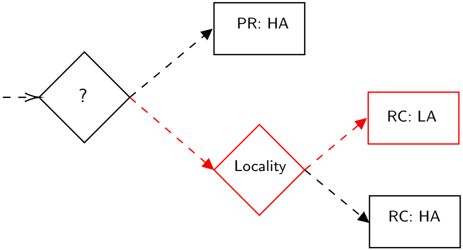


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

- Appear freely with proper names, no relative pronouns, ...
- Verb type restrictions
- Tense/aspect restrictions

The PR First Hypothesis

(24) (Io) Ho visto [la nonna della ragazza che gridava]
(I) have seen the grandma of the girl that screaming
'I saw **[the grandma of the girl that was screaming]**'

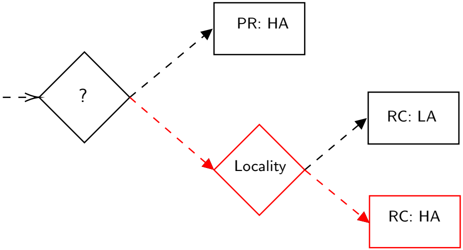


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

- Appear freely with proper names, no relative pronouns, ...
- Verb type restrictions
- Tense/aspect restrictions

The PR First Hypothesis

(24) (Io) Ho visto [la nonna della ragazza che gridava]
(I) have seen the grandma of the girl that screaming
'I saw [the grandma of the girl that was screaming]'

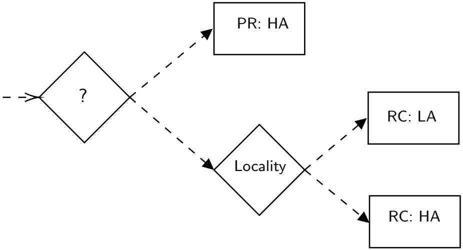


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

- Appear freely with proper names, no relative pronouns, ...
- Verb type restrictions
- Tense/aspect restrictions

The PR First Hypothesis

(24) (Io) Ho visto [la nonna della ragazza che gridava]
(I) have seen the grandma of the girl that screaming
'I saw **[the grandma of the girl that was screaming]**'



Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

- Appear freely with proper names, no relative pronouns, ...
- Verb type restrictions
- Tense/aspect restrictions

Grillo and Costa (2014)

- The daughter of the actress **[that was on the balcony]**
  - RC: HA
  - RC: LA
  - PR: ($\sim$) HA

Online too!

- Italian: De Santo & Lee (2022a)
- Spanish: Aguilar et al. (2020)

Grillo and Costa (2014)

- The daughter of the actress **[that was on the balcony]**
  - RC: HA
  - RC: LA
  - PR: ($\sim$) HA

(57) **Stimuli Experiment II**

a. PR/ RC CONDITION: PR-VERBS
Gianni ha visto il figlio del medico che correva.
G. saw the son of the doctor running.

b. RC ONLY CONDITION: STATIVE VERBS
Gianni vive con il figlio del medico che correva.
G. lives with the son of the doctor running.

Online too!

- Italian: De Santo & Lee (2022a)
- Spanish: Aguilar et al. (2020)

Grillo and Costa (2014)

- The daughter of the actress **[that was on the balcony]**
  - RC: HA
  - RC: LA
  - PR: (∼) HA

(57) **Stimuli Experiment II**

a. PR/ RC CONDITION: PR-VERBS
Gianni ha visto il figlio del medico che correva.
G. saw the son of the doctor running.

b. RC ONLY CONDITION: STATIVE VERBS
Gianni vive con il figlio del medico che correva.
G. lives with the son of the doctor running.

Table 6
Percentage of high attachment preferences.

Eventive	Stative
78.6%	24.2%

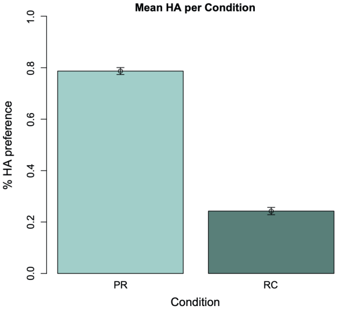


Fig. 2. Summary of attachment preference experiment 2.

