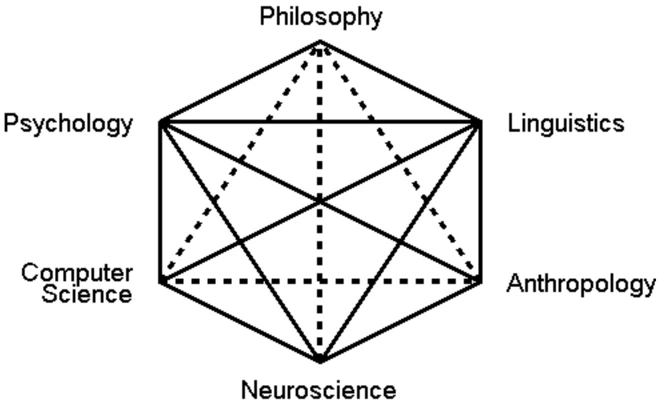
Computation as a Window into Linguistic Cognition

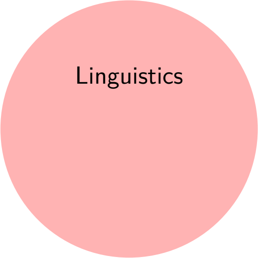
Aniello De Santo
he/him

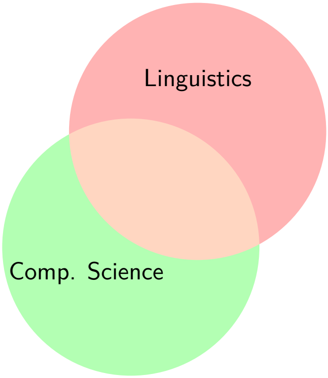
Get the slides!

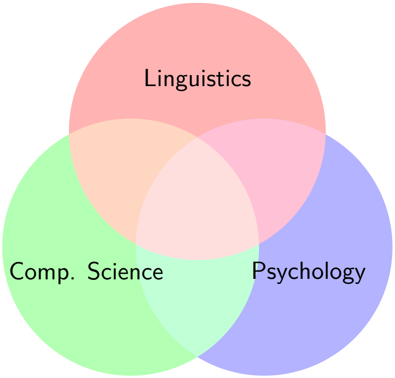
anielloedesanto.github.io/UM24.pdf
aniello.desanto@utah.edu

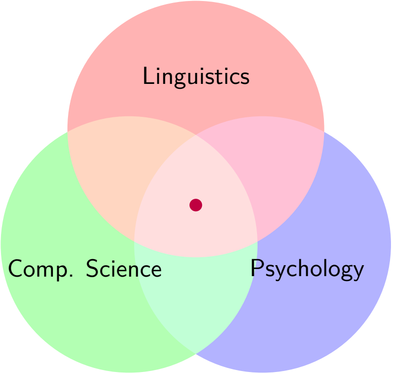
University of Michigan, Ann Arbor
February 2024


Philosophy
Psychology
Linguistics
Computer Science
Anthropology
Neuroscience

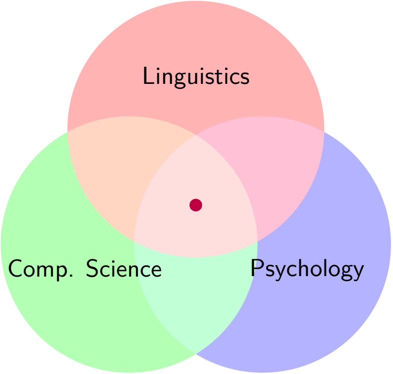

Linguistics


Linguistics
Comp. Science


Linguistics
Comp. Science
Psychology

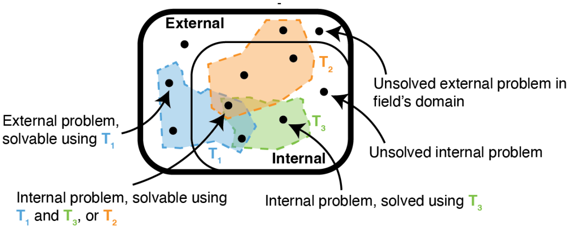


- Theory building (De Santo & Rawski 2022, Baggio, De Santo, Nunez 2024, Levenstein et al. 2024)
- Computational invariants in typology and acquisition
  (De Santo & Graf 2019, De Santo & Aksenova 2021, Johnson and De Santo 2023)
- Computational parallels across linguistic modules
  (Aksenova & De Santo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



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

Problems and Theories[1]



Some Problems

- what are the core representations?
- what do they tell us about processing?
- what do they tell us about learning?

[1] Levenstein, De Santo, ..., et al. (2024), Guest & Martin (2021), a.o.

Computation and Theory Building

> *[...] this is a confusion of two quite separate issues,* ***simulation and explanation****. [...] What we are* ***really*** *interested in [...] is explanation — in developing models that help us* ***understand how it is that people behave*** *that way, not merely demonstrating that we can build an artifact that behaves similarly.*
>
> *(Kaplan, 1995)*

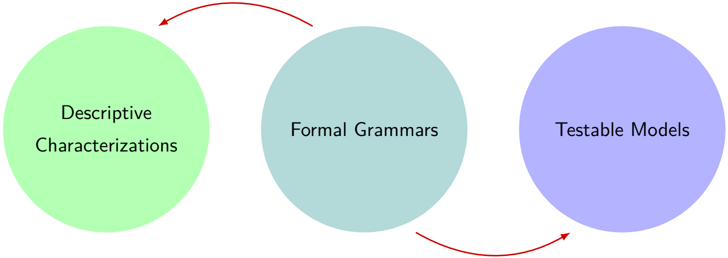
- Invariant properties of phenomena
- Specification of verbal theories

Computation and Theory Building

> *[...] this is a confusion of two quite separate issues, **simulation and explanation**. [...] What we are **really** interested in [...] is explanation — in developing models that help us **understand how it is that people behave** that way, not merely demonstrating that we can build an artifact that behaves similarly.*
>
> *(Kaplan, 1995)*

- Invariant properties of phenomena
- Specification of verbal theories

Building Bridges



Outline

Theories from Data?

Theories of linguistic representations from typological/empirical observations?

*The problem that we cannot deduce [...] theories from data is a limitation, or **perhaps an attribute**, of all empirical science [...] Still, one may abduce hypotheses [...] Abduction is **reasoning from observations** [...] It consists of two steps: generating candidate **hypotheses** (abduction proper), and selecting the "best" explanatory one[**s**] (inference to the **best explanation**).*

(van Roji & Baggio 2020, pg. 9)

Theories from Data?

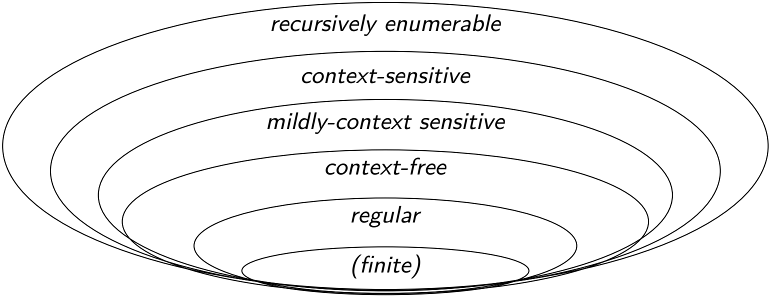
Theories of linguistic representations from typological/empirical observations?

The problem that we cannot deduce [...] theories from data is a limitation, or ***perhaps an attribute****, of all empirical science [...] Still, one may abduce hypotheses [...] Abduction is* ***reasoning from observations*** *[...] It consists of two steps: generating candidate* ***hypotheses*** *(abduction proper), and selecting the "best" explanatory one[****s****] (inference to the* ***best explanation****).*

(van Roji & Baggio 2020, pg. 9)

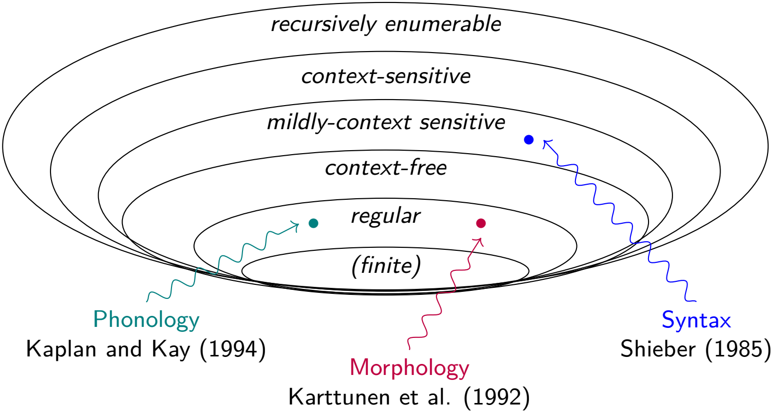
A Lens: Computational Theories of Language

Stringsets can be classified according to the requirements of the grammars that generate them.

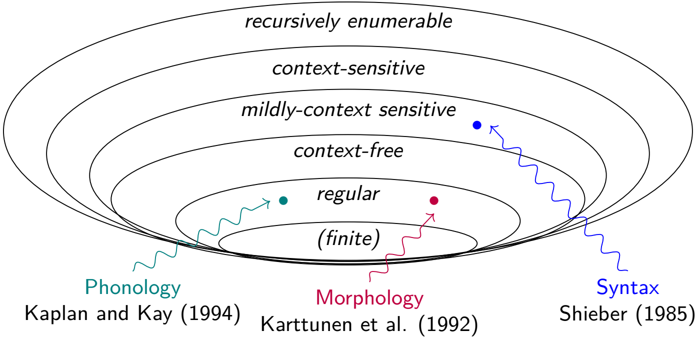


A Lens: Computational Theories of Language

Stringsets can be classified according to the requirements of the grammars that generate them.



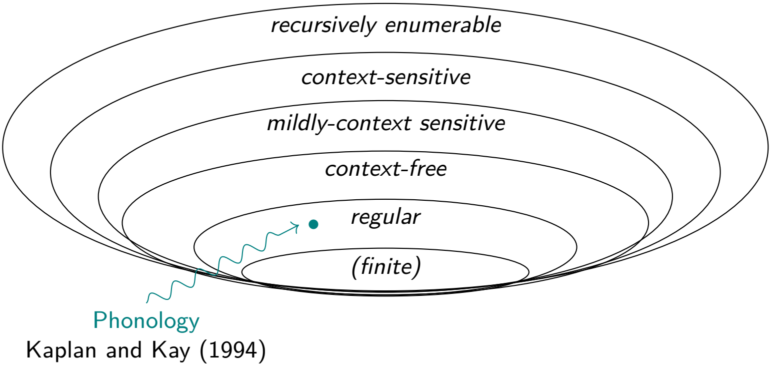
Precise Characterizations ⇒ Precise Predictions



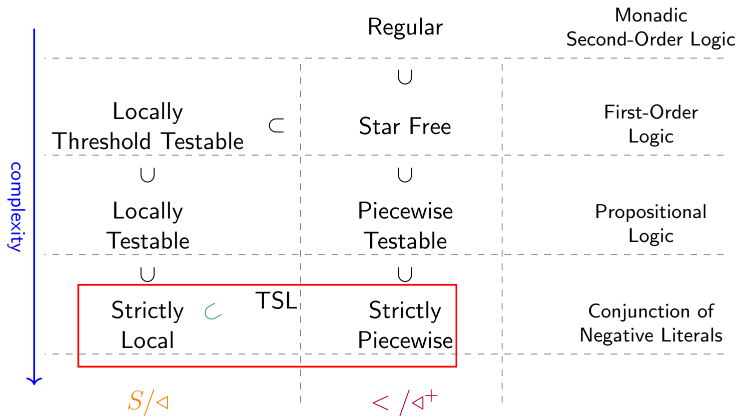
Precise predictions for:

- typology → e.g. no center embedding in phonology
- learnability → e.g. no Gold learning for regular languages
- cognition → e.g. finitely bounded working memory

Spoken Languages' Phonology as a Regular System



Beyond Monolithic Classes: Subregular Languages[2]



[2] McNaughton & Papert (1976), Heinz (2011), Chandlee & Heinz (2014), De Santo & Graf (2019), De Santo & Rawski (2022), a.o.

Local Phonotactic Dependencies

1. **Word-final devoicing**
   Forbid voiced segments at the end of a word

(1) a. * rad
    b. rat

1. **Intervocalic voicing**
   Forbid voiceless segments in between two vowels

(2) a. * faser
    b. fazer

These patters can be described by **strictly local** (SL) constraints.

Local Phonotactic Dependencies

1. **Word-final devoicing**
   Forbid voiced segments at the end of a word

(1) a. * rad
    b. rat

1. **Intervocalic voicing**
   Forbid voiceless segments in between two vowels

(2) a. * faser
    b. fazer

These patters can be described by **strictly local** (SL) constraints.

Local Dependencies in Phonology are SL

Example: Word-final devoicing

- Forbid voiced segments at the end of a word: *[+voice]$
- **German**: ***z**$, ***v**$,***d**$ ($ = word edge).

$ r a **d** $ $ r a t $

Example: Intervocalic voicing

- Forbid voicess segments in-between two vowels: *V[-voice]V
- **German**: *a**s**e, *i**s**e, *e**s**e, *i**s**i, . . .

$ f a **s** e r $ $ f a z e r $

Local Dependencies in Phonology are SL

Example: Word-final devoicing

- Forbid voiced segments at the end of a word: *[+voice]$
- **German**: ***z**$, ***v**$,***d**$ ($ = word edge).

* $ r a **d** $ *ok* $ r a t $

Example: Intervocalic voicing

- Forbid voicess segments in-between two vowels: *V[-voice]V
- **German**: *a**s**e, *i**s**e, *e**s**e, *i**s**i, . . .

$ f a **s** e r $ $ f a z e r $

Local Dependencies in Phonology are SL

Example: Word-final devoicing

- Forbid voiced segments at the end of a word: *[+voice]$
- **German**: ***z**$, ***v**$,***d**$ ($ = word edge).

* $ r a [**d** $] $\quad$ *ok* $ r a [t $]

Example: Intervocalic voicing

- Forbid voicess segments in-between two vowels: *V[-voice]V
- **German**: *a**s**e, *i**s**e, *e**s**e, *i**s**i, . . .

* $ f [a **s** e] r $ $\quad$ *ok* $ f [a z e] r $

Unbounded Dependencies Are Not SL

- **Samala Sibilant Harmony**
  Sibilants must not disagree in anteriority.
  (Applegate 1972)

(3) a. * hasxintilawaʃ
b. * haʃxintilawas
c. haʃxintilawaʃ

Example: Samala

*$ h a s x i n t i l a w a ʃ $

$ h a ʃ x i n t i l a w a ʃ $

Unbounded Dependencies Are Not SL

- **Samala Sibilant Harmony**
  Sibilants must not disagree in anteriority.
  (Applegate 1972)

(3) a. * hasxintilawaʃ
    b. * haʃxintilawas
    c. haʃxintilawaʃ

Example: Samala

* $ h a s x i n t i l a w a ʃ $

$ h a ʃ x i n t i l a w a ʃ $

Unbounded Dependencies Are Not SL

- **Samala Sibilant Harmony**
  Sibilants must not disagree in anteriority.
  (Applegate 1972)

(3) a. * hasxintilawaʃ
    b. * haʃxintilawas
    c. haʃxintilawaʃ

Example: Samala

* $ h a s x i n t i l a w a ʃ $

$ h a ʃ x i n t i l a w a ʃ $

Unbounded Dependencies Are Not SL

- **Samala Sibilant Harmony**
  Sibilants must not disagree in anteriority.
  (Applegate 1972)

(3) a. * hasxintilawaʃ

b. * haʃxintilawas

c. haʃxintilawaʃ

Example: Samala

* $ h a s x i n t i l a w a ʃ $

$ h a ʃ x i n t i l a w a ʃ $

Unbounded Dependencies Are Not SL

- **Samala Sibilant Harmony**
  Sibilants must not disagree in anteriority.
  (Applegate 1972)

(3) a. * hasxintilawaʃ
  b. * haʃxintilawas
  c. haʃxintilawaʃ

Example: Samala

* $ h a s x i n t i l a w a ʃ $

$ h a ʃ x i n t i l a w a ʃ $

- **But:** Sibilants can be arbitrarily far away from each other!

* $ s t a j a n o w o n w a ʃ $

Unbounded Dependencies Are Not SL

- **Samala Sibilant Harmony**
  Sibilants must not disagree in anteriority.
  (Applegate 1972)

(3) a. * hasxintilawaʃ
    b. * haʃxintilawas
    c. haʃxintilawaʃ

Example: Samala

* $ h a s x i n t i l a w a ʃ $

$ h a ʃ x i n t i l a w a ʃ $

- **But:** Sibilants can be arbitrarily far away from each other!

* $ s t a j a n o w o n w a ʃ $

Locality Over Tiers

* $ s t a j a n o w o n w a ʃ $

- Sibilants can be arbitrarily far away from each other!
- **Problem**: SL limited to locality domains of size n;

Tier-based Strictly Local (TSL) Grammars (Heinz et al. 2011)

- Projection of selected segments on a tier T (Goldsmith 1976)
- Strictly local constraints over T determine wellformedness
- Unbounded dependencies are local over **tiers**

Locality Over Tiers

* $ s t a j a n o w o n w a ʃ $

- Sibilants can be arbitrarily far away from each other!
- **Problem**: SL limited to locality domains of size n;

Tier-based Strictly Local (TSL) Grammars (Heinz et al. 2011)

- Projection of selected segments on a tier T (Goldsmith 1976)
- Strictly local constraints over T determine wellformedness
- Unbounded dependencies are local over **tiers**

Unbounded Dependencies are TSL

- Let's revisit Samala Sibilant Harmony

(4) a. * ha**s**xintilawaʃ

b. * haʃxintilawa**s**

c. haʃxintilawaʃ

- What do we need to project? [+strident]
- What do we need to ban? *[+ant][−ant],*[−ant][+ant]

I.E. *sʃ, *sʒ, *zʃ, *zʒ, *ʃs, *ʒs, *ʃz, *ʒz

Example: TSL Samala

s	ʃ		ʃ	ʃ

* $ h a **s** x i n t i l a w **ʃ** $

^{ok} $ h a **ʃ** x i n t i l a w **ʃ** $

Unbounded Dependencies are TSL

- Let's revisit Samala Sibilant Harmony

  (4) a. * ha**s**xintilawaʃ

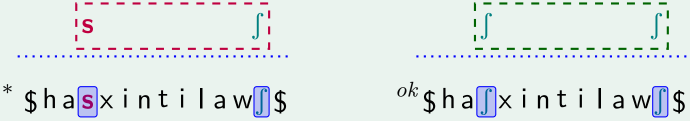
  b. * haʃxintilawa**s**

  c. haʃxintilawaʃ

- What do we need to project? [+strident]
- What do we need to ban? *[+ant][−ant],*[−ant][+ant]

  I.E. *sʃ, *sʒ, *zʃ, *zʒ, *ʃs, *ʒs, *ʃz, *ʒz

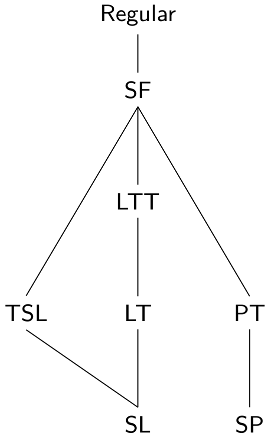
Example: TSL Samala



SL and TSL: So What?

*Descriptive characterizations focus on the **nature of the information** [...] that is needed in order to distinguish [...] a pattern*

Rogers & Pullum (2011)



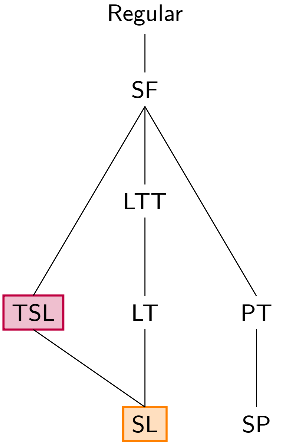
Invariants (De Santo & Rawski 2022)

- SL: adjacency
- TSL: relativized adjancency

- **But** typological variation is complex...
  (McMullin 2016, Mayer & Major 2018, De Santo & Graf 2019)

SL and TSL: So What?

> *Descriptive characterizations focus on the **nature of the information** [...] that is needed in order to distinguish [...] a pattern*
>
> *Rogers & Pullum (2011)*

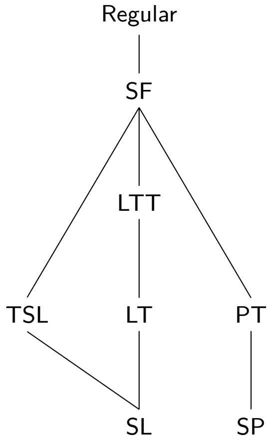


Invariants (De Santo & Rawski 2022)

- SL: adjacency
- TSL: relativized adjancency

- **But** typological variation is complex...
  (McMullin 2016, Mayer & Major 2018, De Santo & Graf 2019)

Refining the Hierarchy via Typological Insights



Observation

TSL is not closed under intersection
(De Santo & Graf, 2019)

- We want to also account for multiple processes
  So we can cover the complete phonotactics of a language
- Multiple non-interacting processes in attested patterns

Concurrent Processes

Sibilant Harmony in Imdlawn Tashlhiyt[3]

1) Underlying causative prefix /s(ː)-/

	Base	*Causative*	
a.	uga	**s**ː-uga	"be evacuated"
b.	a**s**ːtwa	**s**-a**s**ːtwa	"settle, be levelled"

[3] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Concurrent Processes

Sibilant Harmony in IMDLAWN TASHLHIYT[3]

1) Underlying causative prefix /s(ː)-/

	Base	*Causative*	
a.	uga	sː-uga	"be evacuated"
b.	asːtwa	s-asːtwa	"settle, be levelled"

2) Sibilant harmony

	Base	*Causative*	
a.	fiaʃr	ʃ- fiaʃr	"be full of straw, of discord"
b.	nza	zː-nza	"be sold"

[3] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Concurrent Processes

Sibilant Harmony in Imdlawn Tashlhiyt[3]

1) Underlying causative prefix /s(ː)-/

	Base	*Causative*	
a.	uga	**sː**-uga	"be evacuated"
b.	a**s**ːtwa	**s**-a**s**ːtwa	"settle, be levelled"

2) Sibilant harmony

	Base	*Causative*	
a.	fia**ʃ**r	**ʃ**- fia**ʃ**r	"be full of straw, of discord"
b.	n**z**a	**z**ː-n**z**a	"be sold"

3) Sibilant voicing harmony blocked

	Base	*Causative*	
a.	u**kz**	**s**ː-u**kz**	"recognize"
b.	**q**ːu**ʒ**ːi	**ʃ**- **q**u**ʒ**ːi	"be dislocated, broken"

[3] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Sibilant Harmony in Imdlawn Tashlhiyt [4]

Generalization (1/2)

Sibilants must agree in anteriority and voicing.

