Learnability Insights from (and into) the TSL Neighborhood

Aniello De Santo
he/him

aniellodesanto.github.io
aniello.desanto@utah.edu

LSA 2025

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?

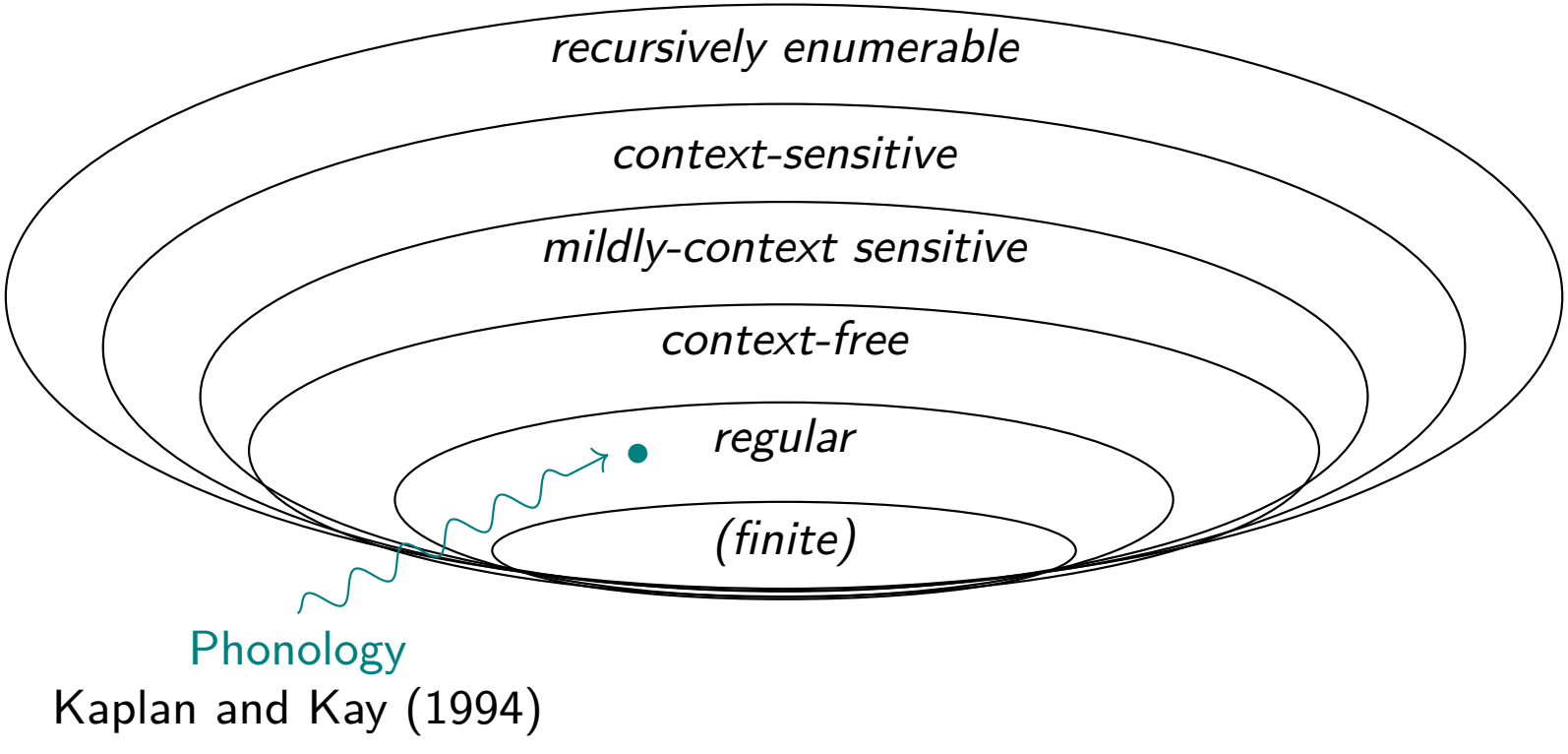
Theories of linguistic representations from typological/empirical observations?