Online too!

- Italian: De Santo & Lee (2022a)
- Spanish: Aguilar et al. (2020)

Grillo and Costa (2014)

- The daughter of the actress **[that was on the balcony]**
  - RC: HA
  - RC: LA
  - PR: (∼) HA

(57) **Stimuli Experiment II**

a. PR/ RC CONDITION: PR-VERBS
Gianni ha visto il figlio del medico che correva.
G. saw the son of the doctor running.

b. RC ONLY CONDITION: STATIVE VERBS
Gianni vive con il figlio del medico che correva.
G. lives with the son of the doctor running.

Table 6
Percentage of high attachment preferences.

Eventive	Stative
78.6%	24.2%

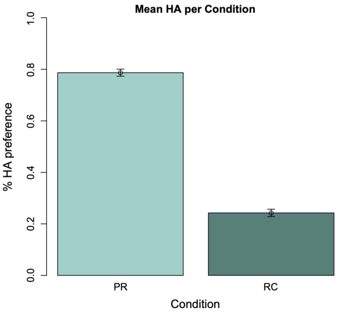


Fig. 2. Summary of attachment preference experiment 2.

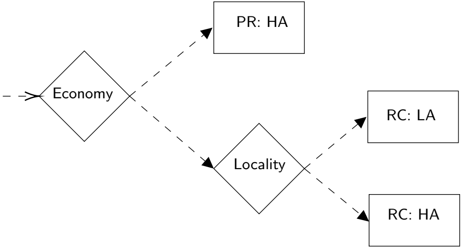
Online too!

- Italian: De Santo & Lee (2022a)
- Spanish: Aguilar et al. (2020)

PR-First: Why?

Question

Why should PRs be preferred?



One Hypothesis: Structural Economy (Grillo & Costa 2014)

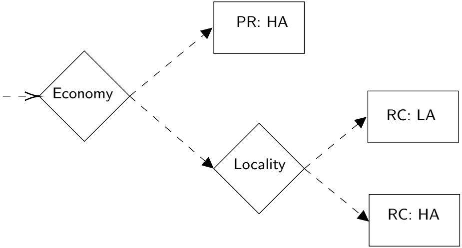
- PR structurally less complex than RC
- RCs: richer and more articulated functional domain

Can we evaluate structural economy quantitatively?

PR-First: Why?

Question

Why should PRs be preferred?



One Hypothesis: Structural Economy (Grillo & Costa 2014)

- PR structurally less complex than RC
- RCs: richer and more articulated functional domain

Can we evaluate structural economy quantitatively?

Modeling PR-First

Why should PRs be easier/preferred?

- Can we evaluate structural economy quantitatively?
- Do different syntactic choices matter?

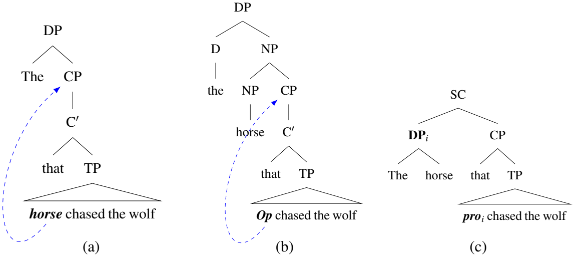


Figure 2: Sketches of the *(a)* RC with Promotion, *(b)* RC with Wh-movement, and *(c)* PR analyses for the sentence *The horse that the wolf chased*.

Modeling Results (De Santo & Lee, 2022b)

	MG Parser: MaxT
Hypothesis	
PR > HA	
PR > LA	
LA > HA	

(25) (Io) Ho visto la nonna della ragazza che gridava
(I) have seen the grandma of the girl that screaming
'I saw the grandma of the girl that was screaming"

- The PR> HA RC depends on syntactic choices
- No metric predicts PR> LA RC
- In sum:
  No immediate support for a parsing economy explanation
- LA>HA arises without explicit locality constraints!