Grammar

$T = \{$ ʒ, s, z, ʃ $\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs, *ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, *ʒʃ, *ʒz $\}$

*z m: ʒ d a w l

ok ʒ m: ʒ d a w l

[4] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Sibilant Harmony in IMDLAWN TASHLHIYT [4]

Generalization (1/2)

Sibilants must agree in anteriority and voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ$\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, *ʒʃ, *ʒz $\}$

z ʒ

*z m: ʒ d a w l

ok ʒ m: ʒ d a w l

[4] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Sibilant Harmony in Imdlawn Tashlhiyt [4]

Generalization (1/2)

Sibilants must agree in anteriority and voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ $\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, *ʒʃ, *ʒz $\}$

*z ʒ

*z m: ʒ d a w l

ok ʒ m: ʒ d a w l

[4] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Sibilant Harmony in IMDLAWN TASHLHIYT [4]

Generalization (1/2)

Sibilants must agree in anteriority and voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ$\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, *ʒʃ, *ʒz $\}$

4 Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Sibilant Harmony in Imdlawn Tashlhiyt [4]

Generalization (1/2)

Sibilants must agree in anteriority and voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ$\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, *ʒʃ, *ʒz $\}$

[4] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Sibilant Harmony in Imdlawn Tashlhiyt [5]

Generalization (2/2)

Voiceless obstruents block agreement in voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ, q$\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, * ʒʃ, *ʒz $\}$

^{ok} ʃ q u ʒ: i   * s q u ʒ: i

[5] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Sibilant Harmony in IMDLAWN TASHLHIYT [5]

Generalization (2/2)

Voiceless obstruents block agreement in voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ, q$\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, * ʒʃ, *ʒz $\}$

ʃ q ʒː

ok ʃ q u ʒː i

* s q u ʒː i

[5] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

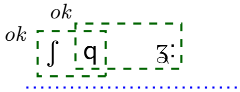
Sibilant Harmony in Imdlawn Tashlhiyt [5]

Generalization (2/2)

Voiceless obstruents block agreement in voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ, q$\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, * ʒʃ, *ʒz $\}$



ok ʃ q u ʒː i

* s q u ʒː i

[5] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Sibilant Harmony in IMDLAWN TASHLHIYT [5]

Generalization (2/2)

Voiceless obstruents block agreement in voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ, q$\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, * ʒʃ, *ʒz $\}$

[5] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

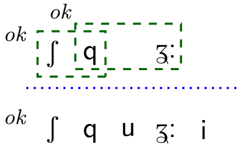
Sibilant Harmony in IMDLAWN TASHLHIYT [5]

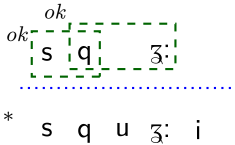
Generalization (2/2)

Voiceless obstruents block agreement in voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ, q$\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, * ʒʃ, *ʒz $\}$





[5] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

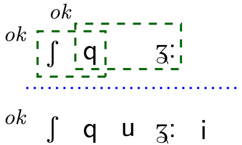
Sibilant Harmony in Imdlawn Tashlhiyt [5]

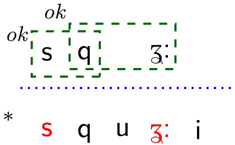
Generalization (2/2)

Voiceless obstruents block agreement in voicing.

Grammar

$T = \{$ ʒ, s, z,ʃ, q$\}$
$S = \{$ *sʒ, *sz, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʃʒ, * ʒʃ, *ʒz $\}$





[5] Elmedlaoui (1995), Hansson (2010), McMullin (2016), De Santo (2018)

Multi-Tier Strictly Local (MTSL) Languages (1/2)[6]

Sibilant Harmony in IMDLAWN TASHLHIYT (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$

^{ok} ʃ q u ʒ: i

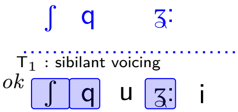
[6] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (1/2)[6]

Sibilant Harmony in IMDLAWN TASHLHIYT (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$



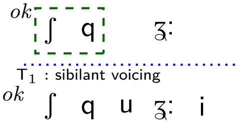
[6] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (1/2)[6]

Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$



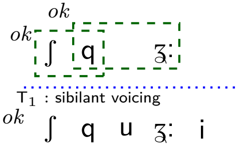
[6] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (1/2)[6]

Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$



[6] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (1/2)[6]

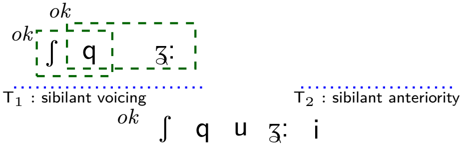
Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$

Unbounded agreement in anteriority:

- $T_2 = \{$ʒ, s, z,ʃ$\}$ $S_2 = \{$*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz $\}$



[6] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (1/2)[6]

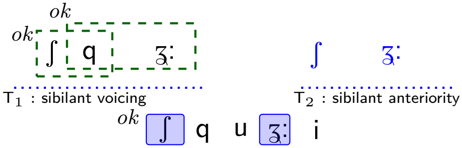
Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$

Unbounded agreement in anteriority:

- $T_2 = \{$ʒ, s, z,ʃ$\}$ $S_2 = \{$*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz $\}$



[6] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (1/2)[6]

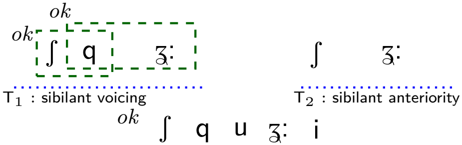
Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$

Unbounded agreement in anteriority:

- $T_2 = \{$ʒ, s, z,ʃ$\}$ $S_2 = \{$*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz $\}$



[6] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (1/2)[6]

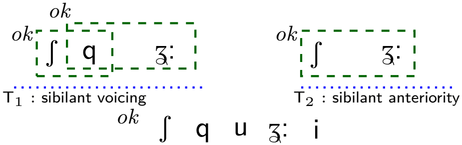
Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$

Unbounded agreement in anteriority:

- $T_2 = \{$ʒ, s, z,ʃ$\}$ $S_2 = \{$*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz $\}$



[6] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (2/2)[7]

Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$

Unbounded agreement in anteriority:

- $T_2 = \{$ʒ, s, z,ʃ$\}$ $S_2 = \{$*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz $\}$

* s q u ʒ: i

[7] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (2/2)[7]

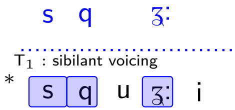
Sibilant Harmony in IMDLAWN TASHLHIYT (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$

Unbounded agreement in anteriority:

- $T_2 = \{$ʒ, s, z,ʃ$\}$ $S_2 = \{$*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz $\}$



[7] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (2/2)[7]

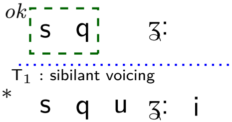
Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$

Unbounded agreement in anteriority:

- $T_2 = \{$ʒ, s, z,ʃ$\}$ $S_2 = \{$*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz $\}$



[7] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (2/2)[7]

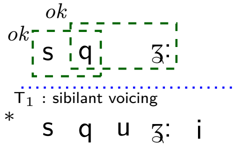
Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 = \{$ʒ, s, z,ʃ, q$\}$ $S_1 = \{$*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ$\}$

Unbounded agreement in anteriority:

- $T_2 = \{$ʒ, s, z,ʃ$\}$ $S_2 = \{$*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz $\}$



[7] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (2/2)[7]

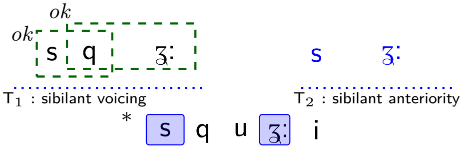
Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- $T_1 =$ {ʒ, s, z,ʃ, q} $S_1 =$ {*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ}

Unbounded agreement in anteriority:

- $T_2 =$ {ʒ, s, z,ʃ} $S_2 =$ {*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz }



[7] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

Multi-Tier Strictly Local (MTSL) Languages (2/2)[7]

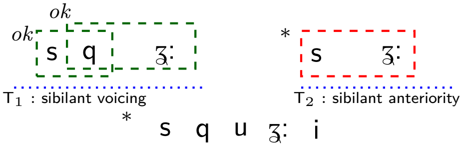
Sibilant Harmony in Imdlawn Tashlhiyt (Revisited)

Voiceless obstruents block agreement in voicing:

- T_1 = {ʒ, s, z,ʃ, q} S_1 = {*sʒ, *sz, *ʒs, *zs, *ʃz, *ʃʒ, *ʒʃ}

Unbounded agreement in anteriority:

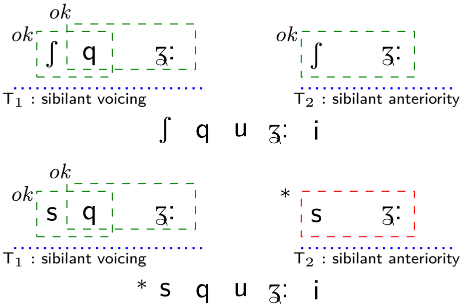
- T_2 = {ʒ, s, z,ʃ} S_2 = {*sʒ, *sʃ, *ʒs,*ʃs, *zs, *zʃ, *zʒ, *ʃz, *ʒz }



[7] McMullin (2016), De Santo (2018), De Santo & Graf (2019)

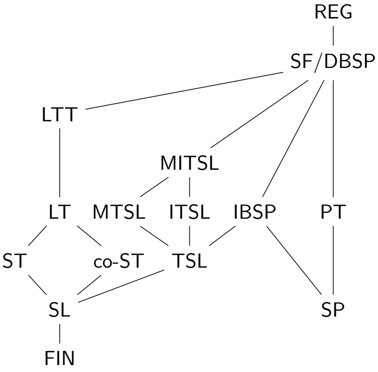
Accounting for Concurrent Processes

- MTSL: TSL closure under intersection
  (De Santo & Graf, 2019)



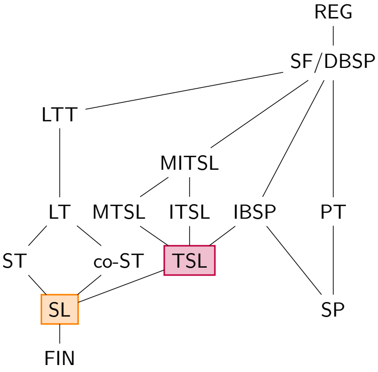
- Intersection closure accounts for multiple concurrent processes
- Can characterize the complete phonotactics of a language

A Plethora of Combination (De Santo & Graf 2019)



- The goal is **not** identifying a single “correct” class
- Pinpoint fundamental properties of the patterns:
  SL: $\triangleleft$, TSL: $\triangleleft_T$, . . .

A Plethora of Combination (De Santo & Graf 2019)

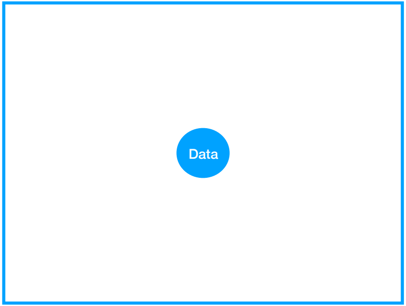


- The goal is **not** identifying a single "correct" class
- Pinpoint fundamental properties of the patterns:
  SL: $\triangleleft$, TSL: $\triangleleft_T$, ...
- What about learnability?

Learning Multiple TSL Grammars [8]

Problem:

- Unrestricted Hypothesis Spaces

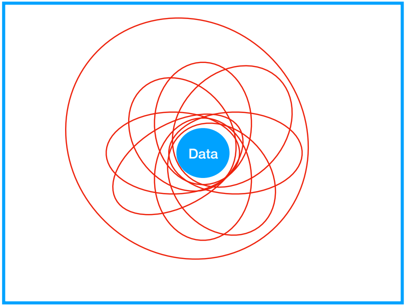


[8] McMullin, Aksenova, De Santo (2020), De Santo & Aksenova (2021)

Learning Multiple TSL Grammars [9]

Problem:

- Unrestricted Hypothesis Spaces

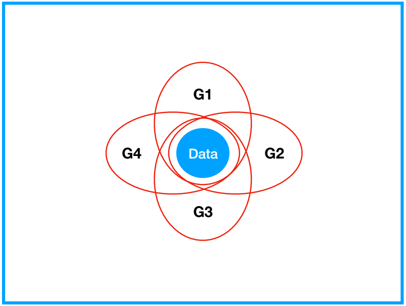


[9] McMullin, Aksenova, De Santo (2020), De Santo & Aksenova (2021)

Learning Multiple TSL Grammars [10]

Solution:

- Structural priors



De Santo & Aksenova (2021):

$\Rightarrow$ Assume relativized locality!

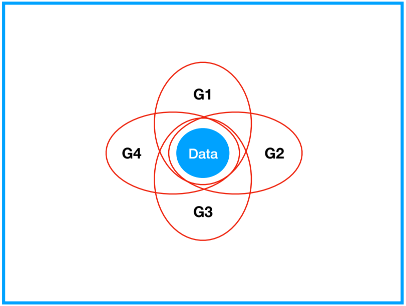
- tiers (but not their content)
- local tier constraints
- characteristic sample!

[10] McMullin, Aksenova, De Santo (2020), De Santo & Aksenova (2021)

Learning Multiple TSL Grammars [10]

Solution:

- Structural priors



De Santo & Aksenova (2021):

⇒ Assume relativized locality!

- tiers (but not their content)
- local tier constraints
- characteristic sample!

Results

- No a priori information on the content of tiers/constraints
- Guaranteed convergence in polynomial time and data

[10] McMullin, Aksenova, De Santo (2020), De Santo & Aksenova (2021)

Evaluating Convergence in Real World Scenarios

	SP	SL	TSL	MTSL	MITSL
Word-final devoicing					
T	✗	✓	✓	✓	✓
A	68%	100%	100%	100%	100%
N_1	58%	100%	100%	100%	100%
Single vowel harmony without blocking					
T	✓	✗	✓	✓	✓
A	100%	83%	100%	100%	100%
N_2	100%	72%	100%	100%	100%
Single vowel harmony with blocking					
T	✗	✗	✓	✓	✓
A	84%	89%	100%	100%	99%
Several vowel harmonies without blocking					
T	✓	✗	✓	✓	✓
A	100%	69%	100%	100%	100%
Several vowel harmonies with blocking					
T	✗	✗	✓	✓	✓
A	76%	59%	100%	100%	99%
N_3	76%	70%	67%	95%	99%
Vowel harmony and consonant harmony without blocking					
T	✓	✗	✗	✓	✓
A	100%	64%	74%	100%	100%
Vowel harmony and consonant harmony with blocking					
T	✗	✗	✗	✓	✓
A	83%	64%	69%	100%	100%
Unbounded tone plateauing					
T	✓	✗	✗	✗	✓
A	100%	85%	90%		100%
Two locally-driven long-distance assimilations (ITSL restrictions)					
T	✗	✗	✗	✗	✓
A					100%

Johnson & De Santo (2023)

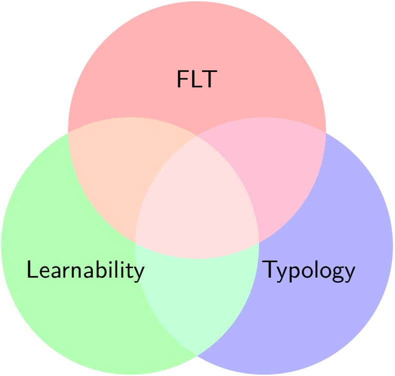
(T)heoretical expectations and performance of 5 subregular learners on (A)rtificial and simplified (N)atural language input data-sets. N_1: German; N_1: Finnish; N_1: Turkish.

Interim Summary

SL and TSL for Spoken Phonotactics

- Linguistically natural (Goldsmith 1976)
- Captures (properties of a) wide range of (spoken) phonotactic dependencies (McMullin 2016, De Santo & Graf 2019)
  What about sign? (Rawski 2017, Rawski forth.)
- Provably correct and efficient learning algorithms
  (De Santo & Aksenova 2021, Johnson & De Santo u.r.)
- Predictions for human learning
  (Lai 2015, Avcu & Hestevic 2021, De Santo & Gutierrez in prep.)
- Generalizes beyond phonotactics
  (Aksenova & De Santo 2017, Graf & De Santo 2019, a.o.)

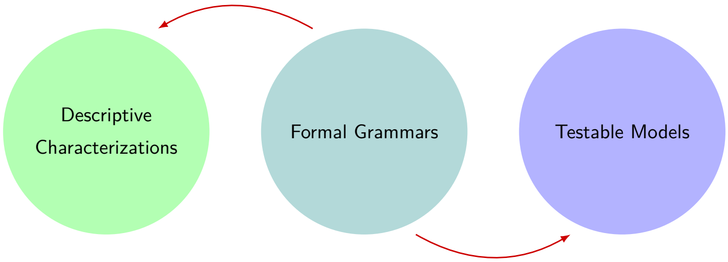
Interim Summary [cont.]



The Tip of the Iceberg:

- Cross-linguistic/cross-domain typological analysis
- Artificial language learning experiments
- New algorithms
- New mathematical insights

Building Bridges



Outline

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

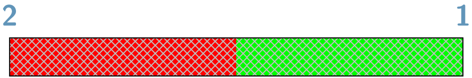
1. What do you think that John bought *t*?
2. *What do you wonder whether John bought *t*?

> *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 ≈ Grammaticality judgments

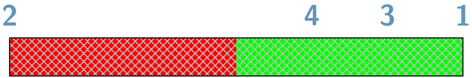
Gradience in Acceptability Judgments

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)*

(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!

Building Linking Hypothesis

We need to link categorical grammars, processing difficulty, and gradience **explicitly**!

(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!

Building Linking Hypothesis

We need to link categorical grammars, processing difficulty, and gradience **explicitly**!

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

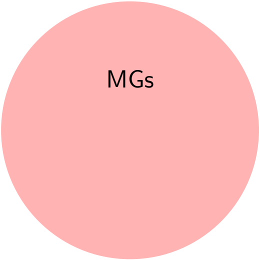
- What is the relation between grammatical operations and cognitive processes?

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.

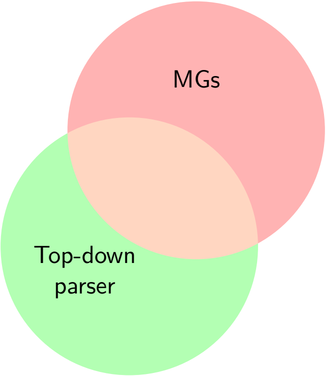
- What is the right notion of syntactic derivation?
- What is costly? And why?

A Formal Model of Sentence Processing



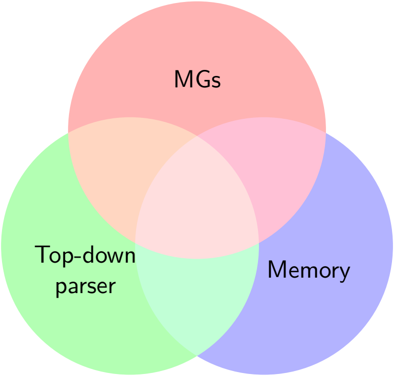
1. An explicit syntactic theory → Minimalist grammars (MGs)

A Formal Model of Sentence Processing



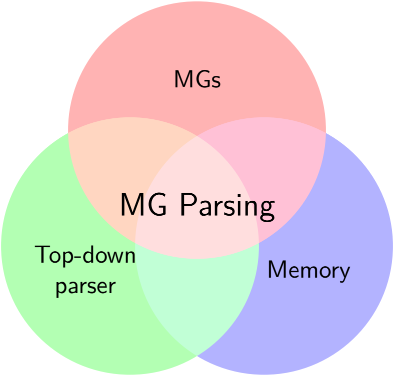
1. An explicit syntactic theory → Minimalist grammars (MGs)
2. A theory of how structures are built → Top-down parser

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

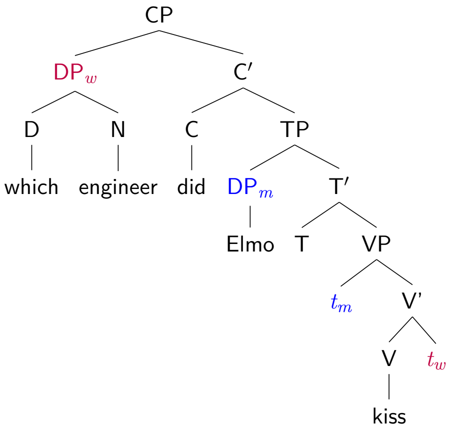
A Formal Model of Sentence Processing



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

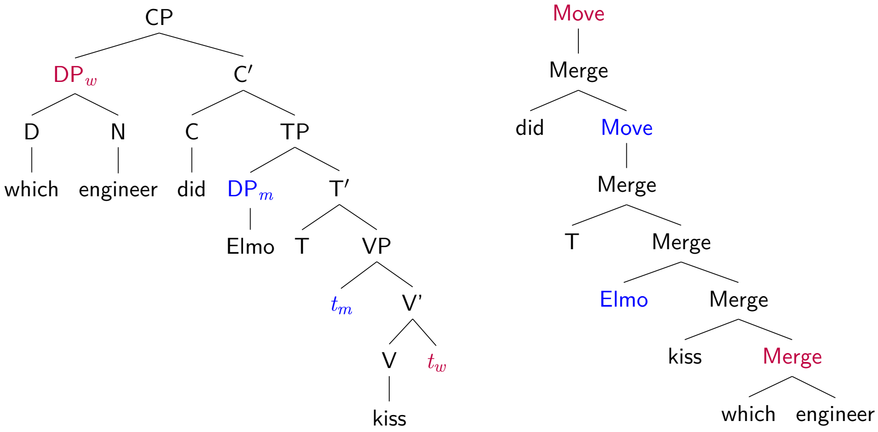
Interpretability for the win!