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)

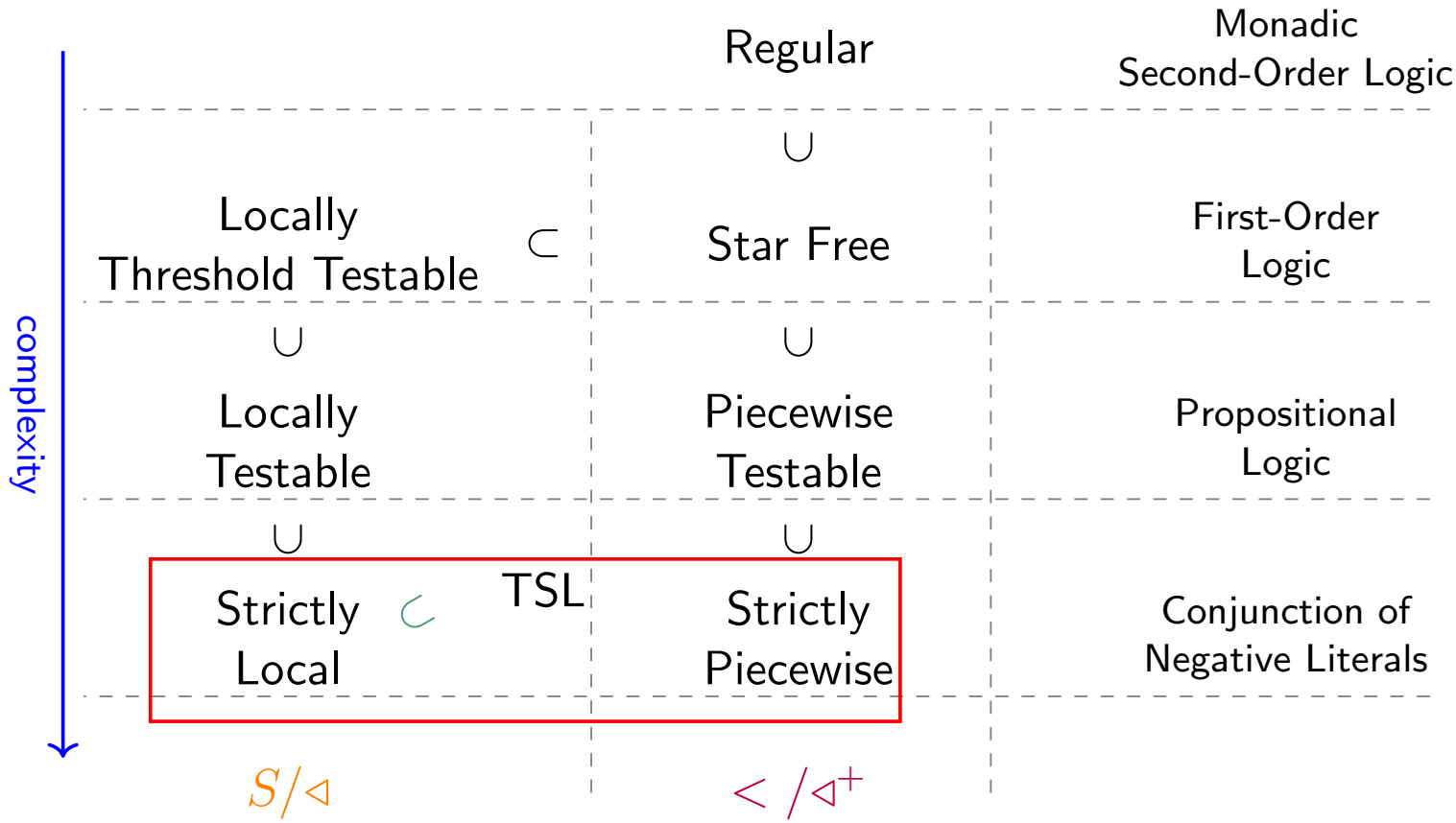
Spoken Languages' Phonotactics as a Regular System