Modeling Results (De Santo & Lee, 2022b)

	MG Parser: MaxT	
Hypothesis	**Promotion**	**Wh-mov**
PR > HA		
PR > LA		
LA > HA		

(25) (Io) Ho visto la nonna della ragazza che gridava
(I) have seen the grandma of the girl that screaming
'I saw the grandma of the girl that was screaming"

- The PR> HA RC depends on syntactic choices
- No metric predicts PR> LA RC
- **In sum:**
  No immediate support for a parsing economy explanation
- LA>HA arises without explicit locality constraints!

Modeling Results (De Santo & Lee, 2022b)

	MG Parser: MaxT	
Hypothesis	**Promotion**	**Wh-mov**
PR > HA	✓	Tie
PR > LA		
LA > HA		

(25) (Io) Ho visto la nonna della ragazza che gridava
(I) have seen the grandma of the girl that screaming
'I saw the grandma of the girl that was screaming"

- The PR> HA RC depends on syntactic choices
- No metric predicts PR> LA RC
- **In sum:**
  No immediate support for a parsing economy explanation
- LA>HA arises without explicit locality constraints!

Modeling Results (De Santo & Lee, 2022b)

	MG Parser: MaxT	
Hypothesis	**Promotion**	**Wh-mov**
PR > HA	✓	Tie
PR > LA	×	×
LA > HA		

(25) (Io) Ho visto la nonna della ragazza che gridava
(I) have seen the grandma of the girl that screaming
'I saw the grandma of the girl that was screaming"

- The PR> HA RC depends on syntactic choices
- No metric predicts PR> LA RC
- **In sum:**
  No immediate support for a parsing economy explanation
- LA>HA arises without explicit locality constraints!

Modeling Results (De Santo & Lee, 2022b)

	MG Parser: MaxT	
Hypothesis	**Promotion**	**Wh-mov**
PR > HA	✓	Tie
PR > LA	×	×
LA > HA	✓	✓

(25) (Io) Ho visto la nonna della ragazza che gridava
(I) have seen the grandma of the girl that screaming
'I saw the grandma of the girl that was screaming"

- The PR> HA RC depends on syntactic choices
- No metric predicts PR> LA RC
- **In sum:**
  No immediate support for a parsing economy explanation
- LA>HA arises without explicit locality constraints!

Modeling Results (De Santo & Lee, 2022b)

	MG Parser: MaxT	
Hypothesis	**Promotion**	**Wh-mov**
PR > HA	✓	Tie
PR > LA	×	×
LA > HA	✓	✓

(25) (Io) Ho visto la nonna della ragazza che gridava
(I) have seen the grandma of the girl that screaming
'I saw the grandma of the girl that was screaming"

- The PR> HA RC depends on syntactic choices
- No metric predicts PR> LA RC
- **In sum:**
  No immediate support for a parsing economy explanation
- LA>HA arises without explicit locality constraints!

TI/Dr: The Value of Formal Models

A fully specified model of syntactic cost:

- Allows evaluation of economy definitions
- Shows that syntactic choices affect "cost" in unexpected ways
- Suggest ways to narrow down the space of plausible accounts

Beyond these results

- Cross-linguistic and cross-analysis validation
- A variety of definitions for *cost* in parsing (Boston, 2012)
  - E.g., # bounding nodes/phases, discourse referents, retrieval
  - Pragmatic Economy?
    E.g. Reference Theory (Altmann & Steedman 1988)
- Investigating economy principles more broadly

Tl/Dr: The Value of Formal Models

A fully specified model of syntactic cost:

- Allows evaluation of economy definitions
- Shows that syntactic choices affect "cost" in unexpected ways
- Suggest ways to narrow down the space of plausible accounts

Beyond these results

- Cross-linguistic and cross-analysis validation
- A variety of definitions for *cost* in parsing (Boston, 2012)
  - E.g., # bounding nodes/phases, discourse referents, retrieval
  - Pragmatic Economy?
    E.g. Reference Theory (Altmann & Steedman 1988)
- Investigating economy principles more broadly

A Look at HA Languages (Grillo & Costa 2015)

Table 4
Attachment preferences and PR availability.