Minimalist Grammars (MGs) & Derivation Trees



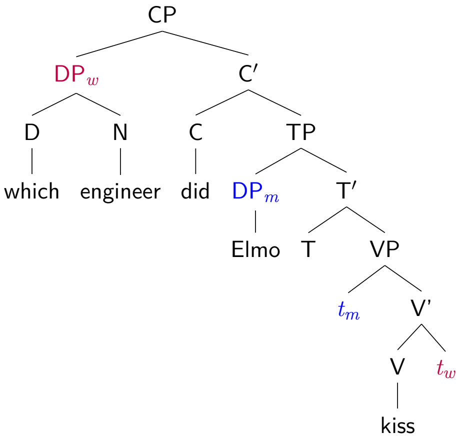
Phrase Structure Tree

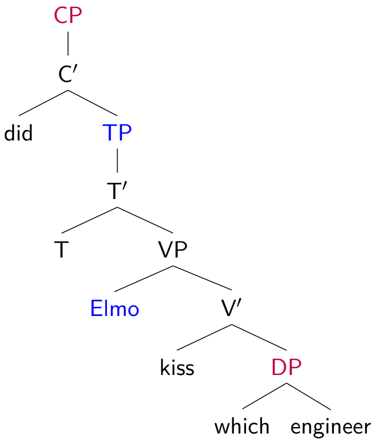
Minimalist Grammars (MGs) & Derivation Trees



Phrase Structure Tree **Derivation Tree**

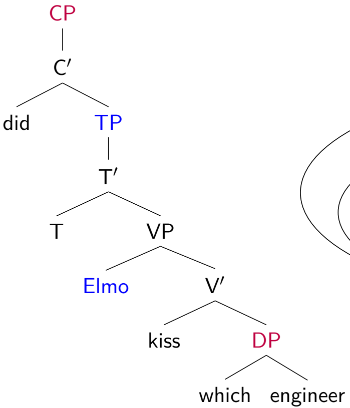
MG Syntax: Derivation Trees

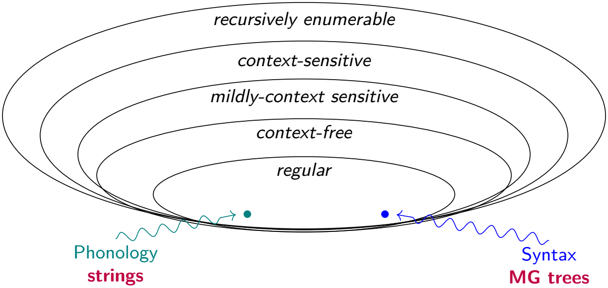




Phrase Structure Tree **Derivation Tree**

Regular Tree Languages[11]

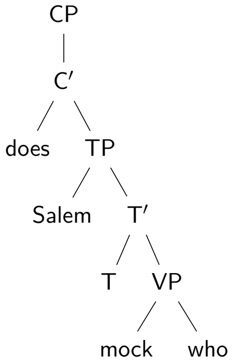




[11] Thatcher (1967), Kobele et al. (2007), Stabler (2013)

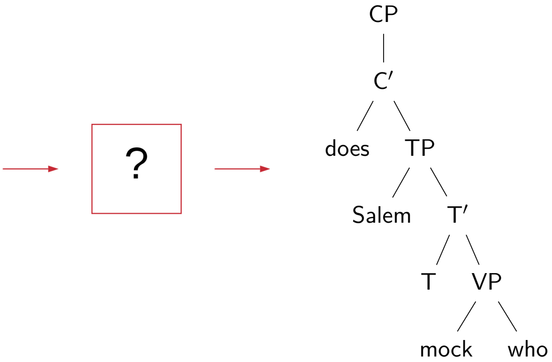
The Job of a Parser

Who does Salem mock?

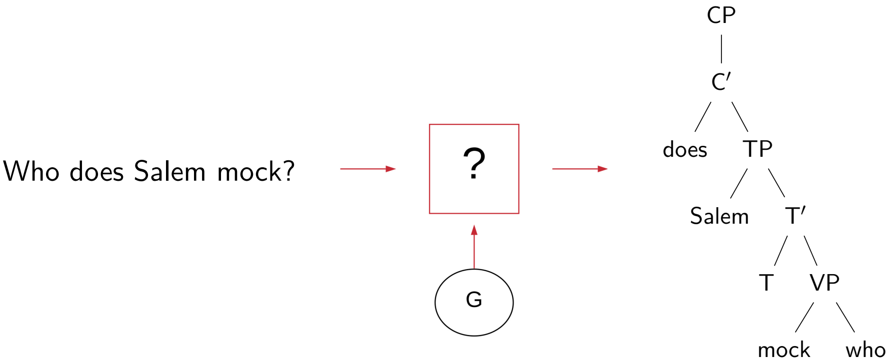


The Job of a Parser

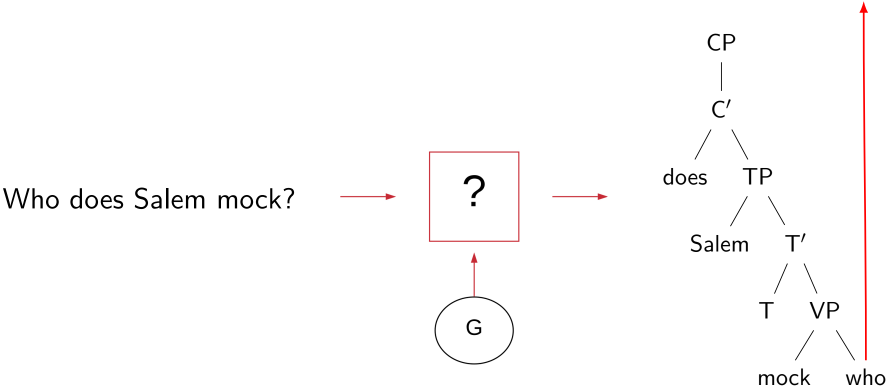
Who does Salem mock?



The Job of a Parser

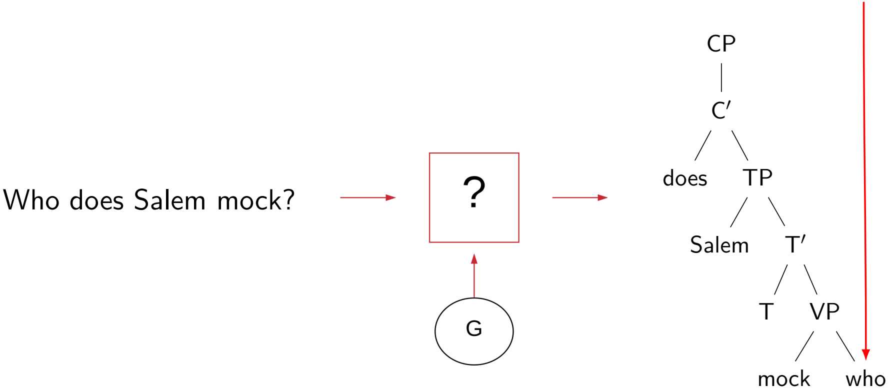


The Job of a Parser



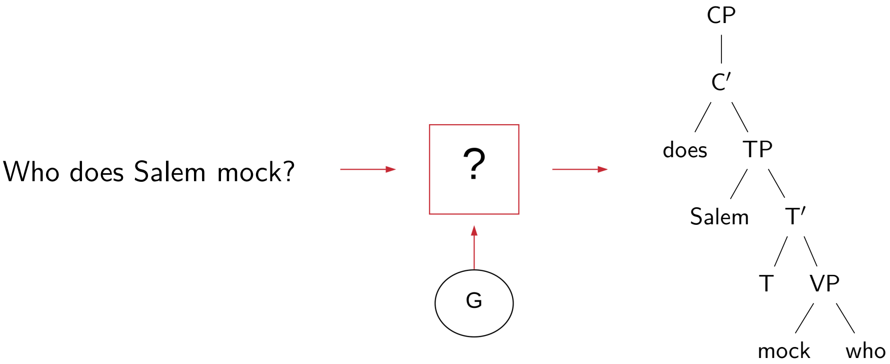
- Bottom-up

The Job of a Parser



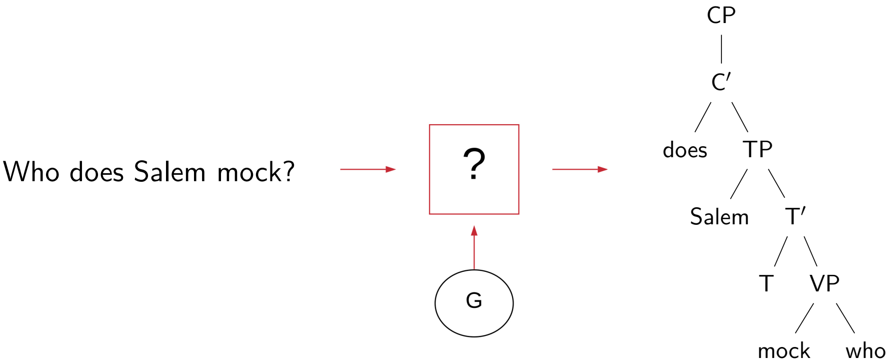
- Bottom-up
- Top-down

The Job of a Parser



- Bottom-up
- Top-down (Stabler, 2013)
  - Psychologically plausible(-ish)

The Job of a Parser

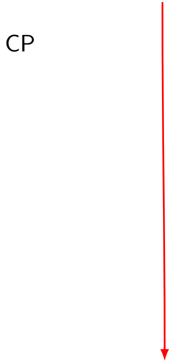


- Bottom-up
- Top-down (Stabler, 2013)
  - Psychologically plausible(-ish)
  - Assumption: Parser as an oracle!

Incremental Top-Down Parsing: The Intuition

Who does Salem mock?

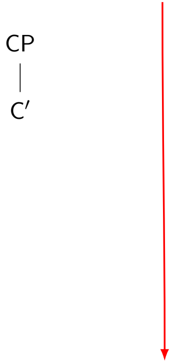
Incremental Top-Down Parsing: The Intuition



Who does Salem mock?

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

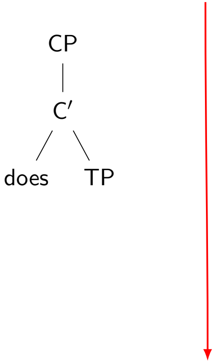
Incremental Top-Down Parsing: The Intuition



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!

Incremental Top-Down Parsing: The Intuition

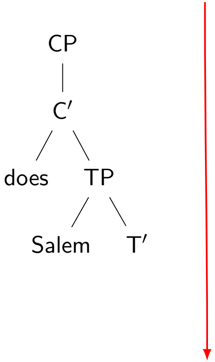


Who does Salem mock?

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

Incremental Top-Down Parsing: The Intuition

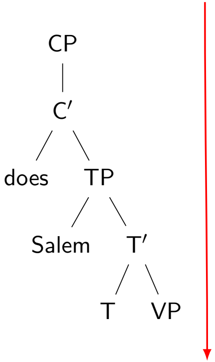
Who does Salem mock?



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

Incremental Top-Down Parsing: The Intuition

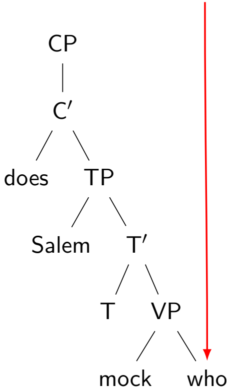
Who does Salem mock?



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

Incremental 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!

Memory-Based Complexity Metrics

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

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!

- Formalized into offline **complexity metrics**

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

Memory-Based Complexity Metrics

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

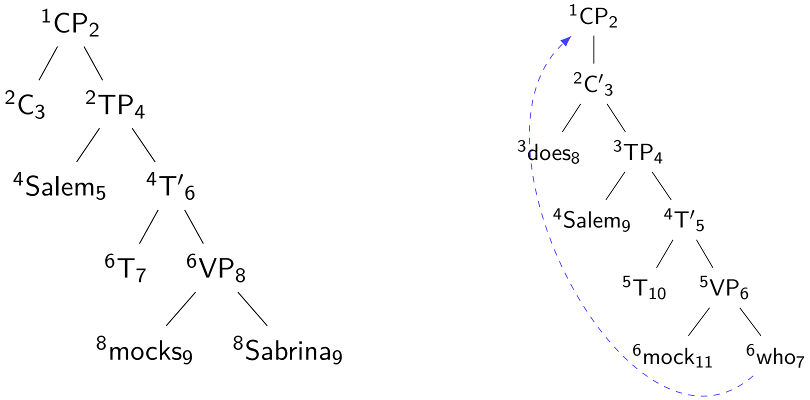
Tenure How long a node is kept in memory
Size How much information is stored in a node
$\Rightarrow$ Intuitively, the length of its movement dependency!

- Formalized into offline **complexity metrics**

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

Contrasting Derivations

MaxTenure $= 2$ **MaxTenure** $= 5$



Summary of the Approach

A Computational Linking Hypothesis (De Santo 2020; in prep.)

Grammar $\rightleftarrows$ MG Parser Effort $\rightleftarrows$ Gradience

General Idea

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

1. Pick two competing derivations
2. Evaluate metrics over each
   - Lowest score means easiest!
3. Compare parser's prediction to experimental data

A Proof of Concept: Back to 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: Back to 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: Back to 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: Back to 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)

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*?

Gap Position × 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

Processing Asymmetries All the Way Down

A variety of processing insights!

Across Many Constructions

- Right > center embedding (Kobele et al. 2012)
- Crossing > nested dependencies (Kobele et al. 2012)
- SRC > ORC
  (Graf et al. 2017; De Santo 2020; Fiorini, Chang, De Santo 2023)
- Priming/Stacked RCs (De Santo 2020, 2022)
- Postverbal subjects
  (De Santo 2019, 2021; Del Valle & De Santo 2023)
- Persian attachment ambiguities (De Santo & Shafiei 2019)
- RC attachment preferences
  (De Santo & Lee 2022; Lee & De Santo 2023)

Across Languages

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

Summary

Gradience from a categorical MG grammar?

- The **first** (quantitative) model of this kind!
- Overall, a success! ⇒ **just** from structural differences!
- Outlier is expected (and makes predictions!)

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 (and makes predictions!)

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 (and makes predictions!)

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)

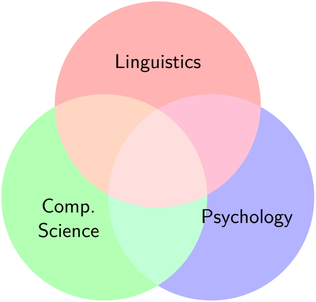
Outline

From the Trees (back) to the Forest

> *[...] this is a confusion of two quite separate issues, **simulation and explanation**. [...] What we are **really** interested in [...] is explanation — in developing models that help us **understand how it is that people behave** that way, not merely demonstrating that we can build an artifact that behaves similarly.*
>
> (Kaplan, 1995)

- Invariant properties of phenomena
- Implementations of verbal theories

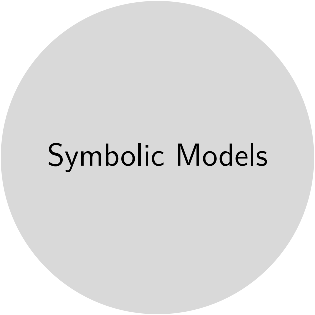
Embracing Multidisciplinarity



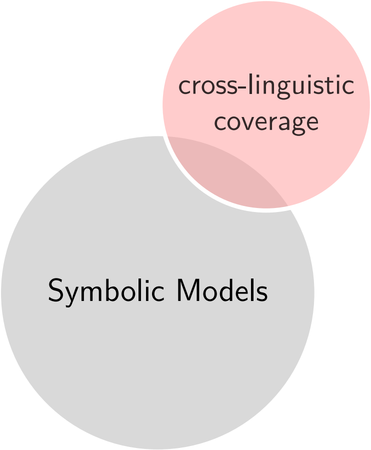
Within the program of research proposed here, joint work by linguists, computer scientists, and psychologists could lead to a deeper scientific understanding of the role of language in cognition.

(Bresnan 1978: pg. 59)

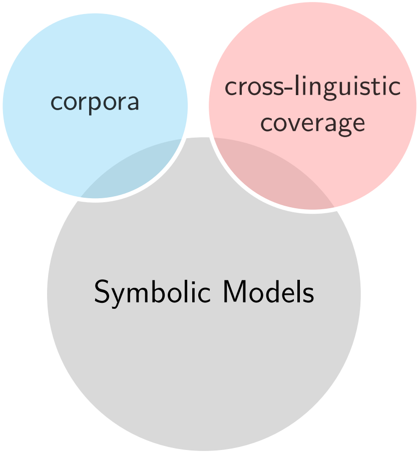
Looking Ahead: A Collaborative Enterprise



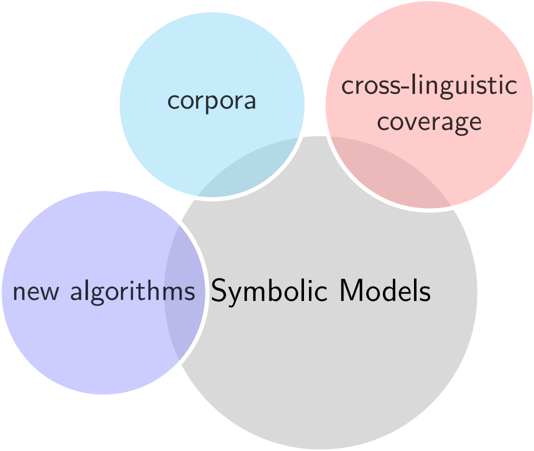
Looking Ahead: A Collaborative Enterprise



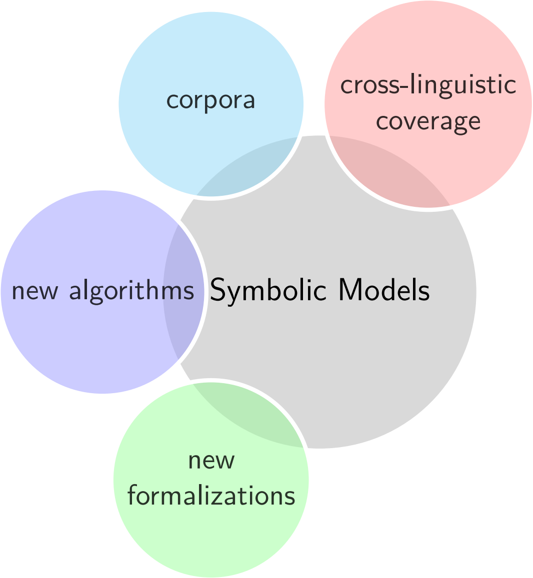
Looking Ahead: A Collaborative Enterprise



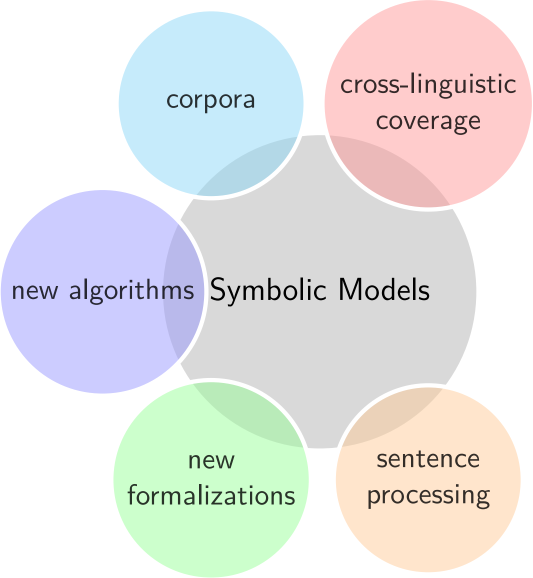
Looking Ahead: A Collaborative Enterprise



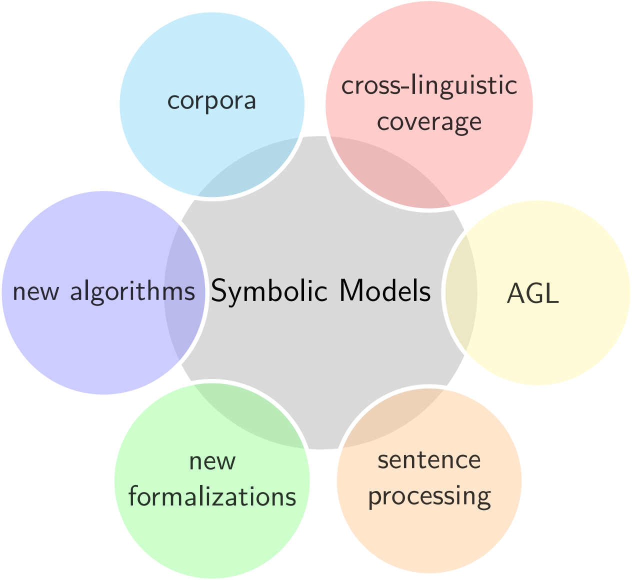
Looking Ahead: A Collaborative Enterprise



Looking Ahead: A Collaborative Enterprise



Looking Ahead: A Collaborative Enterprise



Looking Ahead: A Collaborative Enterprise

Looking Ahead: A Collaborative Enterprise

Selected References I

1 Applegate, R.B. 1972. Ineseno chumash grammar. Doctoral Dissertation, University of California,Berkeley.

2 **De Santo, A.** (2019). Testing a Minimalist grammar 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. In *Proceedings of SCiL 2020*, Jan 2-5, New Orleans.

4 **De Santo, A.** & Aksenova, A. (2021). Learning Interactions of Local and Non-Local Phonotactic Constraints from Positive Input. In Proceedings of SCiL

5 **De Santo, A.** and Graf, T. 2019. Structure sensitive tier projection: Applications and formal properties. Proceedings of FG.

6 **De Santo, A.** & Rawski, J. (2022). Mathematical Linguistics and Cognitive Complexity. In Handbook of Cognitive Mathematics (pp. 1-38).