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



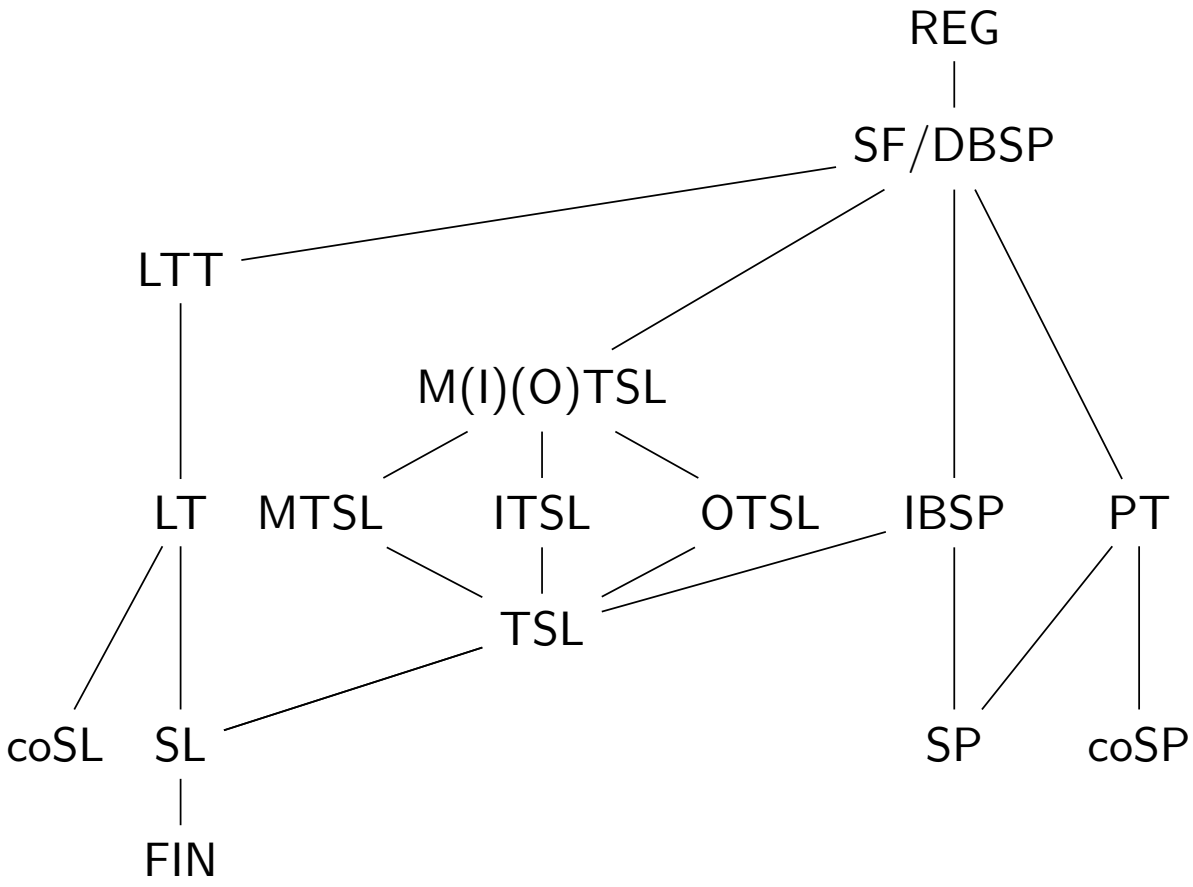
Kaplan and Kay (1994)

From Subclasses...[1]