Language	Attachment	PRs
English	Low	*
Romanian	Low	*
Basque	Low	*
Chinese	Low	*
German (?)	High/Low	*
Russian (?)	High	*
Bulgarian (?)	High/Low	*
Norwegian (?)	Low	✓
Swedish (?)	Low	✓
Spanish	High	✓
Galician	High	✓
Dutch	High	✓
Italian	High	✓
French	High	✓
Serbo-Croatian	High	✓
Japanese	High	✓
Korean	High	✓
Greek	High	✓
Portuguese	High	✓

Figure: Survey of Attachment preferences from Grillo & Costa (2014)

PRs: Modeling Results 1

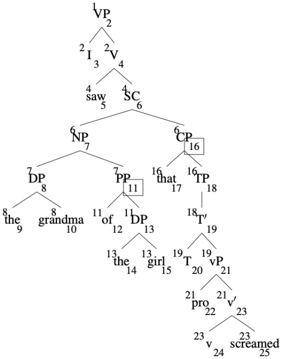


Figure 3: Annotated derivation trees for the Italian sentence *I saw the grandma of the girl that screamed*, according to a pseudo-relaive clause analysis. The tree is treated as a VP since additional structure in the matrix clause would be identical across comparisons.

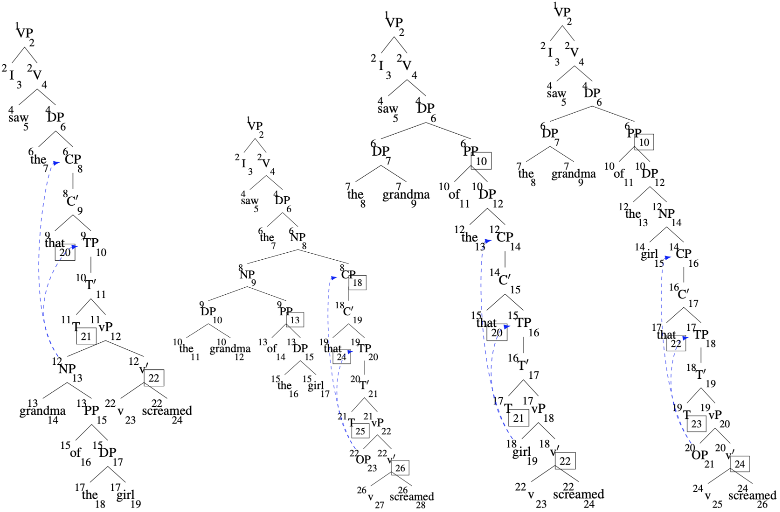
	MG Parser	
Hypothesis	**Promotion**	**Wh-mov**
PR < HA	✓	Tie
PR < LA	×	×
LA < HA	✓	✓

Table 1: Summary of the predictions made by a *pseudo-relative first* account, and corresponding parser's predictions based on MAXTENURE, as pairwise comparisons (x < y: x is preferred over y).

	MAXT	
	Promotion	**Wh-mov**
PR	10/CP	
HA	11/that	10/CP
LA	5/that	7/that

Table 2: MAXT values (*value/node*) by construction, with RCs modulated across a promotion and wh-movement analysis.

PRs: Modeling Results 2



(a) (b) (c) (d)

Our Study

Question: Online effects of PR availability in Italian?

- Modulating:
  - Type of Verb: Perceptual vs. Non-perceptual
  - Attachment: HA vs. LA
- Temporal ambiguity HA/LA until # agreement on the **verb**

(2)

	Verb	Interpretation		before	target	after	
a.	PR/RC (Perceptual)	LA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
b.	PR/RC (Perceptual)	HA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correva was running-SG	la the	maratona marathon
c.	RC only	LA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
d.	RC only	HA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correva was running-SG	la the	maratona marathon

- Counterbalancing # features (singular vs plural) on DP_1/DP_2

Our Study

Question: Online effects of PR availability in Italian?