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

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

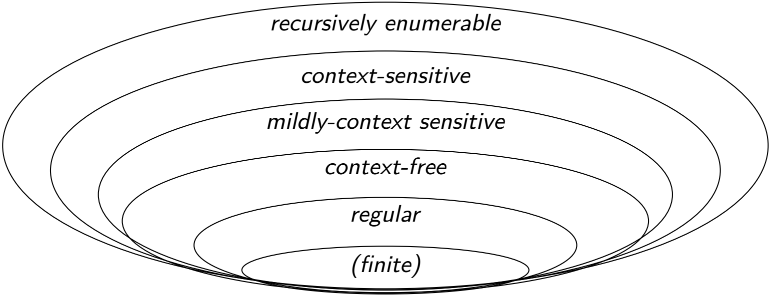
9 **Stabler, E.P.** (2013). Bayesian, minimalist, incremental syntactic analysis. *Topics in Cognitive Science* 5:611–633.

Appendix

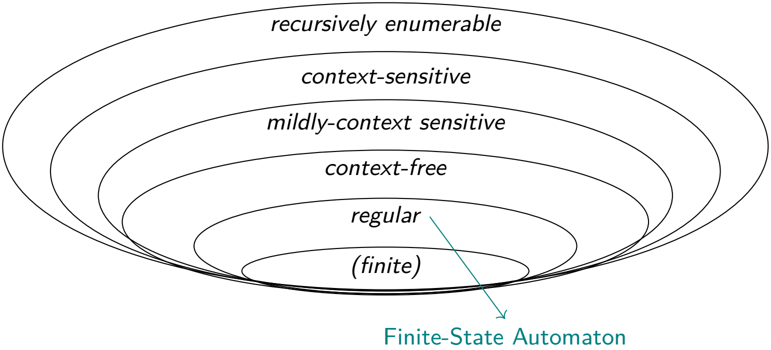
Kaplan's Full Quote

> *[...] this is a confusion of two quite separate issues, simulation and explanation. As scientists, we are not merely interested in simulating human behavior [...] What we are really interested in [...] is explanation — in developing models that help us* ***understand how it is that people behave*** *that way, not merely demonstrating that we can build an artifact that behaves similarly. [...] We should look for modular theories that account for the observed interactions in terms of the interleaving of information from separate, scientifically comprehensible systems*
>
> *(Kaplan, 1995)*

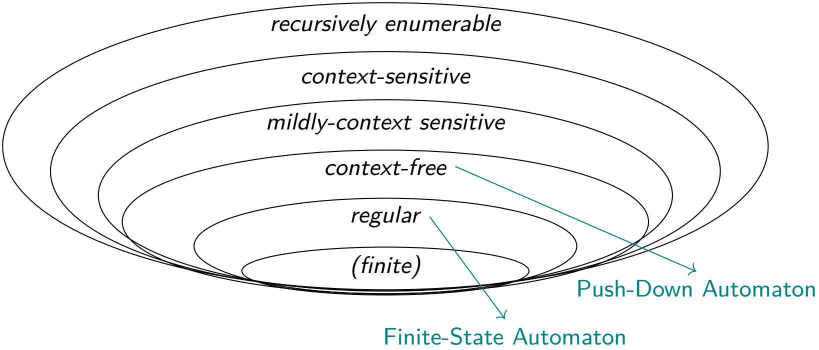
Chomsky Hierarchy and Automata Theory



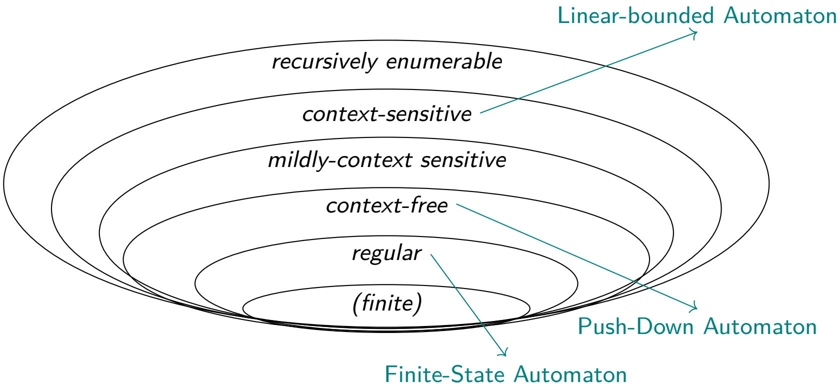
Chomsky Hierarchy and Automata Theory



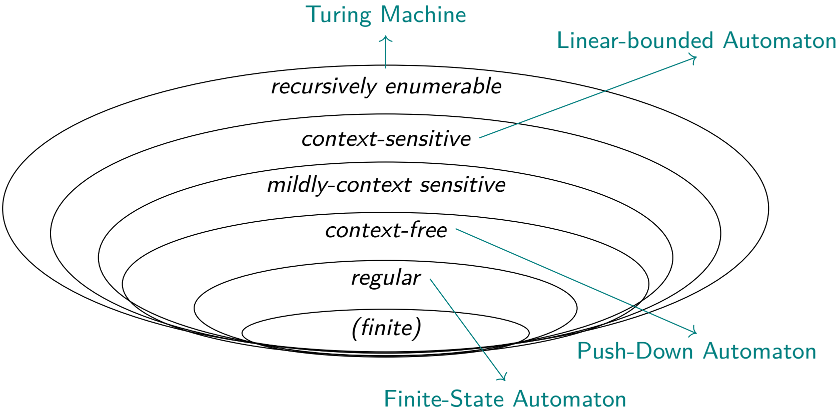
Chomsky Hierarchy and Automata Theory



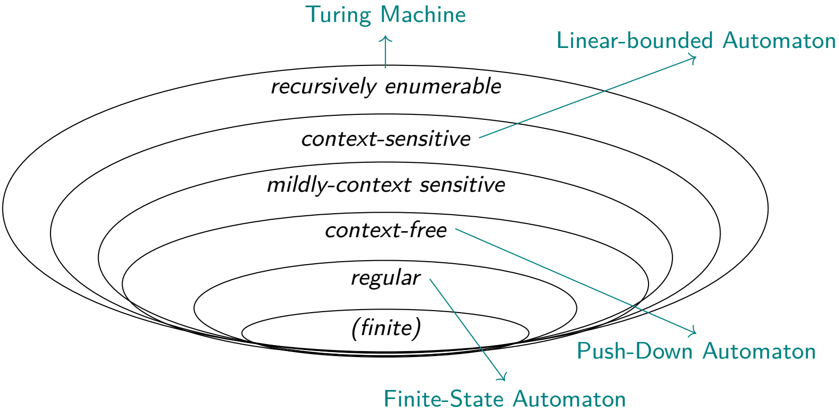
Chomsky Hierarchy and Automata Theory



Chomsky Hierarchy and Automata Theory



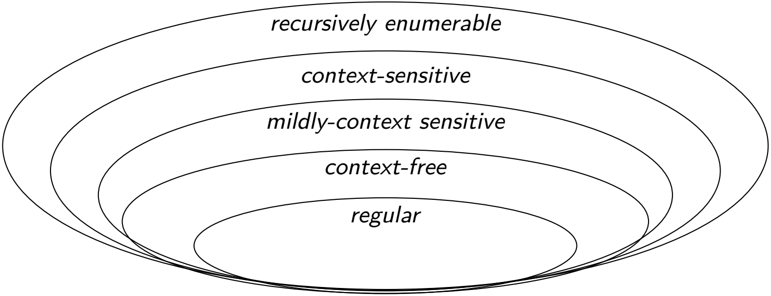
Chomsky Hierarchy and Automata Theory



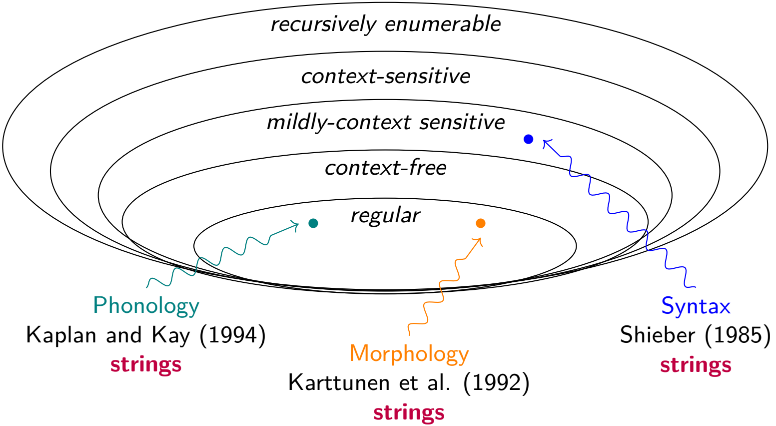
Automata theoretic classes seem to presuppose [...] specific classes of recognition mechanisms, raising questions about whether these are necessarily relevant to the cognitive mechanisms under study.

Rogers & Pullum 2011

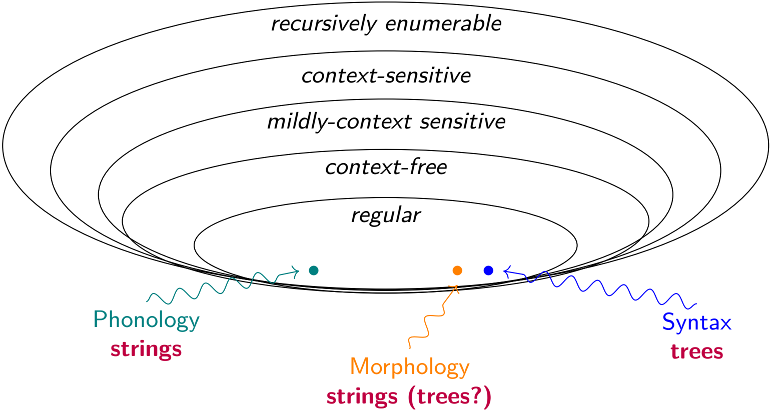
Cross-domain Parallels



Cross-domain Parallels



Cross-domain Parallels



Some Insights

Parallels between phonology and syntax?

- What would a computational linguist tell you?
  Well, it depends!
- What will I show you?
  They are fundamentally similar!

The Take-Home Message

- **Two kind of dependencies**: local and non-local
- The core mechanisms are **the same** cross-domain
- That is: linguistic dependencies are **local** over the right **structural representations**

Some Insights

Parallels between phonology and syntax?

- What would a computational linguist tell you?
  Well, it depends!
- What will I show you?
  They are fundamentally similar!

The Take-Home Message

- **Two kind of dependencies**: local and non-local
- The core mechanisms are **the same** cross-domain
- That is: linguistic dependencies are **local** over the right **structural representations**

Some Insights

Parallels between phonology and syntax?

- What would a computational linguist tell you?
  Well, it depends!
- What will I show you?
  They are fundamentally similar!

The Take-Home Message

- **Two kind of dependencies**: local and non-local
- The core mechanisms are **the same** cross-domain
- That is: linguistic dependencies are **local** over the right **structural representations**

What about Syntax?

We need a model for syntax ...

- Minimalist grammars (MGs) are a formalization of Minimalist syntax. (Stabler 1997, 2011)
- Operations: **Merge** and **Move**
- Adopt Chomsky-Borer hypothesis:
  Grammar is just a finite list of feature-annotated lexical items

Local dependencies in syntax

- Merge is a **feature-driven** operation:
  category feature N^-, D^-, ...
  selector feature N^+, D^+, ...
- Subcategorization as formalized by Merge is **strictly local**.

What about Syntax?

We need a model for syntax ...

- Minimalist grammars (MGs) are a formalization of Minimalist syntax. (Stabler 1997, 2011)
- Operations: **Merge** and **Move**
- Adopt Chomsky-Borer hypothesis:
  Grammar is just a finite list of feature-annotated lexical items

Local dependencies in syntax

- Merge is a **feature-driven** operation:
  category feature N^-, D^-, ...
  selector feature N^+, D^+, ...
- Subcategorization as formalized by Merge is **strictly local**.

Local Dependencies in Syntax

Merge is a **feature-driven** operation:

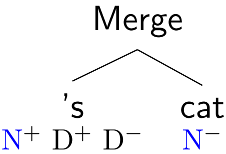
- category feature N^-, D^-, ...
- selector feature N^+, D^+, ...

's	cat
N^+ D^+ D^-	N^-

Local Dependencies in Syntax

Merge is a **feature-driven** operation:

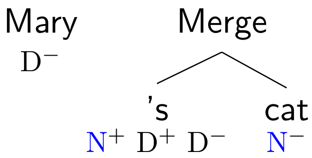
- category feature N^-, D^-, ...
- selector feature N^+, D^+, ...



Local Dependencies in Syntax

Merge is a **feature-driven** operation:

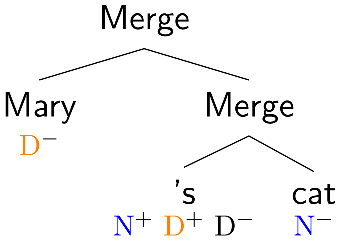
- category feature N^-, D^-, ...
- selector feature N^+, D^+, ...



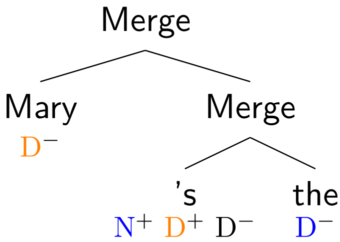
Local Dependencies in Syntax

Merge is a **feature-driven** operation:

- category feature N^-, D^-, ...
- selector feature N^+, D^+, ...

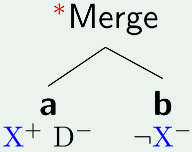


Merge is SL (Graf 2012)

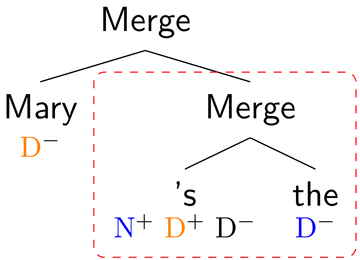


SL constraints on Merge

- We lift constraints from string n-grams to tree n-grams
- We get SL constraints over subtrees.

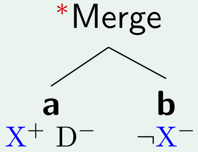


Merge is SL (Graf 2012)



SL constraints on Merge

- We lift constraints from string n-grams to tree n-grams
- We get SL constraints over subtrees.



Interim Summary

	Local	**Data Structure**
Phonology	?	?
Syntax	?	?

Local phenomena modeled by n-grams of bounded size:

- computationally very simple
- learnable from positive examples of strings/trees
- plausible cognitive requirements

Interim Summary

	Local	**Data Structure**
Phonology	**SL**	**Strings**
Syntax	**SL**	**Trees**

Local phenomena modeled by n-grams of bounded size:

- computationally very simple
- learnable from positive examples of strings/trees
- plausible cognitive requirements

Interim Summary

	Local	**Non-local**	**Data Structure**
Phonology	**SL**	?	**Strings**
Syntax	**SL**	?	**Trees**

Local phenomena modeled by n-grams of bounded size:

- computationally very simple
- learnable from positive examples of strings/trees
- plausible cognitive requirements

TSL Phonology: Accounting for Context

- **Unbounded Tone Plateauing in Luganda** (UTP)
  No L may occur within an interval spanned by H.
  (Hyman 2011)

(7)
a. LHLLLL
b. LLLLHL
c. *LHLLHL
d. LHHHHL

Example

TSL Phonology: Accounting for Context

- **Unbounded Tone Plateauing in Luganda** (UTP)
  No L may occur within an interval spanned by H.
  (Hyman 2011)

(7) a. **LHLLLL**

b. **LLLLHL**

c. * **LHLLHL**

d. **LHHHHL**

Example

TSL Phonology: Accounting for Context

- **Unbounded Tone Plateauing in Luganda** (UTP)
  No L may occur within an interval spanned by H.
  (Hyman 2011)

(7)
a. **LHLLLL**
b. **LLLLHL**
c. * **LHLLHL**
d. **LHHHHL**

Example

***L H L L H L**

TSL Phonology: Accounting for Context

- **Unbounded Tone Plateauing in Luganda** (UTP)
  No L may occur within an interval spanned by H.
  (Hyman 2011)

(7)
a. LHLLLL
b. LLLLHL
c. *LHLLHL
d. LHHHHL

Example

L H L L H L

....

*L H L L H L

TSL Phonology: Accounting for Context

- **Unbounded Tone Plateauing in Luganda** (UTP)
  No L may occur within an interval spanned by H.
  (Hyman 2011)

(7)
a. **LHLLLL**
b. **LLLLHL**
c. * **LHLLHL**
d. **LHHHHL**

Example

L H L L H L

** L H L L H L*

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

^{ok}**L H L L L L** ***L H L L H L**

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H

ok **L H** L L L L

* L H L L H L

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H L

. . . .

^{ok}**L H L L L L** $\quad$ ***L H L L H L**

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H L

^{ok}**L H L L L L**

***L H L L H L**

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H L

....

^{ok}**L H L L L L**

***L H L L H L**

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H L

^{ok}**L H L L L L** ***L H L L H L**

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H L

^{ok}L H L L L L

*L H L L H L

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

HL

....

^{ok} L H L L L L

....

* L H L L H L

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

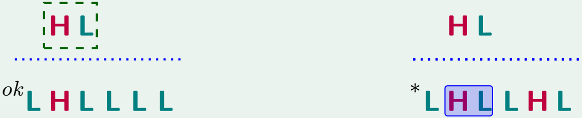
H L	**H**
^{ok}**L H L L L L**	***L H L L H L**

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example



Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H L
^{ok}L H L L L L

H L
*L H L L H L

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H L	**H L H**
^{ok}**L H L L L L**	***L H L L H L**

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H L	H L H L
^{ok}L H L L L L	*L H L L H L

Accounting for Context [cont.]

A TSL analysis for UTP (De Santo and Graf 2017):

- Project every **H**; project **L** iff immediately follows **H**
- Ban: **HLH**

Example

H L	**H L H L**
^{ok}**L H L L L L**	***L H L L H L**

TSL for Phonology

Most non-local dependencies in phonology are TSL

- Linguistically natural (Goldsmith 1976)
- Captures wide range of phonotactic dependencies (McMullin 2016)
- Provably correct and efficient learning algorithms (Jardine and McMullin 2017)
- Rules out unattested patterns (cf. Lai 2015, Aksenova et al. 2016, Graf & De Santo 2019, a.o.)

What about syntax?

TSL for Phonology

Most non-local dependencies in phonology are TSL

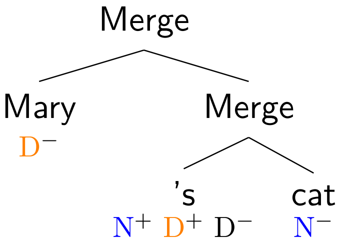
- Linguistically natural (Goldsmith 1976)
- Captures wide range of phonotactic dependencies (McMullin 2016)
- Provably correct and efficient learning algorithms (Jardine and McMullin 2017)
- Rules out unattested patterns (cf. Lai 2015, Aksenova et al. 2016, Graf & De Santo 2019, a.o.)

What about syntax?

Non-Local Dependencies in Syntax

Let's stick to core operations:

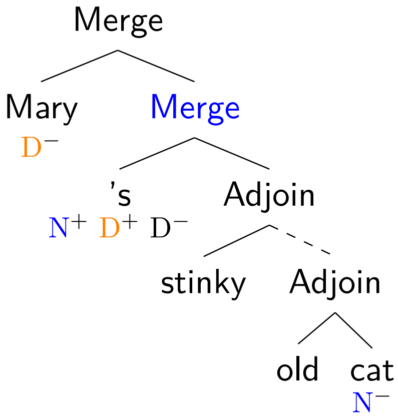
- Move
- Merge?