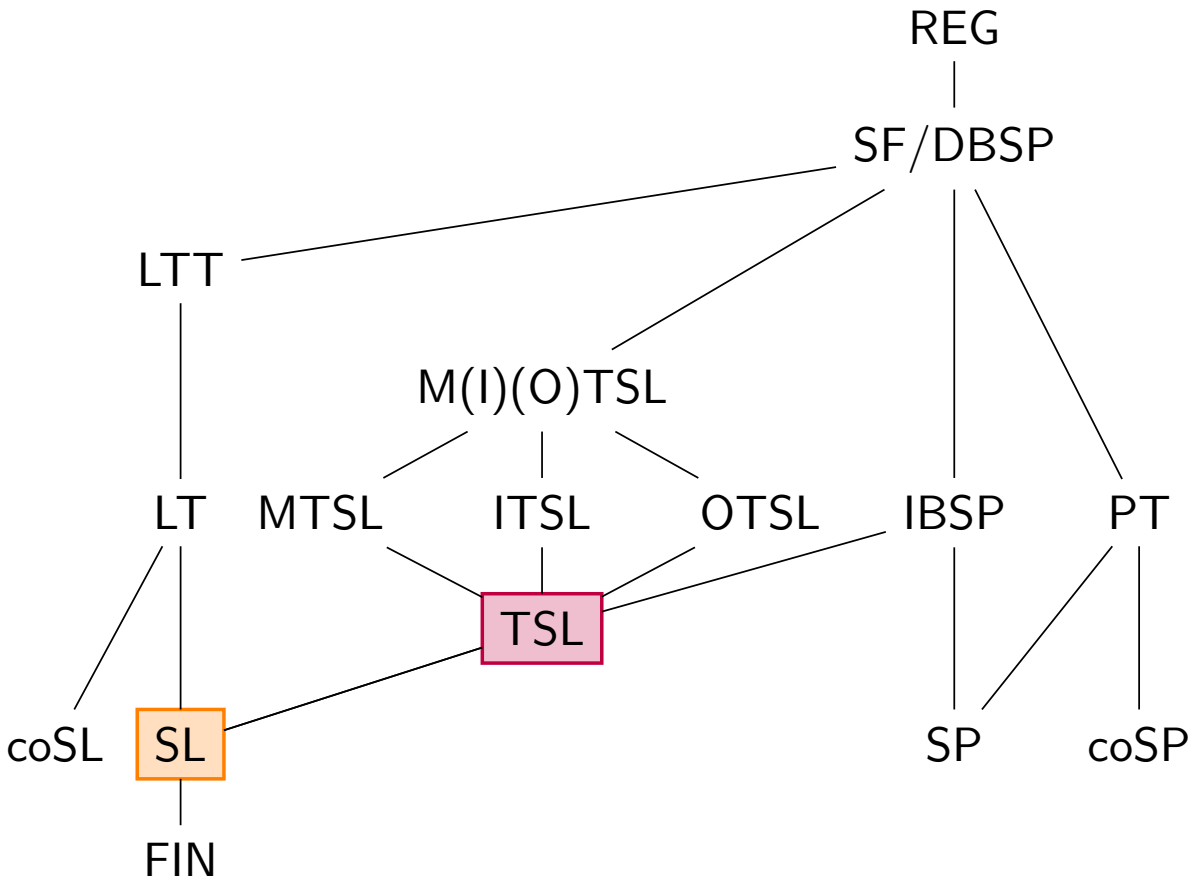
[1] Subregular classes as of Heinz (2011)

... To Subclasses[2]



[2]Subregular classes as of (Lambert 2022)

... To Subclasses[2]



[2]Subregular classes as of (Lambert 2022)