- Modulating:
  - Type of Verb: Perceptual vs. Non-perceptual
  - Attachment: HA vs. LA
- Temporal ambiguity HA/LA until # agreement on the **verb**
  - **Perceptual Verbs**: costly LA disambiguation (on verb)
  - **Non-Perceptual Verbs**: costly HA disambiguation (on verb)

(2)	Verb	Interpretation		before	target	after	
a.	PR/RC (Perceptual)	LA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
b.	PR/RC (Perceptual)	HA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correva was running-SG	la the	maratona marathon
c.	RC only	LA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
d.	RC only	HA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correva was running-SG	la the	maratona marathon

- Counterbalancing # features (singular vs plural) on DP_1/DP_2

Decomposing the Hypothesis: Perceptual Verbs

- Temporal HA/LA ambiguity until # agreement on the **verb**

(2)

	Verb	Interpretation		before	target	after	
a.	PR/RC (Perceptual)	LA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
b.	PR/RC (Perceptual)	HA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correva was running-SG	la the	maratona marathon
c.	RC only	LA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
d.	RC only	HA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correva was running-SG	la the	maratona marathon

Perceptual Verbs

- PR vs RC
- PR-first: HA-like interpretation is preferred
- LA disambiguation (on verb) should be costly

Decomposing the Hypothesis: Perceptual Verbs

- Temporal HA/LA ambiguity until # agreement on the **verb**

(2)

	Verb	Interpretation		before	target	after	
a.	PR/RC (Perceptual)	LA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
b.	PR/RC (Perceptual)	HA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correva was running-SG	la the	maratona marathon
c.	RC only	LA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
d.	RC only	HA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correva was running-SG	la the	maratona marathon

Non-Perceptual Verbs

- Just RC
- LA interpretation (more local) is preferred
- HA disambiguation (on verb) should be costly

Study Details: Summary of Predictions

- Temporarily ambiguous sentences modulating:
  - Type of Verb: Perceptual vs. Non-perceptual
  - Attachment: HA vs. LA

Hypothesis

Perceptual Verbs

- LA disambiguation (on verb) should be costly

Non-Perceptual Verbs

- HA disambiguation (on verb) should be costly

- 74 participants (recruited through Prolific, run on Ibex Farm)
- 24 item sets, 48 fillers
- Self-paced reading

Results: Behavioral Data

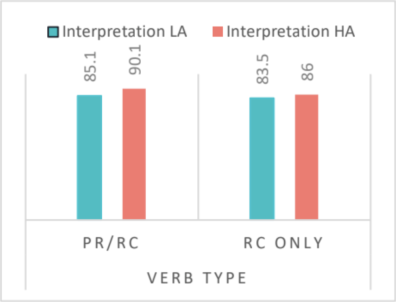
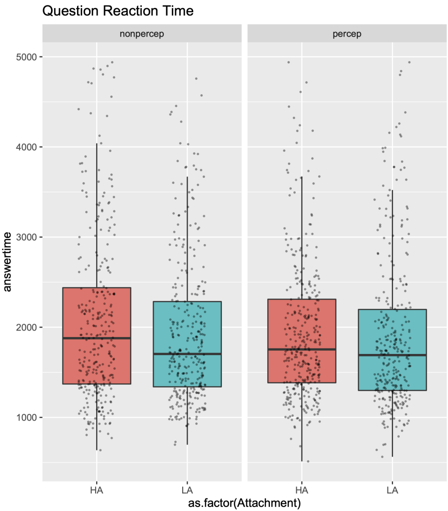
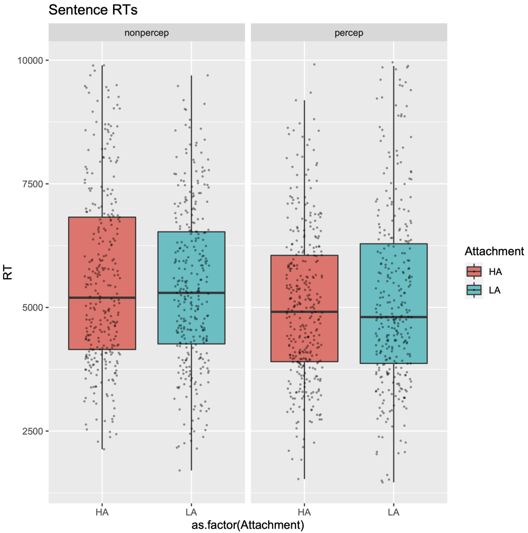