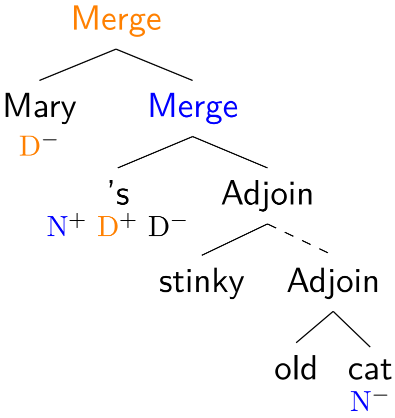
Non-Local Dependencies in Syntax

Let's stick to core operations:

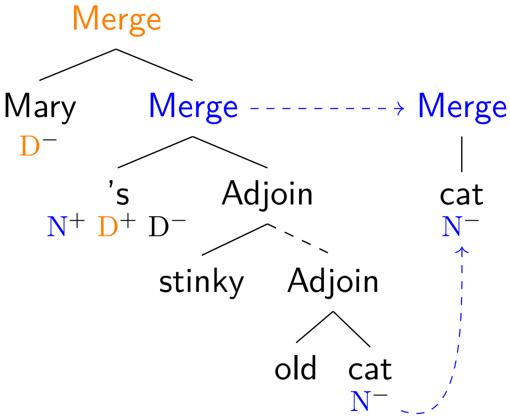
- Move
- **Merge**: Unbounded adjunction
  Frey and Gärtner (2002); Graf (2017)



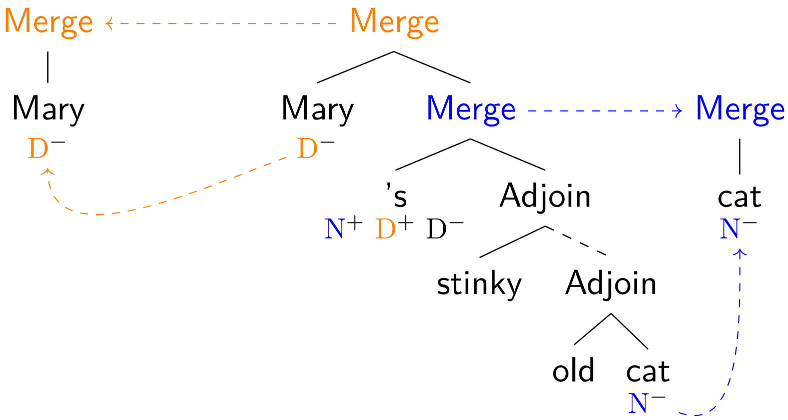
TSL over Trees: Projecting Tiers



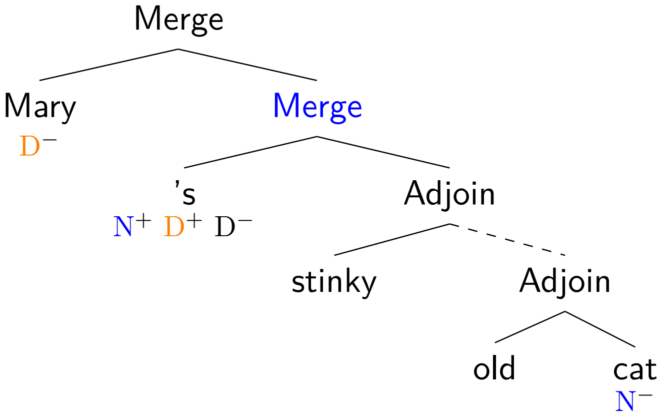
TSL over Trees: Projecting Tiers



TSL over Trees: Projecting Tiers

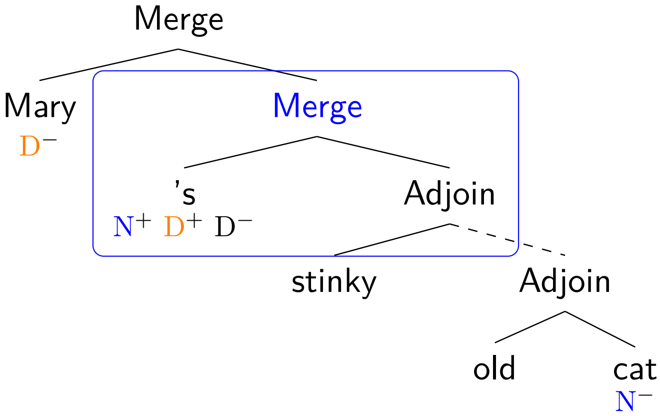


Merge with Adjunction is TSL



A TSL grammar for Merge

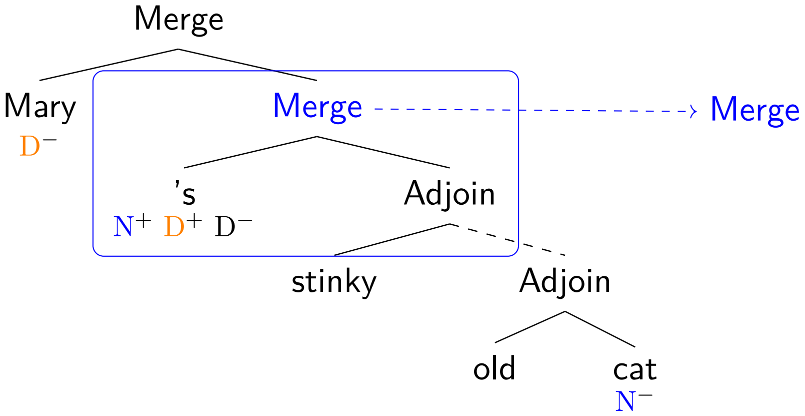
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^{+} (e.g. $X = N$)

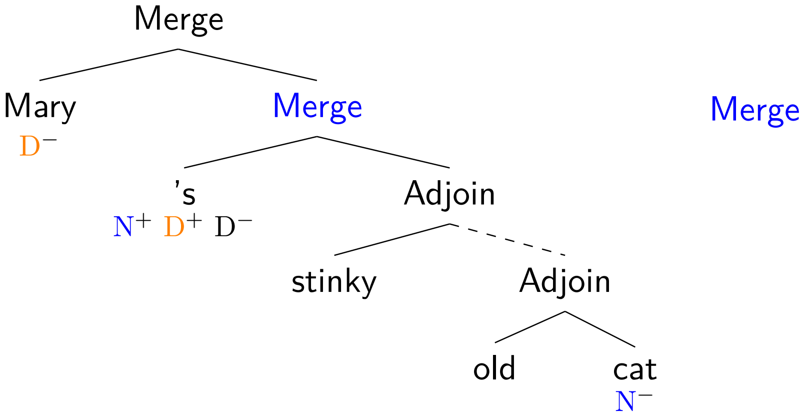
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^+ (e.g. $X = N$)

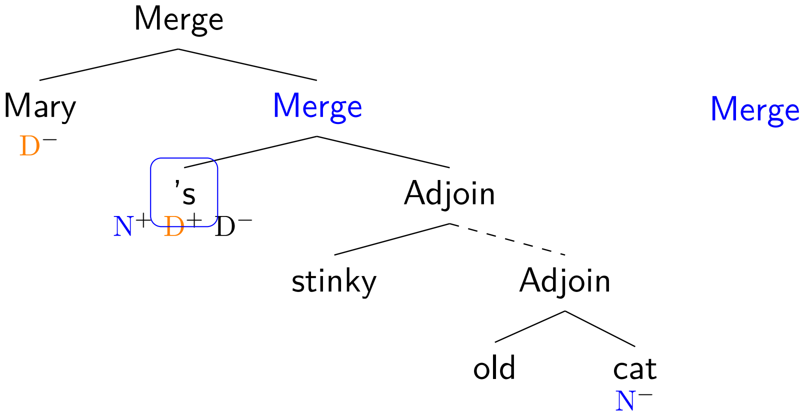
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^+ (e.g. $X = N$)
2. Project any node which has X^- (e.g. $X = N$)

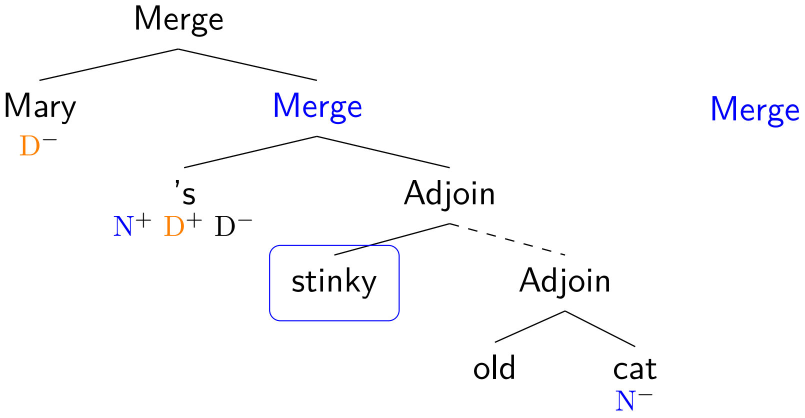
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^{+} (e.g. $\mathrm{X} = \mathrm{N}$)
2. Project any node which has X^{-} (e.g. $\mathrm{X} = \mathrm{N}$)

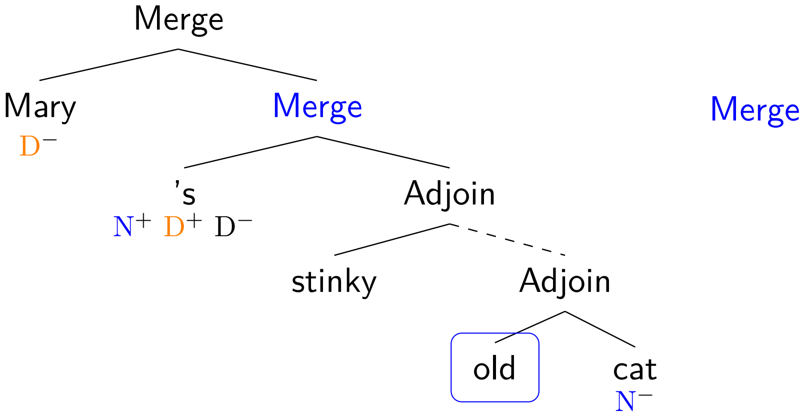
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^+ (e.g. $X = N$)
2. Project any node which has X^- (e.g. $X = N$)

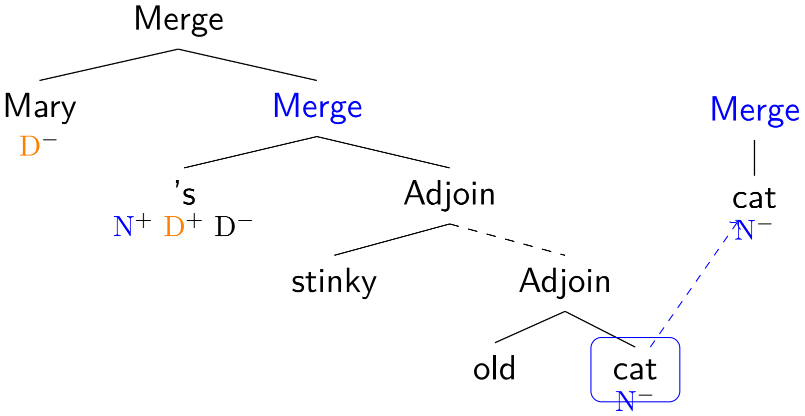
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^{+} (e.g. $\mathrm{X} = \mathrm{N}$)
2. Project any node which has X^{-} (e.g. $\mathrm{X} = \mathrm{N}$)

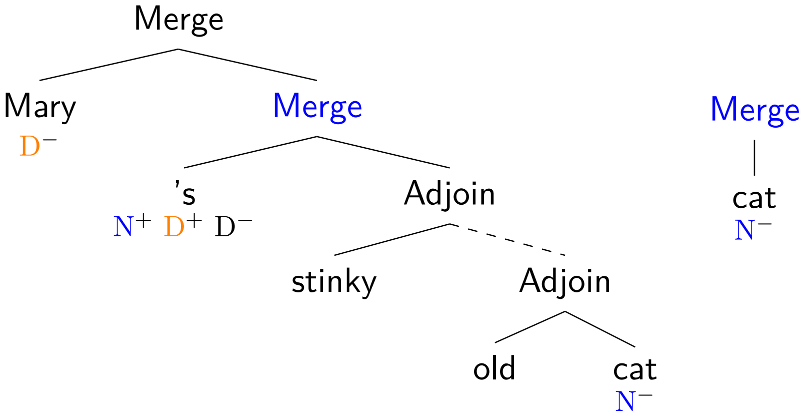
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^+ (e.g. $\mathrm{X} = \mathrm{N}$)
2. Project any node which has X^- (e.g. $\mathrm{X} = \mathrm{N}$)

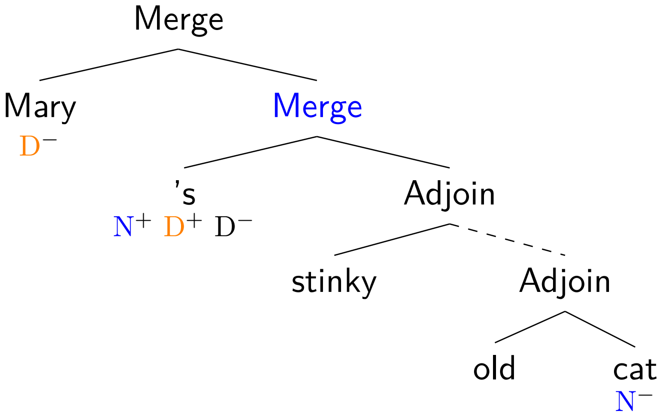
Merge with Adjunction is TSL

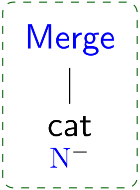


A TSL grammar for Merge

1. Project **Merge** iff a child has X^+ (e.g. $X = N$)
2. Project any node which has X^- (e.g. $X = N$)

Merge with Adjunction is TSL

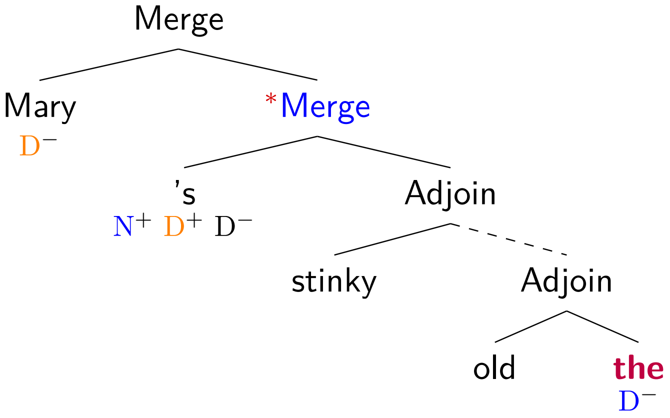




A TSL grammar for Merge

1. Project **Merge** iff a child has X^+ (e.g. $X = N$)
2. Project any node which has X^- (e.g. $X = N$)
3. No Merge without exactly one LI among its daughters.

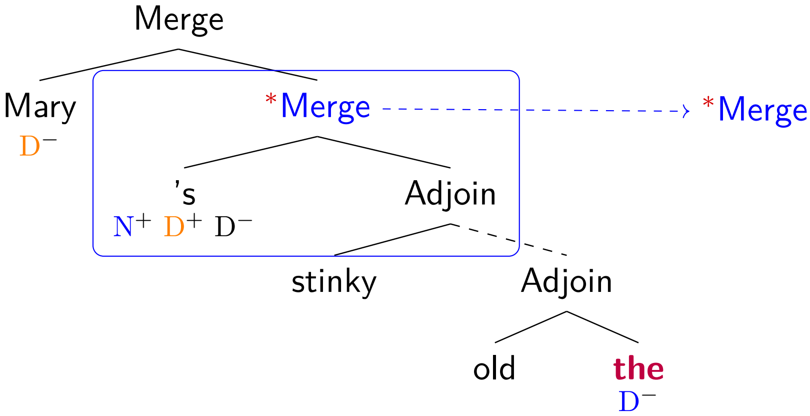
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^+ (e.g. $\mathrm{X} = \mathrm{V}$)
2. Project any node which has X^- (e.g. $\mathrm{X} = \mathrm{V}$)
3. No Merge without exactly one LI among its daughters.

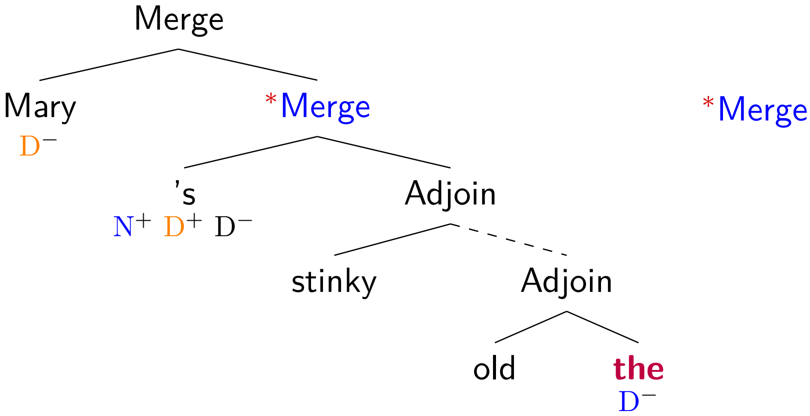
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^+ (e.g. $X = V$)
2. Project any node which has X^- (e.g. $X = V$)
3. No Merge without exactly one LI among its daughters.

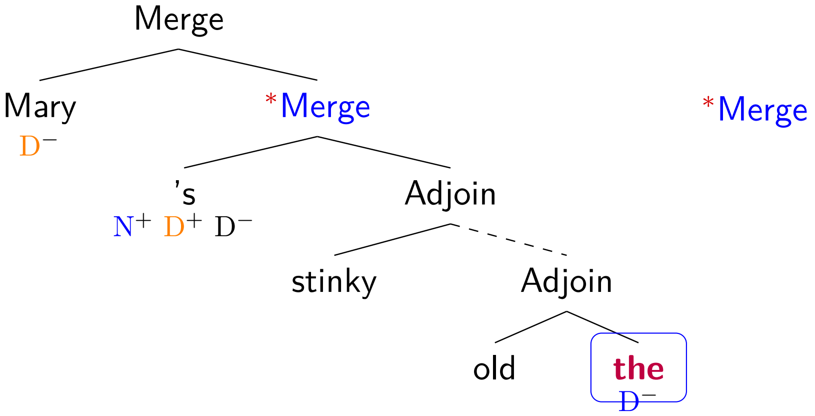
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^- (e.g. $\mathrm{X} = \mathrm{V}$)
2. Project any node which has X^+ (e.g. $\mathrm{X} = \mathrm{V}$)
3. No Merge without exactly one LI among its daughters.

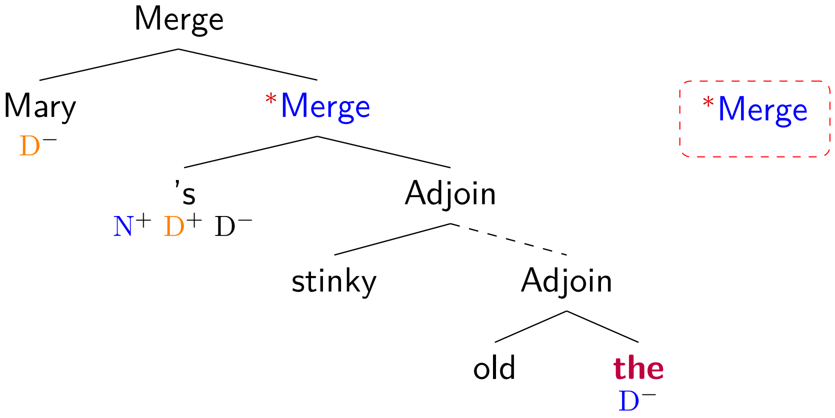
Merge with Adjunction is TSL



A TSL grammar for Merge

1. Project **Merge** iff a child has X^- (e.g. $\mathrm{X} = \mathrm{V}$)
2. Project any node which has X^+ (e.g. $\mathrm{X} = \mathrm{V}$)
3. No Merge without exactly one LI among its daughters.

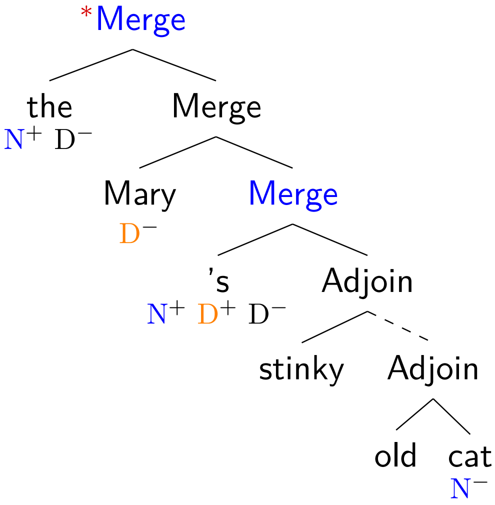
Merge with Adjunction is TSL



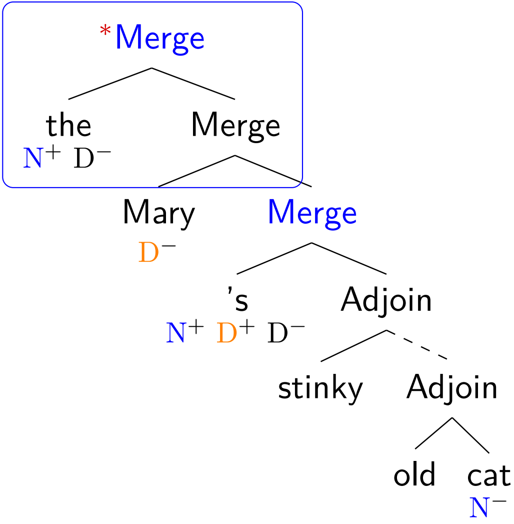
A TSL grammar for Merge

1. Project **Merge** iff a child has X^- (e.g. $X = V$)
2. Project any node which has X^+ (e.g. $X = V$)
3. No Merge without exactly one LI among its daughters.

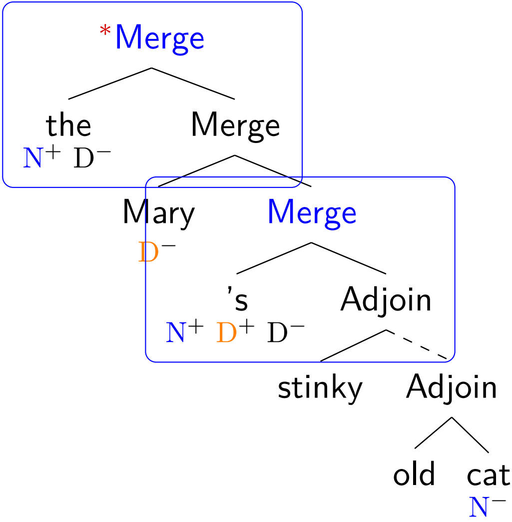
TSL Merge: Understanding the Constraint



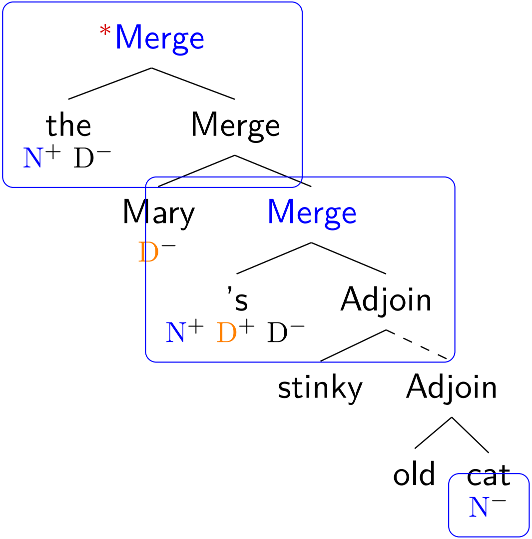
TSL Merge: Understanding the Constraint



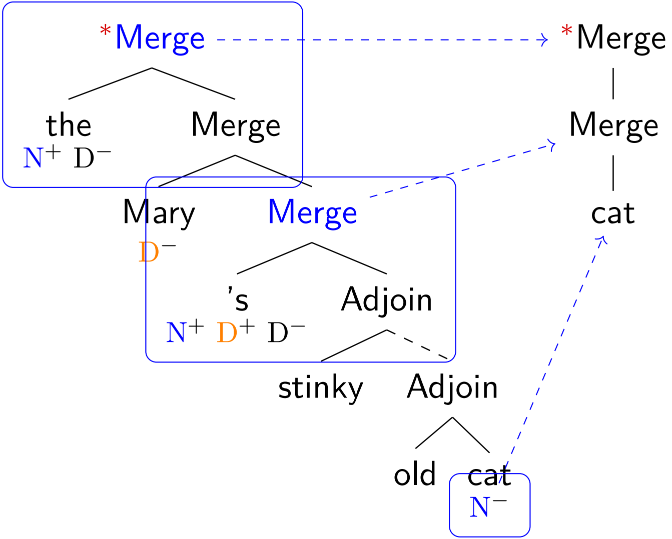
TSL Merge: Understanding the Constraint



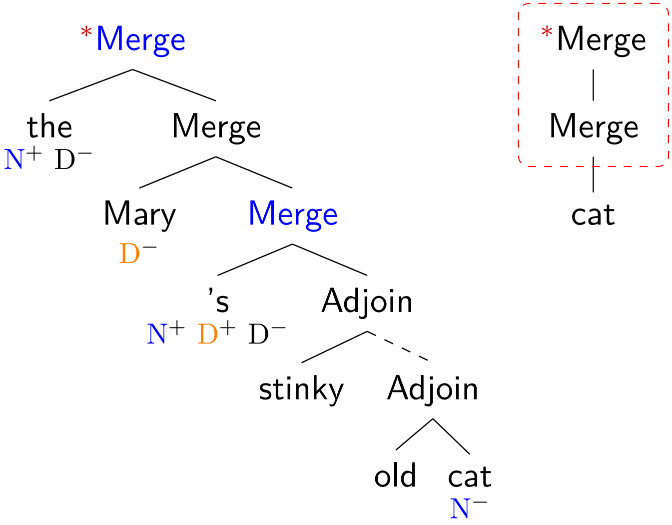
TSL Merge: Understanding the Constraint



TSL Merge: Understanding the Constraint



TSL Merge: Understanding the Constraint



Parallels Between Phonology And Syntax

	Local	Non-local
Phonology	?	?
Syntax	?	?