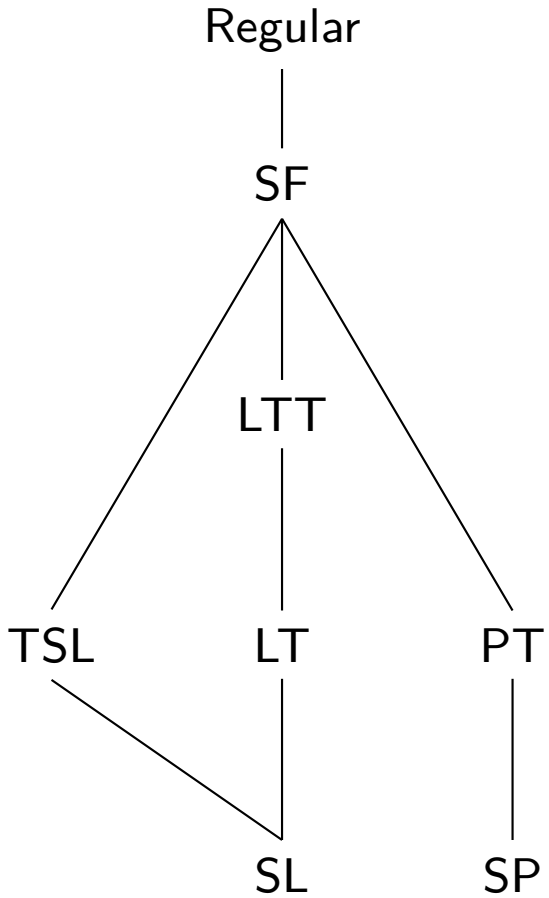
Boundaries vs. Invariants

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

(Rogers & Pullum 2011)

Invariants

- SL: adjacency
- SP: precedence
- TSL: relativized adjacency
- ...



- **So** let's look at subtle differences between classes!

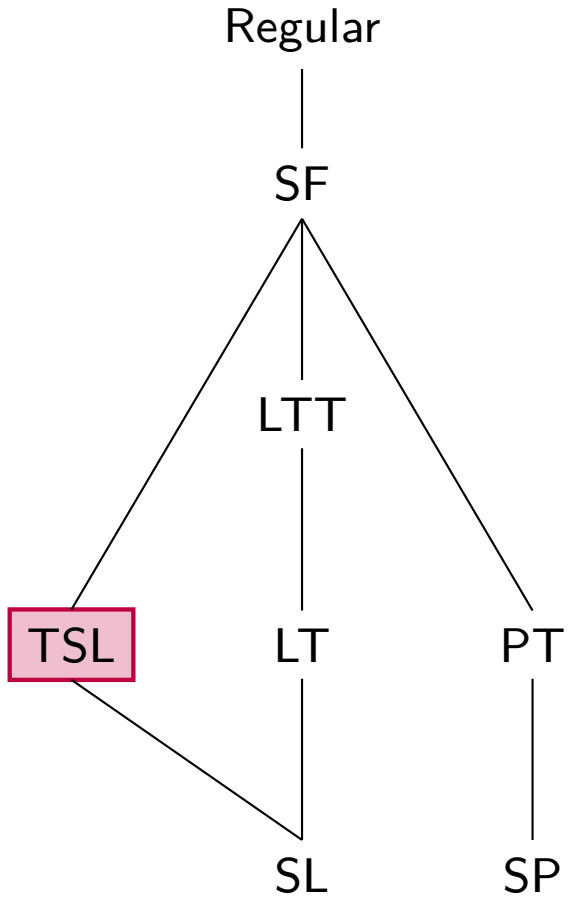
Boundaries vs. Invariants

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

(Rogers & Pullum 2011)

Invariants

- SL: adjacency
- SP: precedence
- TSL: relativized adjacency
- ...



- **So** let's look at subtle differences between classes!