Figure 2. The results of the comprehension test



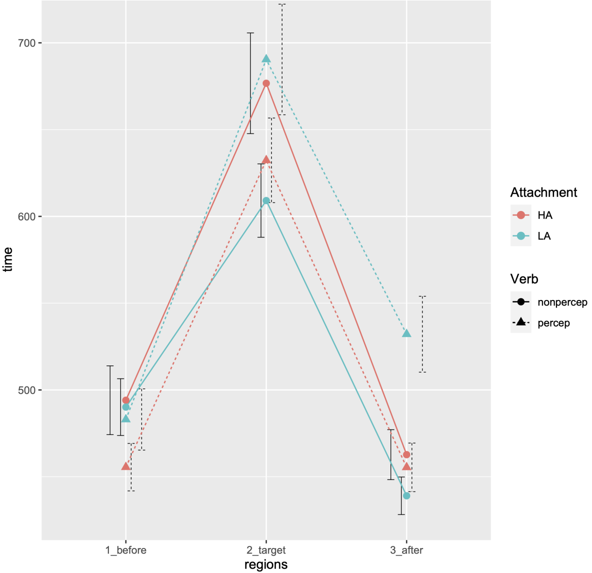
- No effect of Verb, Attachment, or Interaction

Results: Sentence Reading Time



- Effect of the Verb ($p<0.01$) and Verb*Attachment ($p<0.05$)

Results: RTs by ROI



Hypothesis

- **Percep**: LA costly
- **Non-Perc**: HA costly

- Pre-Target: No Effect
- Target: Verb*Attachment ($p < 0.01$)
- Spillover: Verb*Attachment ($p < 0.001$) and Verb ($p<0.001$)

Online Effects: Stimuli and RTs

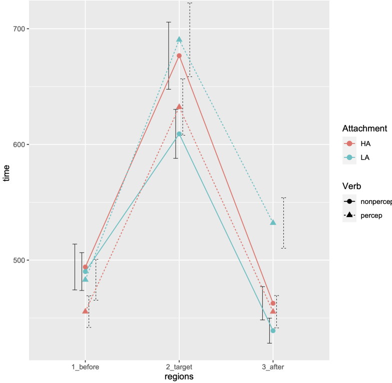
(2)

	Verb	Interpretation		before	target	after	
a.	PR/RC (Perceptual)	LA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
b.	PR/RC (Perceptual)	HA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correva was running-SG	la the	maratona marathon
c.	RC only	LA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
d.	RC only	HA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correva was running-SG	la the	maratona marathon

Online Effects: Stimuli and RTs

(2)

	Verb	Interpretation		before	target	after	
a.	PR/RC (Perceptual)	LA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
b.	PR/RC (Perceptual)	HA	Gianni vide il figlio dei medici Gianni saw the son-SG of the doctors-PL	che who	correva was running-SG	la the	maratona marathon
c.	RC only	LA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correvano were running-PL	la the	maratona marathon
d.	RC only	HA	Gianni visse con il figlio dei medici Gianni lived with the son-SING of the doctors-PL	che who	correva was running-SG	la the	maratona marathon



Hypothesis (@ verb)

- **Percep**: LA costly
- **Non-Perc**: HA costly

See also Aguilar et al. (2021)

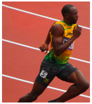
PRs vs RCs: Interpretative Differences

(6) RC: *John saw the man that runs*

∃e [see(e) & EXPERIENCER(e)(John) & STIMULUS(the unique man that ran)(e)]