- **Relativized Locality:**
  Non-local dependencies are local over a simple relativization domain.

Strong Cognitive Parallelism Hypothesis

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

Parallels Between Phonology And Syntax

	Local	Non-local
Phonology	**SL**	?
Syntax	**SL**	?

- **Relativized Locality:**
  Non-local dependencies are local over a simple relativization domain.

Strong Cognitive Parallelism Hypothesis

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

Parallels Between Phonology And Syntax

	Local	Non-local
Phonology	**SL**	**TSL**
Syntax	**SL**	**TSL**

- **Relativized Locality:**
  Non-local dependencies are local over a simple relativization domain.

Strong Cognitive Parallelism Hypothesis

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

Parallels Between Phonology And Syntax

	Local	Non-local	Data Structure
Phonology	**SL**	**TSL**	**Strings**
Syntax	**SL**	**TSL**	**Trees**

- **Relativized Locality:**
  Non-local dependencies are local over a simple relativization domain.

Strong Cognitive Parallelism Hypothesis

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

Parallels Between Phonology And Syntax

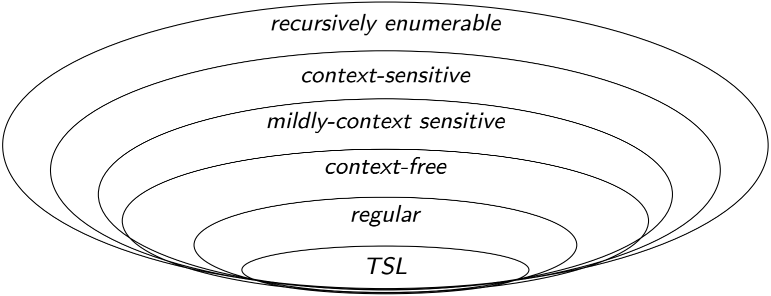
	Local	Non-local	Data Structure
Phonology	**SL**	**TSL**	**Strings**
Syntax	**SL**	**TSL**	**Trees**

- **Relativized Locality:**
  Non-local dependencies are local over a simple relativization domain.

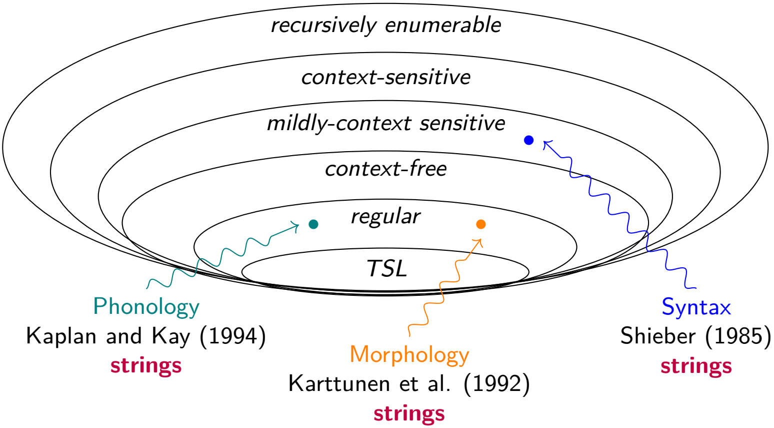
Strong Cognitive Parallelism Hypothesis

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

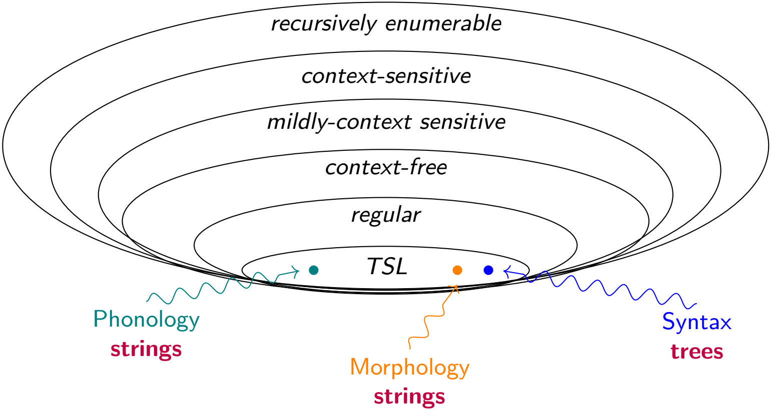
A Bird's-Eye View of the Framework



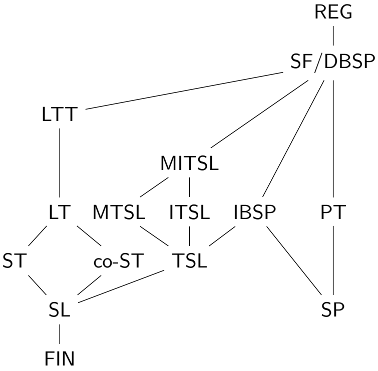
A Bird's-Eye View of the Framework



A Bird's-Eye View of the Framework

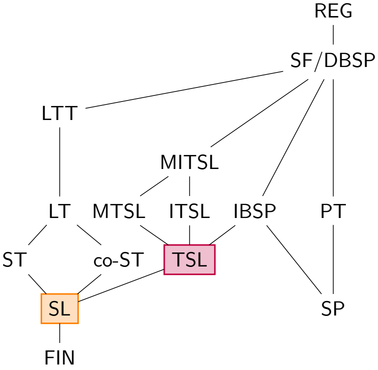


Refining the Hierarchy via Typological Insights



- The goal is not identifying a single "correct" class
- Pinpoint fundamental properties of the patterns:
  SL: $\triangleleft$, TSL: $\triangleleft_T$, etc

Refining the Hierarchy via Typological Insights

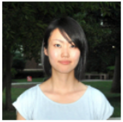


- The goal is not identifying a single "correct" class
- Pinpoint fundamental properties of the patterns:
  SL: $\triangleleft$, TSL: $\triangleleft_T$, etc

Syntax beyond Merge and Move

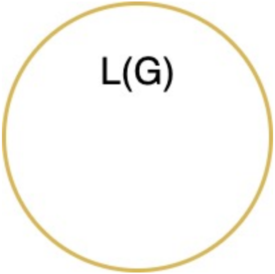
- regular tree languages
  (Michaelis 2004; Kobele et al. 2007)
- subregular operations (Graf 2018)
- subregular dependencies/constraints
  (Vu et al. 2019; Shafiei and Graf 2019)
- tree automata and parsing restrictions
  (Graf & De Santo 2020)



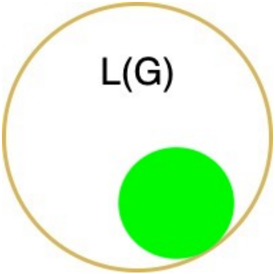
Artificial Grammar Learning (AGL)

- Can be used to test implicit learning abilities (Reber, 1976)



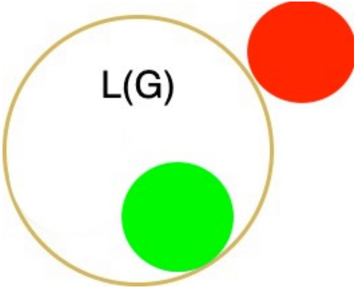
Artificial Grammar Learning (AGL)

- Can be used to test implicit learning abilities (Reber, 1976)



Artificial Grammar Learning (AGL)

- Can be used to test implicit learning abilities (Reber, 1976)



Reber (1976)

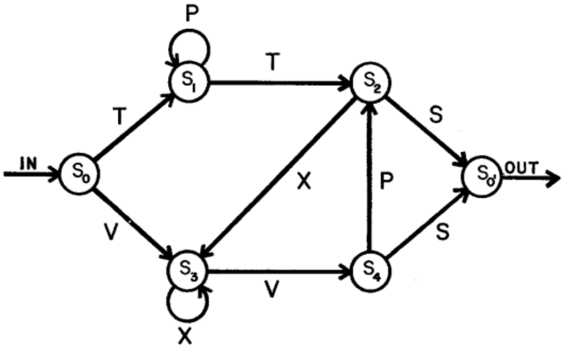


FIG. 1. Schematic state diagram of the grammar used to generate the grammatical stimulus items.

- Stimuli generated from an FST or randomly
  - 28 sentences per group, in sets of four sentences each
  - Participants asked to reproduce the sentences in a group
  - Participants informed of correct/incorrect reproductions, but not of error type

Reber (1976) [cont.]

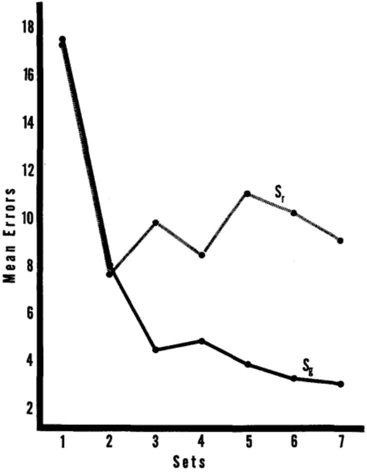
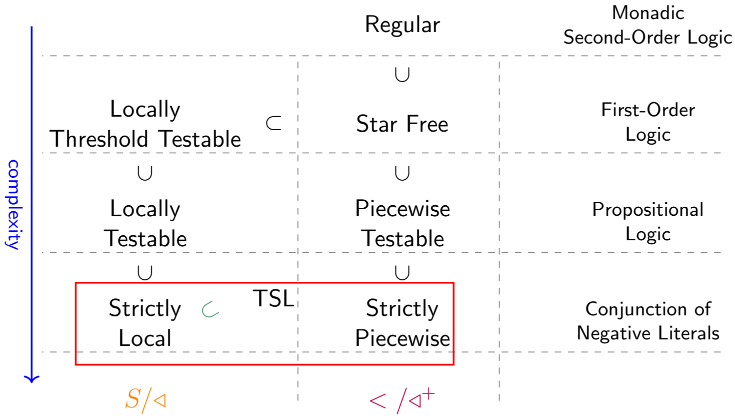


Fig. 2. Mean number of errors to criterion on each of the seven learning sets.

- Stimuli generated from an FST or randomly
  - Significant differences between learning trajectories across participant group

Testing Subregular Predictions



Example: Attested vs. Unattested Patterns

Attested: Unbounded Sibilant Harmony

- Every sibilant needs to harmonize

s ʃ ʃ ʃ

* $hasxintilawʃ$ ok $haʃxintilawʃ$

Unattested: First-Last Harmony

- Harmony only holds between initial and final segments

s ʃ s ʃ

ok $hasxintilawʃ$ * $satxintilawʃ$

Lai (2015)

Learnable vs. Unlearnable Harmony Patterns

Regine Lai

Posted Online July 09, 2015
https://doi.org/10.1162/LING_a_00188

Linguistic Inquiry
Volume 46 | Issue 3 | Summer 2015
p.425-451

Keywords: phonotactics, learnability, computational phonology, formal theory, typology, dependencies

Lai (2015): Stimuli

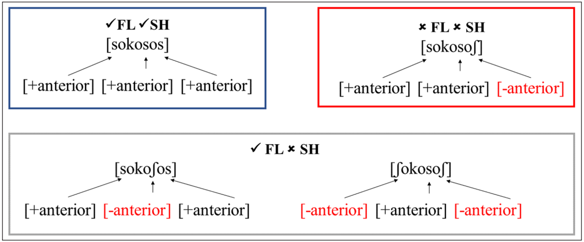


Figure 3: Comparison of SH and FL stimuli.

Lai (2015): Stimuli

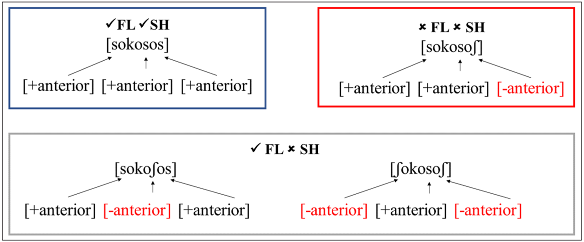
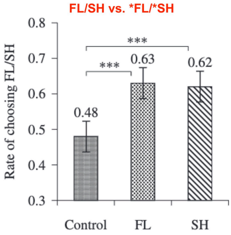


Figure 3: Comparison of SH and FL stimuli.

Table 6
Predicted results with respect to the control group for each test pairing if Sibilant Harmony and First-Last Assimilation grammars were internalized

	Pairs		
Conditions	FL/*SH vs. *FL/*SH (e.g., [s . . . ʃ . . . s] vs. [s . . . s . . . ʃ]) Rate of FL/*SH	FL/SH vs. *FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . s . . . ʃ]) Rate of FL/SH	FL/SH vs. FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . ʃ . . . s]) Rate of FL/SH
SH	~ Control	> Control	> Control
FL	> Control	> Control	~ Control

Lai (2015): Results



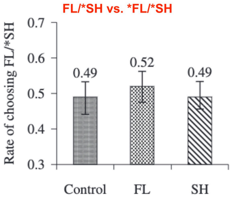


Table 6
Predicted results with respect to the control group for each test pairing if Sibilant Harmony and First-Last Assimilation grammars were internalized

	Pairs		
Conditions	FL/*SH vs. *FL/*SH (e.g., [s . . . ʃ . . . s] vs. [s . . . s . . . ʃ]) Rate of FL/*SH	FL/SH vs. *FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . s . . . ʃ]) Rate of FL/SH	FL/SH vs. FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . ʃ . . . s]) Rate of FL/SH
SH	∼ Control	> Control	> Control
FL	> Control	> Control	∼ Control

- See Avcu and Hestvik (2020), Avcu et al. (2019) for replications

Lai (2015): Results

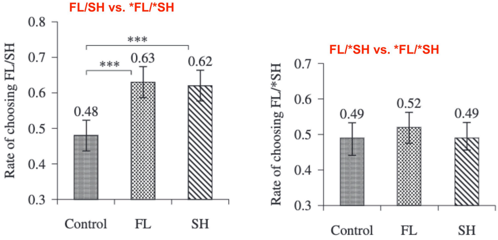


Table 6
Predicted results with respect to the control group for each test pairing if Sibilant Harmony and First-Last Assimilation grammars were internalized

	Pairs		
Conditions	FL/*SH vs. *FL/*SH (e.g., [s . . . ʃ . . . s] vs. [s . . . s . . . ʃ]) Rate of FL/*SH	FL/SH vs. *FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . s . . . ʃ]) Rate of FL/SH	FL/SH vs. FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . ʃ . . . s]) Rate of FL/SH
SH	~ Control	> Control	> Control
FL	> Control	> Control	~ Control

- See Avcu and Hestvik (2020), Avcu et al. (2019) for replications

Lai (2015): Full Results

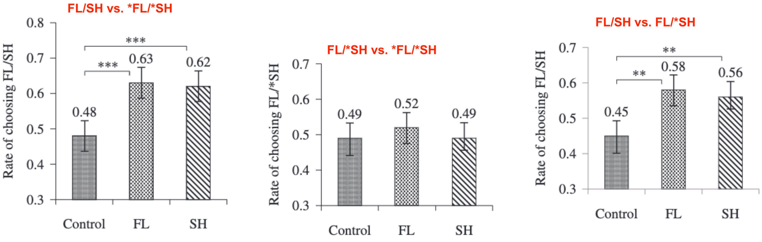


Table 6
Predicted results with respect to the control group for each test pairing if Sibilant Harmony and First-Last Assimilation grammars were internalized

	Pairs		
Conditions	FL/*SH vs. *FL/*SH (e.g., [s . . . ʃ . . . s] vs. [s . . . s . . . ʃ]) Rate of FL/*SH	FL/SH vs. *FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . s . . . ʃ]) Rate of FL/SH	FL/SH vs. FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . ʃ . . . s]) Rate of FL/SH
SH	~ Control	> Control	> Control
FL	> Control	> Control	~ Control

AGL and Syntax/Semantics

> *distinctions between mechanisms for recognizing non-Finite-State stringsets depend on the way in which the additional structure, beyond the string itself, is organized; these are issues that show up in the analysis of the string, not in its form as a sequence of events.*
>
> *Rogers & Pullum 2011*

In other words:

- Questions of complexity confounded by representations
- Questions of representations confounded by procedures

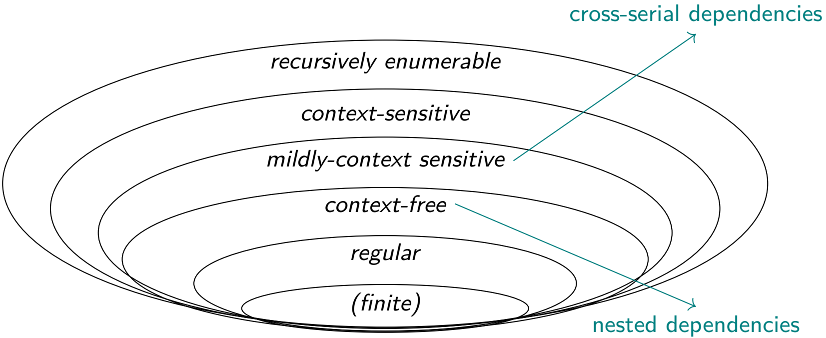
AGL and Syntax/Semantics

> *distinctions between mechanisms for recognizing non-Finite-State stringsets depend on the way in which the additional structure, beyond the string itself, is organized; these are issues that show up in the analysis of the string, not in its form as a sequence of events.*
>
> *Rogers & Pullum 2011*

In other words:

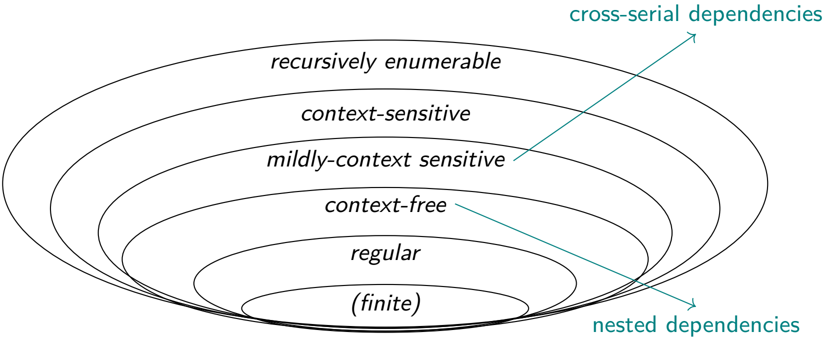
- Questions of complexity confounded by representations
- Questions of representations confounded by procedures

Syntactic Expressivity



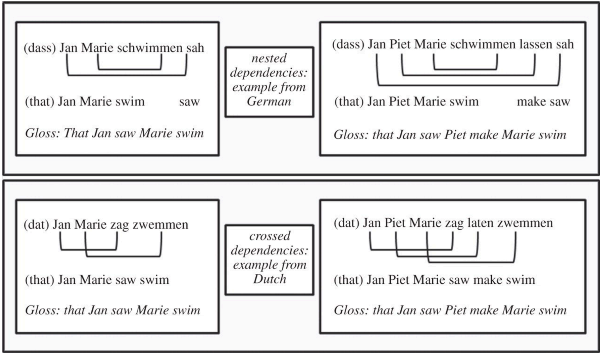
- cross-serial preferred over nested (Bach et al. 1986)
- against predictions from the CH? (Chesi & Moro 2014; de Vries et al. 2012)

Syntactic Expressivity



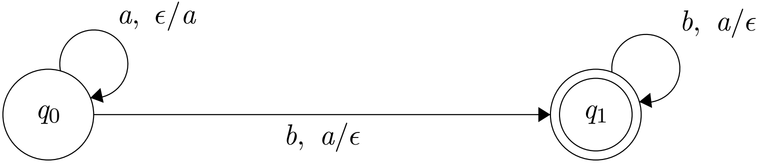
- cross-serial preferred over nested (Bach et al. 1986)
- against predictions from the CH?
  (Chesi & Moro 2014; de Vries et al. 2012)

Expressivity vs. Procedures



- cross-serial preferred over nested (Bach et al. 1986)
- against predictions from the CH? (Chesi & Moro 2014; de Vries et al. 2012)
- BUT: this can easily be derived via processing mechanisms (Savitch 1989; Joshi, 1990; Rainbow and Joshi,1994)
- recognition complexity requires a precise theory of parsing cost

AGL and Syntax/Semantics [cont.]