Unbounded Dependencies as TSL

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

  (1) 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 Ineseño Chumash

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 as TSL

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

  (1) a. * ha**s**xintilawa∫
  b. * ha∫xintilawa**s**
  c. ha∫xintilawa∫

- What do we need to project? [+strident]
- What do we need to ban? *[+ant][−ant],*[−ant][+ant]
  I.E. *s∫, *sʒ, *z∫, *zʒ, *∫s, *ʒs, *∫z, *ʒz

Example: TSL Ineseño Chumash

Ineseño Chumash: The Full Picture

Sibilant Harmony in Ineseño Chumash (McMullin 2016)

1) Unbounded sibilant harmony

a.	/k-su-ʃojin/	kʃuʃojin	"I darken it"
b.	/k-su-k'ili-mekeken-ʃ/	kʃuk'ilimekeketʃ	"I straighten up"

2) /s/→ [ʃ] when preceding (adjacent) [t, n, l]

a.	/s-lok'in/	ʃlok'in	"he cuts it"
b.	/s-tepuʔ/	ʃtepuʔ	"he gambles"

3) Long-distance agreement overrides local disagreement

a.	/s-iʃt-iʃti-jep-us/	sististijepus	"they show him"
b.	/s-net-us/	snetus	"he does it to him"

Ineseño Chumash: The Full Picture

Sibilant Harmony in Ineseño Chumash (McMullin 2016)

1) Unbounded sibilant harmony

a.	/k-su-ʃojin/	kʃuʃojin	"I darken it"
b.	/k-su-k'ili-mekeken-ʃ/	kʃuk'ilimekeketʃ	"I straighten up"

2) /s/→ [ʃ] when preceding (adjacent) [t, n, l]

a.	/s-lok'in/	ʃlok'in	"he cuts it"
b.	/s-tepuʔ/	ʃtepuʔ	"he gambles"

3) Long-distance agreement overrides local disagreement

a.	/s-iʃt-iʃti-jep-us/	sististijepus	"they show him"
b.	/s-net-us/	snetus	"he does it to him"

Ineseño Chumash: The Full Picture

Sibilant Harmony in Ineseño Chumash (McMullin 2016)

1) Unbounded sibilant harmony

a.	/k-su-ʃojin/	kʃuʃojin	"I darken it"
b.	/k-su-k'ili-mekeken-ʃ/	kʃuk'ilimekeketʃ	"I straighten up"

2) /s/→ [ʃ] when preceding (adjacent) [t, n, l]

a.	/s-lok'in/	ʃlok'in	"he cuts it"
b.	/s-tepuʔ/	ʃtepuʔ	"he gambles"

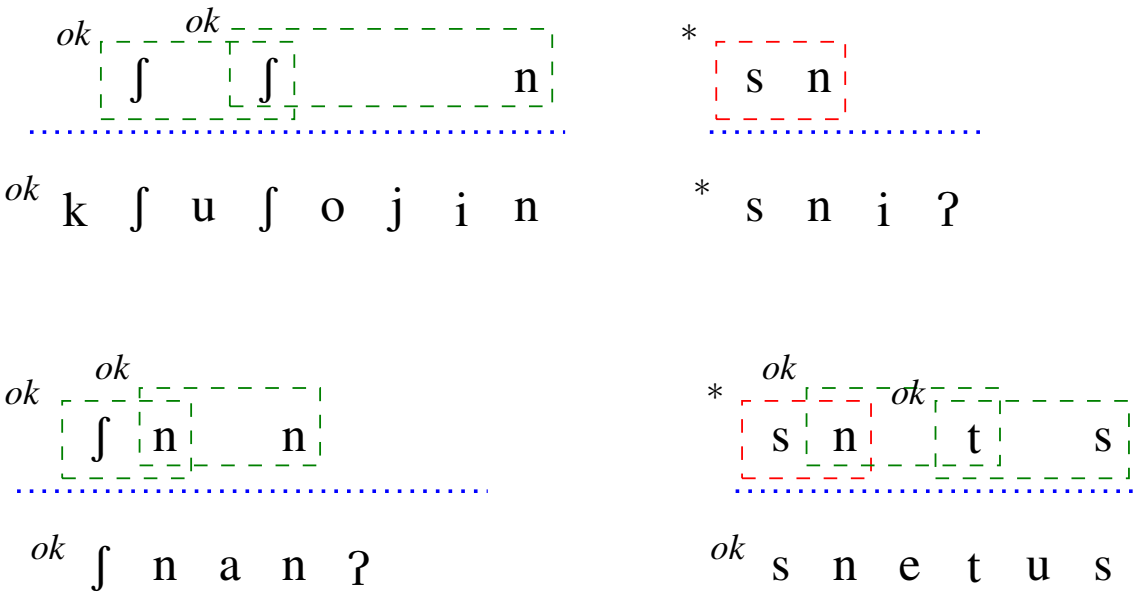
3) Long-distance agreement overrides local disagreement