There is an event of *seeing* and the experiencer of that event is *John* and the stimulus of the event is *the unique man that ran*.[8]

(7) PR: *John saw the man running*



∃e∃e'[see(e) & EXPERIENCER(e)(John) & STIMULUS(e')(e) & run(e') & AGENT(e')(the man)]

There is an event of *seeing* and the experiencer of that event is *John* and the stimulus of the event is *an event of running* and the agent of running is *the man*.[9]

PRs vs RCs 1

i. PRs appear freely with proper names (13-a), contrary to RCs (13-b).[7]

(13) a. Ho visto Gianni che correva (Italian)
He visto a [$_{PR}$ Juan que corría] (Spanish)
J'ai vu [$_{PR}$ Jean qui courait] (French)
'I saw Gianni running.'
b. *I saw John that ran.
c. Ho visto Gianni, che correva. Appositive

ii. Relative pronouns are banned from PRs, but obviously not from RCs:

(14) *Ho visto Gianni il quale correva.
Have.I seen Gianni the which run.IMPF.
'I saw Gianni who was running.'

iii. Just like other types of Small Clauses (see ungrammatical translation), PRs are only available with embedded subjects and cannot be construed with embedded objects (15-a), this restriction obviously does not apply to RCs (15-b)[8]:

(15) a. *Luigi ha visto [$_{PR}$ Gianni$_i$ che Maria baciava EC$_i$].
Luigi saw Gianni that Maria kissed EC.
'Luigi saw John Mary kissing EC.'
b. Luigi ha visto il ragazzo che Maria ha baciato <ragazzo>.
'Luigi saw the boy that Mary kissed.'

PRs vs RCs 2: Tense and Aspect Restrictions

(16) Ho visto il ragazzo/ *Gianni che correrà.
Have.I seen the boy/ *Gianni that run.FUT 'I saw the boy/*Gianni that will run.'

v. Restrictions to both inner and outer aspect hold for PRs. PRs require imperfective, but not perfective, aspect (17-a), as they denote ongoing events. They are further restricted to stage level properties and cannot denote individual level properties (17-b). Neither of these restrictions applies to RCs.

(17) a. Ho visto Gianni che correva/ *che è corso a casa.
'I saw Gianni running/ that had run home.'

b. Ho visto Gianni che *aveva gli occhi rossi*/ **aveva gli occhi blu*.
I saw Gianni that had the eyes red/ had the eyes blue.
'I saw Gianni with red eyes/ with blue eyes.' (Casalicchio, 2013, p. 117, ex. 160)

PRs vs RCs 3

Additionally, PRs and SCs can be freely coordinated (20-a,b), while neither of them can be coordinated with RCs (which is further evidence against a RC analysis of PRs or other types of clausal complements (20-c,d).

(20) a. SC & PR:
Ho visto [Gianni depresso] e [Piero che cercava di risollevarlo].
'I saw G. depressed and P. that was trying to cheer him up.'

b. SC & PR:
Ho visto [Gianni [depresso] e [che piangeva]].
'I saw G. depressed and that was crying.'

c. *RC & PR/SC:
*Ho visto [Gianni, [che vive con Maria], e [depresso/ che piangeva]].
'I saw G., who lives with M. and depressed/ that was crying.'

d. *PR/SC & FINITE CP:
*Ho visto [Gianni [che piangeva/ depresso] e [che P. cercava di risollevarlo]].
'I saw G. crying/ depressed and that P. tried to cheer him up.'

PRs vs RCs 4

iii. Just like other types of Small Clauses (see ungrammatical translation), PRs are only available with embedded subjects and cannot be construed with embedded objects (15-a), this restriction obviously does not apply to RCs (15-b)[8]:

(15) a. *Luigi ha visto [$_{PR}$ Gianni$_i$ che Maria baciava EC$_i$].
Luigi saw Gianni that Maria kissed EC.
'Luigi saw John Mary kissing EC.'

b. Luigi ha visto il ragazzo che Maria ha baciato <ragazzo>.
'Luigi saw the boy that Mary kissed.'