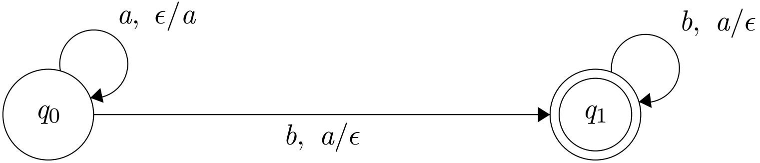


- A^nB^n does not necessarily imply a proper stack
  a PDA with a single counter is enough (Counter Machines)
- Same for the language of strings of **well-nested parentheses**
- Phrase-structure analyses often depend on distinctions based on the meaning of the strings

Complicated questions:

- What **representations** are relevant?
- How are they connected to **tasks**?
- How do we probe them?

AGL and Syntax/Semantics [cont.]

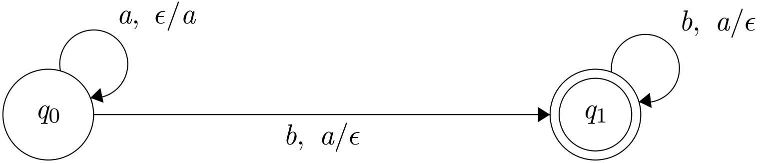


- A^nB^n does not necessarily imply a proper stack
  a PDA with a single counter is enough (Counter Machines)
- Same for the language of strings of **well-nested parentheses**
- Phrase-structure analyses often depend on distinctions based on the meaning of the strings

Complicated questions:

- What **representations** are relevant?
- How are they connected to **tasks**?
- How do we probe them?

AGL and Syntax/Semantics [cont.]

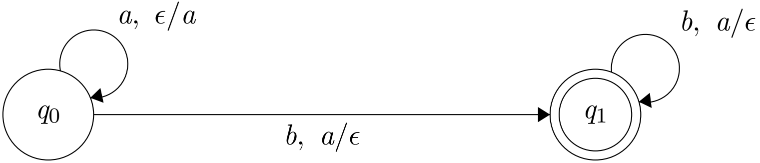


- $A^n B^n$ does not necessarily imply a proper stack
  a PDA with a single counter is enough (Counter Machines)
- Same for the language of strings of **well-nested parentheses**
- Phrase-structure analyses often depend on distinctions based on the meaning of the strings

Complicated questions:

- What **representations** are relevant?
- How are they connected to **tasks**?
- How do we probe them?

AGL and Syntax/Semantics [cont.]



- $A^n B^n$ does not necessarily imply a proper stack
  a PDA with a single counter is enough (Counter Machines)
- Same for the language of strings of **well-nested parentheses**
- Phrase-structure analyses often depend on distinctions based on the meaning of the strings

Complicated questions:

- What **representations** are relevant?
- How are they connected to **tasks**?
- How do we probe them?

A Plethora of Testable Predictions

Observation

- Attested patterns **A** and **B** are TSL.
- But combined pattern **A**+**B** is not TSL.

Prediction

- **A**+**B** should be harder to learn than **A** and **B**

A Plethora of Testable Predictions

Observation

- Attested patterns **A** and **B** are TSL.
- But combined pattern **A**+**B** is not TSL.

Prediction

- **A**+**B** should be harder to learn than **A** and **B**

Morphotactics as Tier-Based Strictly Local Dependencies

Alëna Aksënova Thomas Graf Sedigheh Moradi

Example: Compounding Markers

- Russian has an infix -*o*- that may occur between parts of compounds.
- Turkish has a single suffix -*sı* that occurs at end of compounds.

(8)

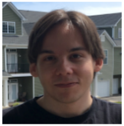
vod	**-o-**	voz	**-o-**	voz
water	-COMP-	carry	-COMP-	carry

'carrier of water-carriers'

(9)

türk	bahçe	kapı	**-sı**	(*-**sı**)
turkish	garden	gate	-COMP	(*-COMP)

'Turkish garden gate'



Example: Compounding Markers [cont.]

- Russian and Turkish are TSL.

Tier$_1$	COMP affix and stem edges #
Russian n-grams	**oo**, **$o**, **o$**
Turkish n-grams	**sisi**, **$si**, **si#**

- The combined pattern would yield **Ruskish**: stem^{n+1}-si^{n}
- This pattern is not regular and hence **not TSL either**.
- **Hypothesis (Aksenova et al, 2016)**
  If a language allows unboundedly many compound affixes, they are **infixes**.

Testable Predictions

- Can naive subjects learn Russian-like, Turkis-like, and Ruskish-like compounding?

Complexity as a Magnifying Lens

- We can compare patterns and predictions across classes
- We can also compare patterns within a same class

Proceedings of the Society for Computation in Linguistics

Volume 1 Article 8

2018

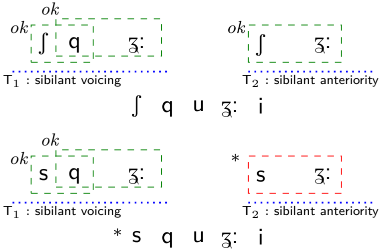
Formal Restrictions On Multiple Tiers

Alena Aksenova
Stony Brook University, alena.aksenova@stonybrook.edu

Sanket Deshmukh
Stony Brook University, sanket.deshmukh@stonybrook.edu

Testing Harmony Systems

- We can also account for multiple processes
- Thus we can cover the complete phonotactics of a language



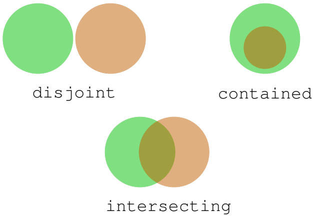


Figure 2: Theoretically possible tier alphabet relations

Testing Harmony Systems (cont.)

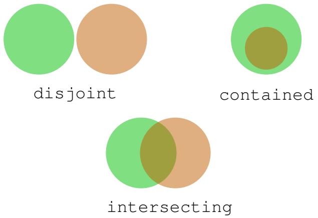
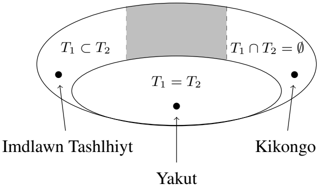


Figure 2: Theoretically possible tier alphabet relations



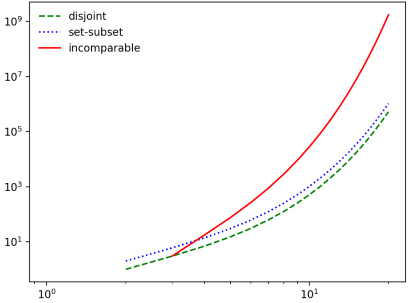
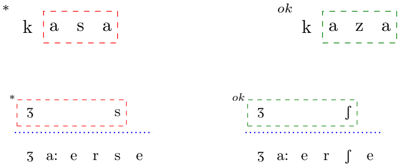


Figure 7: Growth of number of partitions of sets containing up to 20 elements (loglog scale)

The Fallacy of Generalization

- Imagine we want to test the ability to learn long-distance dependencies:



- Assuming an alphabet $\Sigma = \{a, b, c, d, e\}$, the training samples could look like the following:

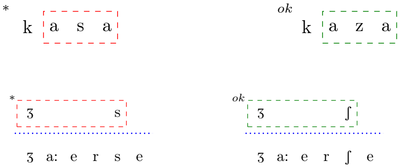
$$L_{loc} = \{abcd, aabcd, baacd, bcaae, \dots\}$$
$$L_{dist} = \{abacd, bacad, bcada, bcaea, \dots\}$$

What happens if we test on stimuli with similar distances?

$$L_{test} = \{abcad, abcad, bacda, abcea, \dots\}$$

The Fallacy of Generalization

- Imagine we want to test the ability to learn long-distance dependencies:



- Assuming an alphabet $\Sigma = \{a, b, c, d, e\}$, the training samples could look like the following:

$$L_{loc} = \{abcd, aabcd, baacd, bcaae, \dots\}$$
$$L_{dist} = \{abacd, bacad, bcada, bcaea, \dots\}$$

What happens if we test on stimuli with similar distances?

$$L_{test} = \{abcad, abcad, bacda, abcea, \dots\}$$

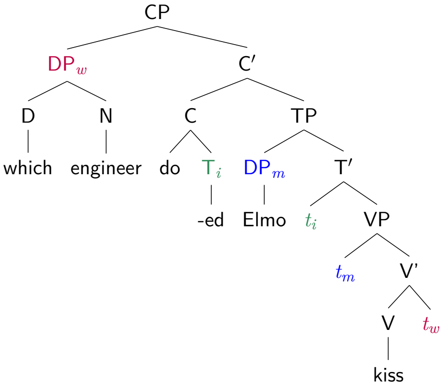
Why MGs?

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
     $\Rightarrow$ cf. TAG (Rambow & Joshi, 1995; Demberg, 2008)

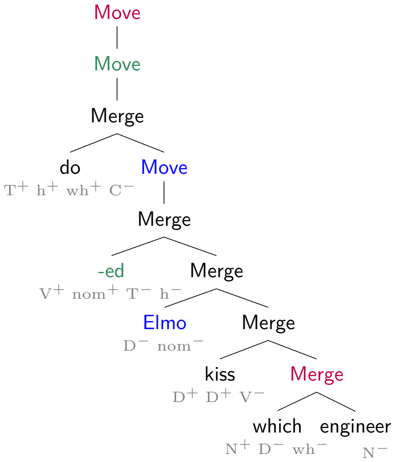
Some Important Properties of MGs

- 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)

- • **Who** • **does** • **Salem** • **T** • **mock**

[1]CP

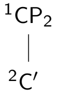
step 1	*CP* is conjectured
step 2	*CP* expands to *C'*
step 3	*C'* expands to *does* and *TP*
step 4	*TP* expands to *Salem* and *T'*
step 5	*T'* expands to *T* and *VP*
step 6	*VP* expands to *mock* and *who*
step 7	***who*** is found
step 8	***does*** is found
step 9	***Salem*** is found
step 10	***T***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

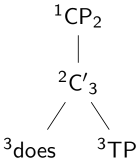
Incremental Top-Down Parsing

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



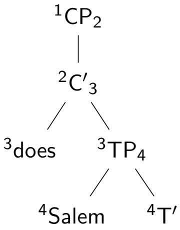
Incremental Top-Down Parsing

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



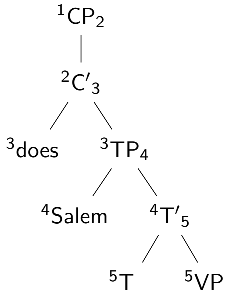
Incremental Top-Down Parsing

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



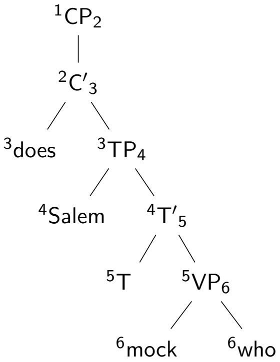
Incremental Top-Down Parsing

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



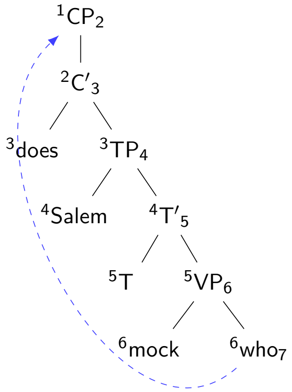
Incremental Top-Down Parsing

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



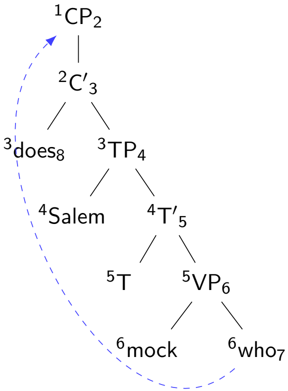
Incremental Top-Down Parsing

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



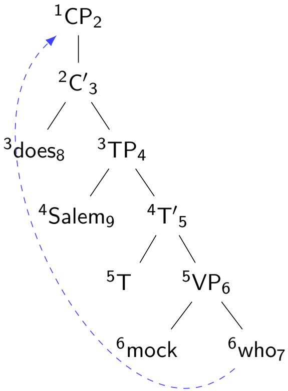
Incremental Top-Down Parsing

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



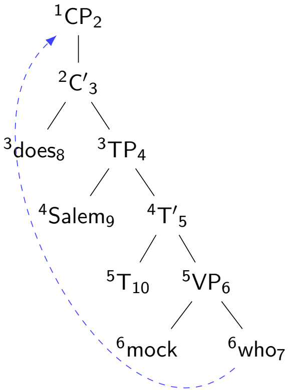
Incremental Top-Down Parsing

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



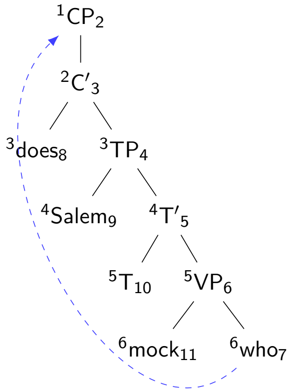
Incremental Top-Down Parsing

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



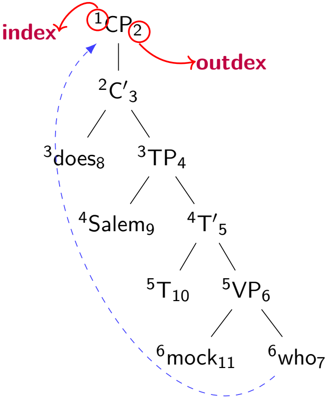
Incremental Top-Down Parsing

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



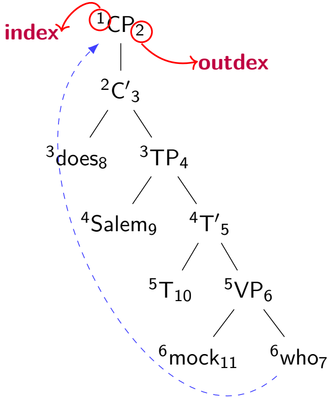
Incremental Top-Down Parsing

Technical details!

- String-driven recursive descent parser (Stabler 2013)

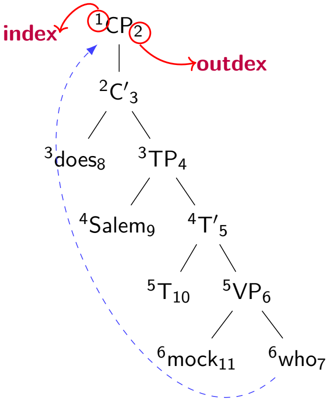
- • **Who** • **does** • **Salem** • **T** • **mock**

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



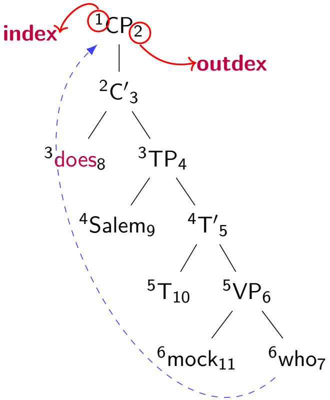
Index and Outdex are our connection to memory!

Computing Metrics: An Example



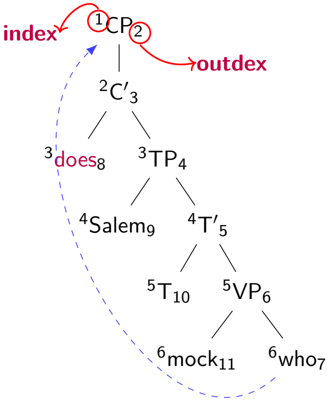
Tenure how long a node is kept in memory

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**$(Salem), \dots\} = 5$

A Case Study: Italian Postverbal Subjects

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)

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

(11) 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):

(12) 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:

(13) 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):

(12) 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:

(13) 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):

(12) 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:

(13) 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**

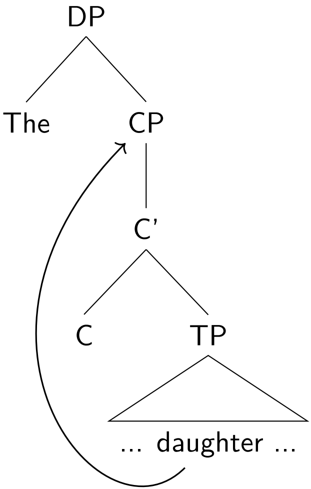
(2) Il cavallo che i leoni hanno inseguito
The horse that the lions have chased
"The horse that the lions chased" **ORC**

(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**

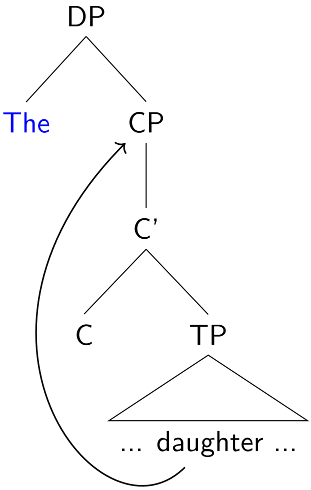
Kayne's Promotion Analysis (Kayne 1994)



- 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]]]

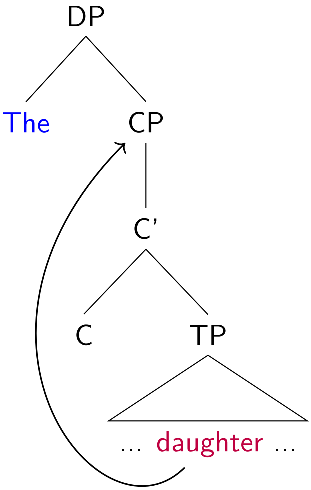
Kayne's Promotion Analysis (Kayne 1994)



- 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]]]

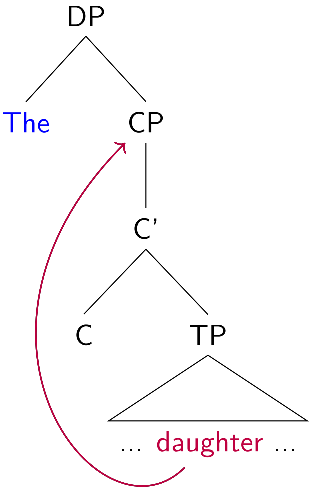
Kayne's Promotion Analysis (Kayne 1994)



- 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)

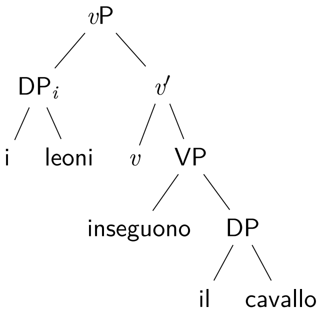


- RC is selected by an external D^0
- the RC head is a nominal constituent
- the RC head raises from its base position to [Spec, CP]

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

Postverbal Subjects (Belletti & Leonini 2004)

(5) 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

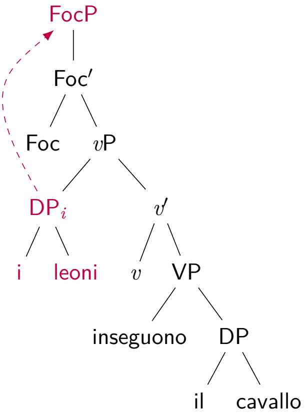
Postverbal Subjects (Belletti & Leonini 2004)

(6) 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



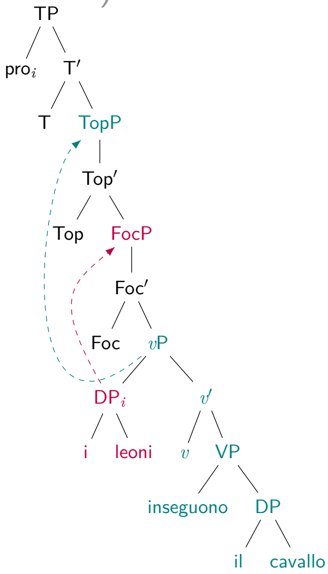
Postverbal Subjects (Belletti & Leonini 2004)

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

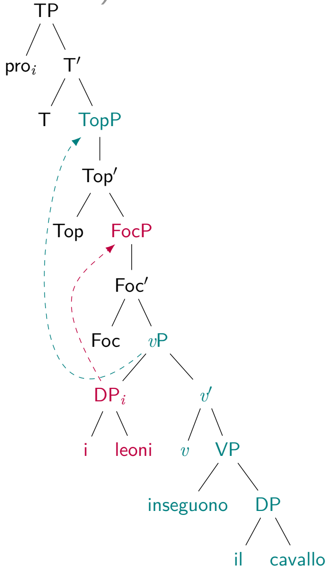
- the subject DP raises to Spec, FocP
- The whole vP raises to Spec, TopP

Technical details!

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



Postverbal Subjects (Belletti & Leonini 2004)



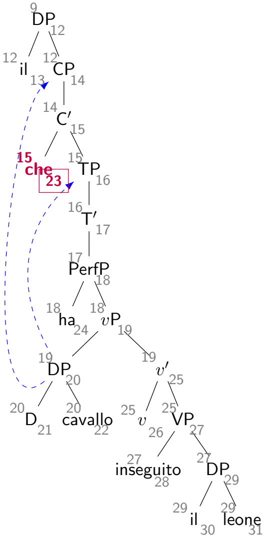
(7) Inseguono il cavallo i leoni
Chase the horse the lions
"The lions chase the horse"

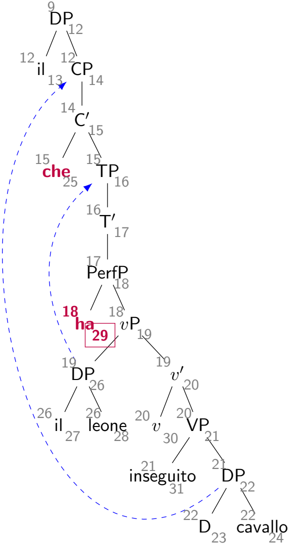
- the subject DP raises to Spec, FocP
- The whole vP raises to Spec,TopP

Technical details!

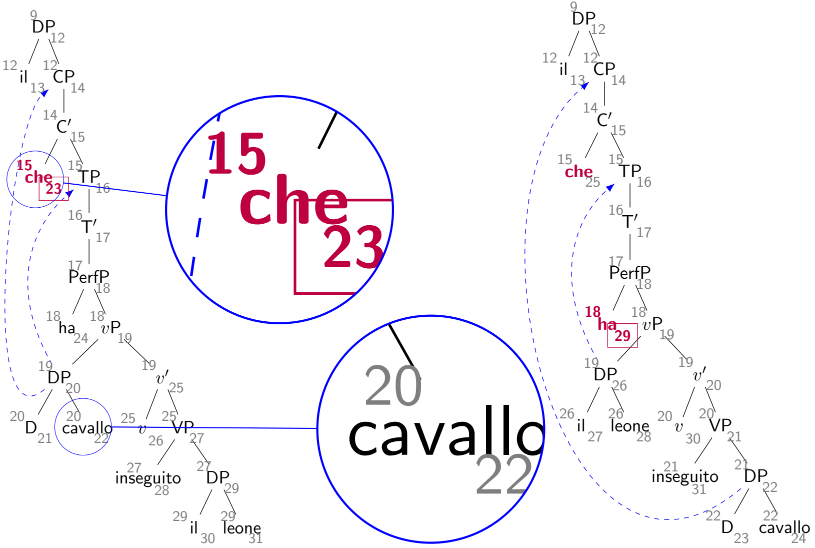
- an expletive *pro* is base generated in Spec,TP

Results: SRC > ORC

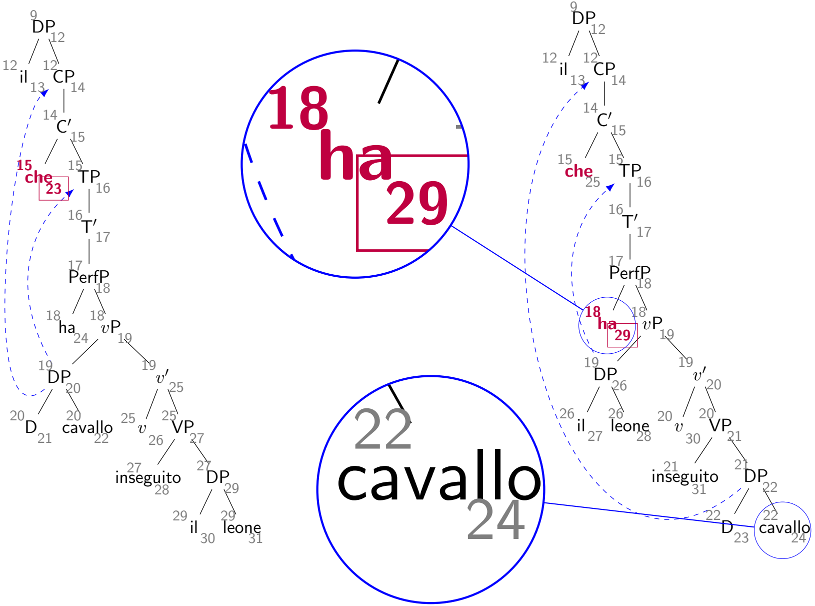




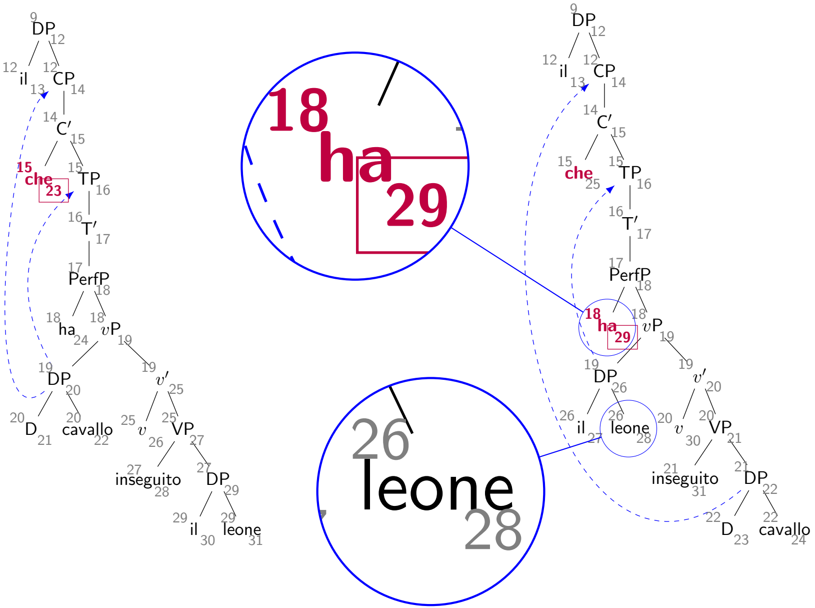
Results: SRC > ORC



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

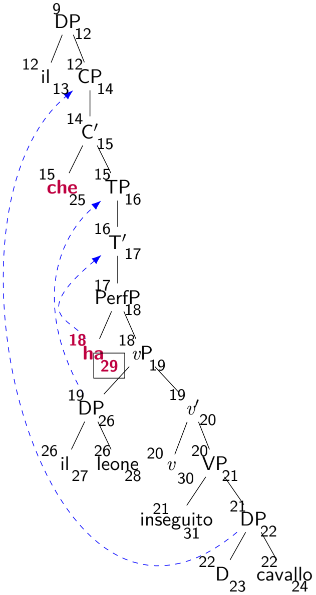
- DLT, active-filler strategy, Competition model, ...

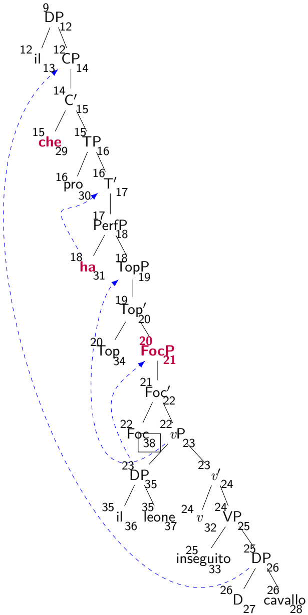
ORC > ORCp

- 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

(8) Ha chiamato Gio
Has called Giovanni

a. "He/she/it called Gio" **SVO**

b. "Gio called" **VS**

- Unaccusatives vs. Unergatives

(9) È arrivato Gio
Is arrived Gio

"Gio arrived" **Unaccusative**

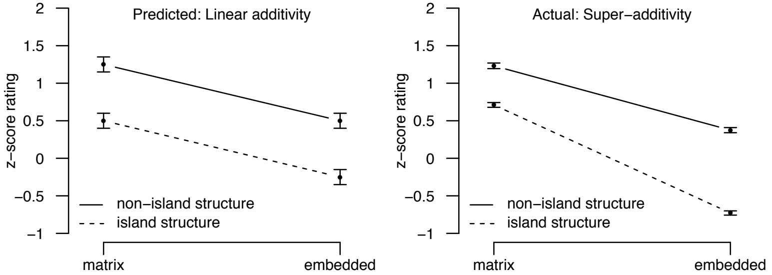
(10) Ha corso Gio
Has ran Gio

"Gio ran" **Unergative**

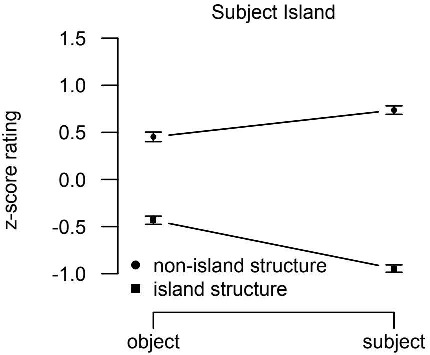
Gradience in Islands

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:

(11) 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:

(12) 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

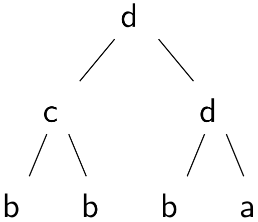
c. Who *t* thinks the speech about global warming interrupted the primetime TV show? Matrix | Island

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

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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

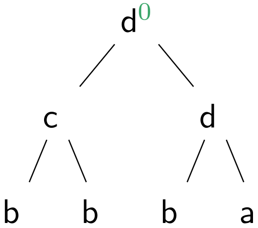


- $0(b) \to b$; $1(b) \to b$
- $1(a) \to a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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

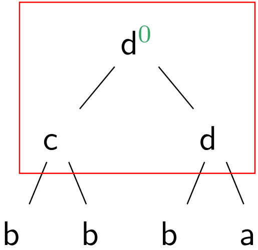


- $0(b) \to b$; $1(b) \to b$
- $1(a) \to a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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

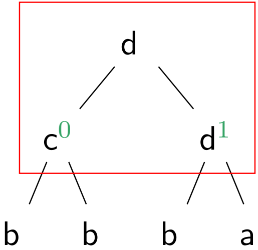


- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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

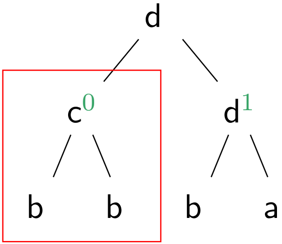


- $0(b) \to b$; $1(b) \to b$
- $1(a) \to a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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

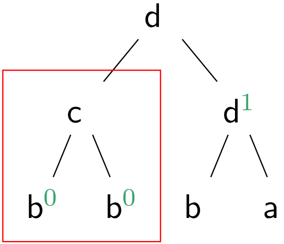


- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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

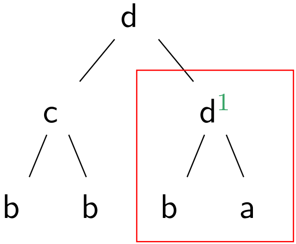


- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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

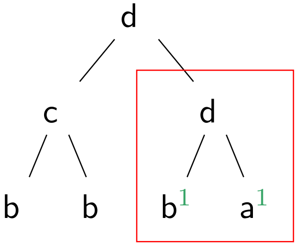


- $0(b) \to b$; $1(b) \to b$
- $1(a) \to a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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

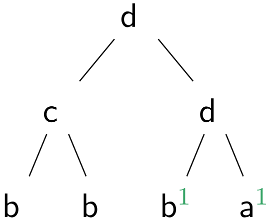


- $0(b) \to b$; $1(b) \to b$
- $1(a) \to a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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

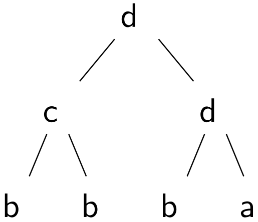


- $0(b) \to b$; $1(b) \to b$
- $1(a) \to a$

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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



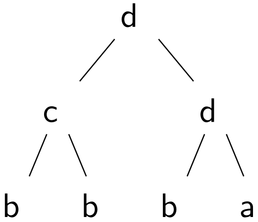
- $0(b) \to b$; $1(b) \to b$
- $1(a) \to a$

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

Top-down Parsing + Grammaticalized Constraints?

Graf & De Santo (2019)

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



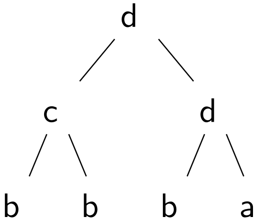
- $0(b) \rightarrow b$; $1(b) \rightarrow b$
- $1(a) \rightarrow a$

- 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.



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

- 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.)

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

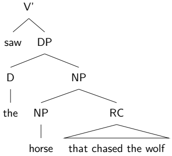
(13) (Io) Ho visto **[la nonna della ragazza che gridava]**
(I) have seen the grandma of the girl that screaming
'I saw **[the grandma of the girl that was screaming]**"

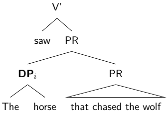
- RC: HA
- RC: LA

Grillo & Costa: Pseudo-RCs in Italian

(13) (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

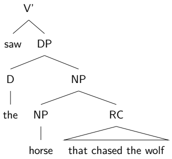


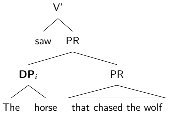


Grillo & Costa: Pseudo-RCs in Italian

(13) (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: ∼ 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

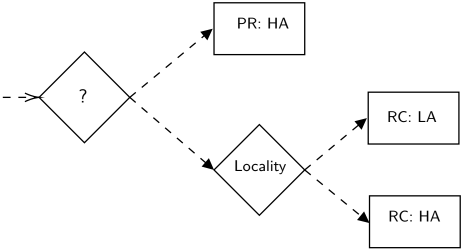
(14) (Io) Ho visto **[la nonna della ragazza che gridava]**
(I) have seen the grandma of the girl that screaming
'I saw **[the grandma of the girl that was screaming]**'

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

(14) (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]'

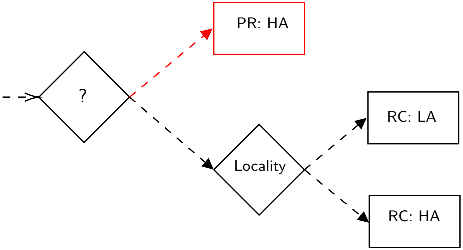


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

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

The PR First Hypothesis

(14) (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]**'

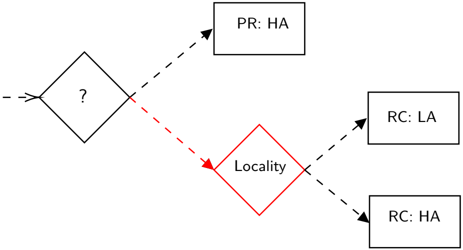


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

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

The PR First Hypothesis

(14) (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]**'

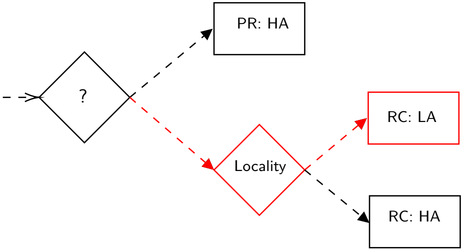


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

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

The PR First Hypothesis

(14) (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]**'

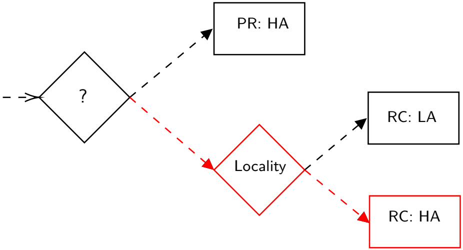


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

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

The PR First Hypothesis

(14) (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]**'

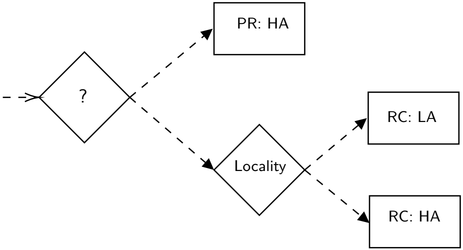


Syntactic tests (Guasti 1999, Cinque 1992, Casalicchio 2013: a.o.)

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

The PR First Hypothesis

(14) (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: (∼) 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%

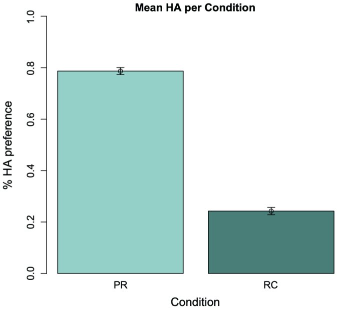


Fig. 2. Summary of attachment preference experiment 2.

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%

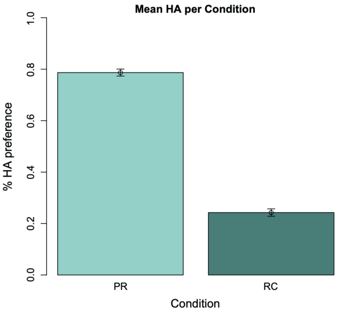


Fig. 2. Summary of attachment preference experiment 2.

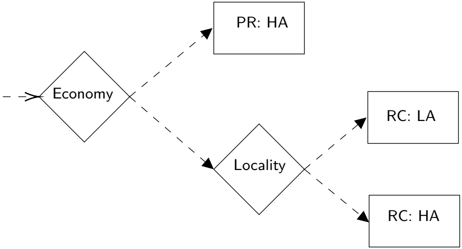
Online too!

- 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)

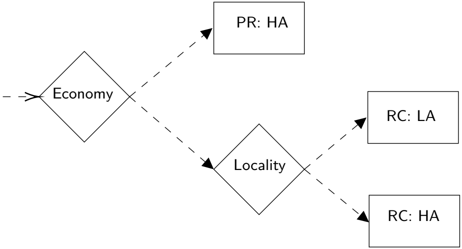
- PR structurally less complex than RC
- RCs: richer and more articulated functional domain

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?

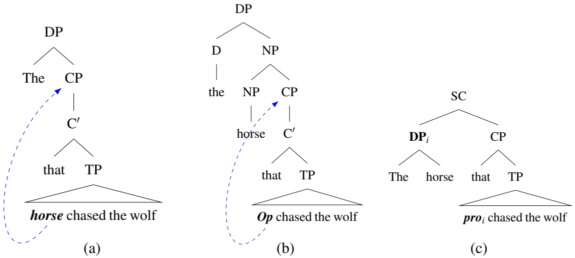


Figure 2: Sketches of the *(a)* RC with Promotion, *(b)* RC with Wh-movement, and *(c)* PR analyses for the sentence *The horse that the wolf chased*.

Modeling Results (De Santo & Lee, 2022b)

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

(15) (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		

(15) (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		

(15) (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		

(15) (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	✓	✓

(15) (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	✓	✓

(15) (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

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

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

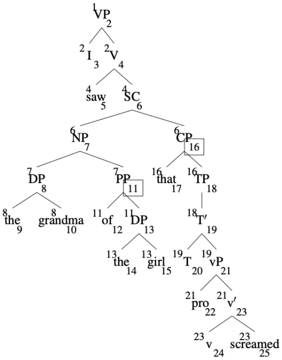


Figure 3: Annotated derivation trees for the Italian sentence *I saw the grandma of the girl that screamed*, according to a pseudo-relaive clause analysis. The tree is treated as a VP since additional structure in the matrix clause would be identical across comparisons.

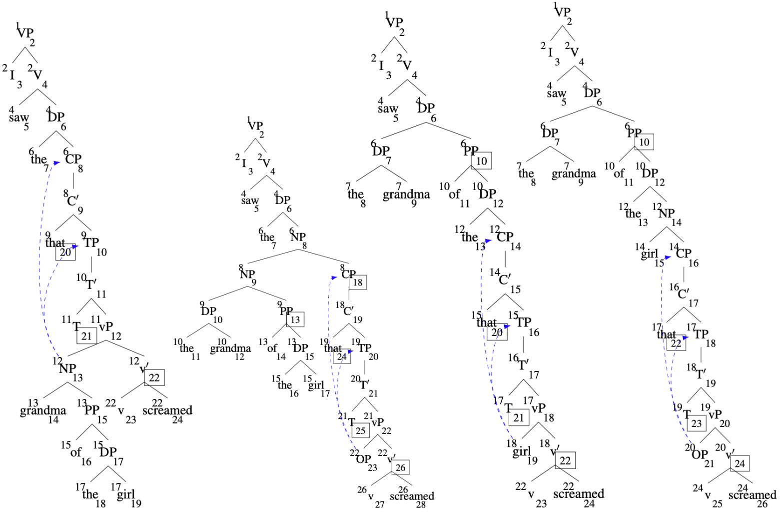
MG Parser		
Hypothesis	**Promotion**	**Wh-mov**
PR < HA	✓	Tie
PR < LA	×	×
LA < HA	✓	✓

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)

Selected References I

Applegate, R.B. 1972. *Ineseno chumash grammar*. Doctoral Dissertation, University of California,Berkeley.

Avcu, Enes, and Arild Hestvik. 2020. Unlearnable phonotactics. *Glossa: a journal of general linguistics* 5.

De Santo, Aniello, and Thomas Graf. 2017. Structure sensitive tier projection: Applications and formal properties. Ms., Stony Brook University.

Frey, Werner, and Hans-martin Gärtner. 2002. On the treatment of scrambling and adjunction in minimalist grammars. In *In Proceedings, Formal Grammar?02*. Citeseer.

Graf, Thomas. 2012. Locality and the complexity of Minimalist derivation tree languages. In *Formal Grammar 2010/2011*, ed. Philippe de Groot and Mark-Jan Nederhof, volume 7395 of *Lecture Notes in Computer Science*, 208–227. Heidelberg: Springer. URL `http://dx.doi.org/10.1007/978-3-642-32024-8_14`.

Graf, Thomas. 2017. Why movement comes for free once you have adjunction. In *Proceedings of CLS 53*. URL `http://ling.auf.net/lingbuzz/003943`, (to appear).

Graf, Thomas. 2018. Why movement comes for free once you have adjunction. In *Proceedings of CLS 53*, ed. Daniel Edmiston, Marina Ermolaeva, Emre Hakgüder, Jackie Lai, Kathryn Montemurro, Brandon Rhodes, Amara Sankhagowit, and Miachel Tabatowski, 117–136.

Selected References II

Heinz, Jeffrey, Chetan Rawal, and Herbert G. Tanner. 2011. Tier-based strictly local constraints in phonology. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics*, 58–64. URL `http://www.aclweb.org/anthology/P11-2011`.

Hyman, Larry M. 2011. Tone: Is it different? *The Handbook of Phonological Theory, Second Edition* 197–239.

Kobele, Gregory M., Christian Retoré, and Sylvain Salvati. 2007. An automata-theoretic approach to Minimalism. In *Model Theoretic Syntax at 10*, ed. James Rogers and Stephan Kepser, 71–80.

Lai, Regine. 2015. Learnable vs. unlearnable harmony patterns. *Linguistic Inquiry* 46:425–451.

Michaelis, Jens. 2004. Observations on strict derivational minimalism. *Electronic Notes in Theoretical Computer Science* 53:192–209.

Shafiei, Nazila, and Thomas Graf. 2019. The subregular complexity of syntactic islands. Ms., Stony Brook University.

Stabler, Edward P. 1997. Derivational Minimalism. In *Logical aspects of computational linguistics*, ed. Christian Retoré, volume 1328 of *Lecture Notes in Computer Science*, 68–95. Berlin: Springer.

Stabler, Edward P. 2011. Computational perspectives on Minimalism. In *Oxford handbook of linguistic Minimalism*, ed. Cedric Boeckx, 617–643. Oxford: Oxford University Press.

Selected References III

Vu, Mai Ha, Nazila Shafiei, and Thomas Graf. 2019. Case assignment in TSL syntax: A case study. In *Proceedings of the Society for Computation in Linguistics (SCiL) 2019*, ed. Gaja Jarosz, Max Nelson, Brendan O'Connor, and Joe Pater, 267–276.