a.	/s-iʃt-iʃti-jep-us/	sististijepus	"they show him"
b.	/s-net-us/	snetus	"he does it to him"

Ineseño Chumash is not TSL

Ineseño Chumash Sibilant Harmony (Revisited)

- anticipatory sibilant harmony [*sʃ, *sʃ]
- palatalization to avoid local restriction [*sn, *st, *sl]
- sibilant harmony overrides palatalization



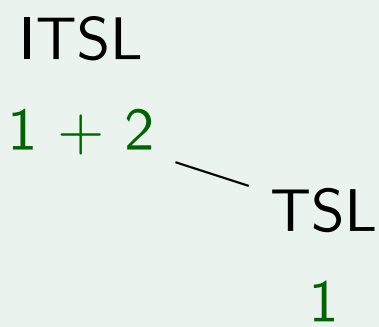
Input-Sensitive TSL (ITSL) Languages

TSL languages are characterized by:

- a 1-local projection function E_T
- strictly k-local constraints applied on T

ITSL (De Santo & Graf 2019)

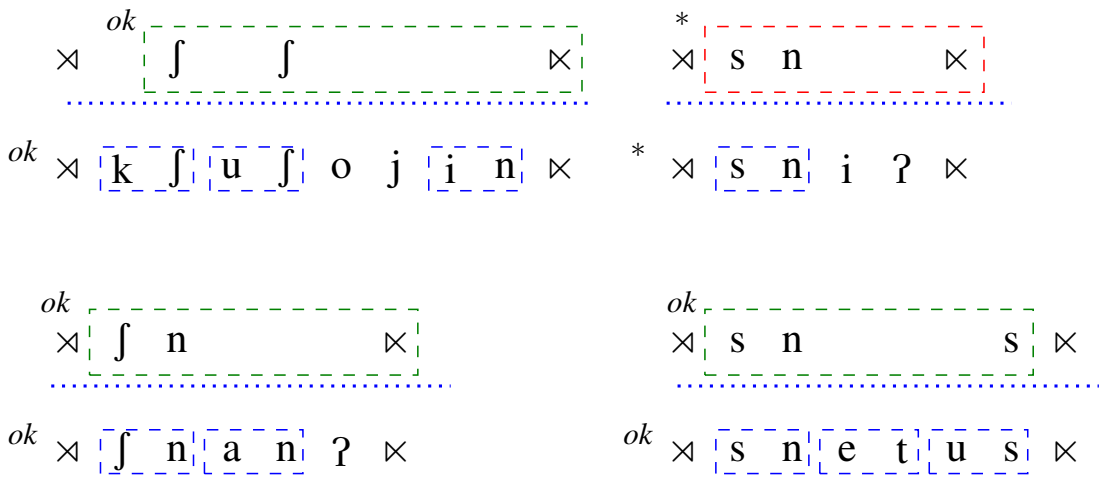
- Tier projection controlled by:
  1. label of segment
  2. n-local context
- strictly k-local constraints applied on T



An ITSL Account of Ineseño Chumash

Ineseño Chumash Sibilant Harmony (Revisited)

- anticipatory sibilant harmony [*sʃ, *sʃ]
- palatalization to avoid local restriction [*sn, *st, *sl]
- sibilant harmony overrides palatalization



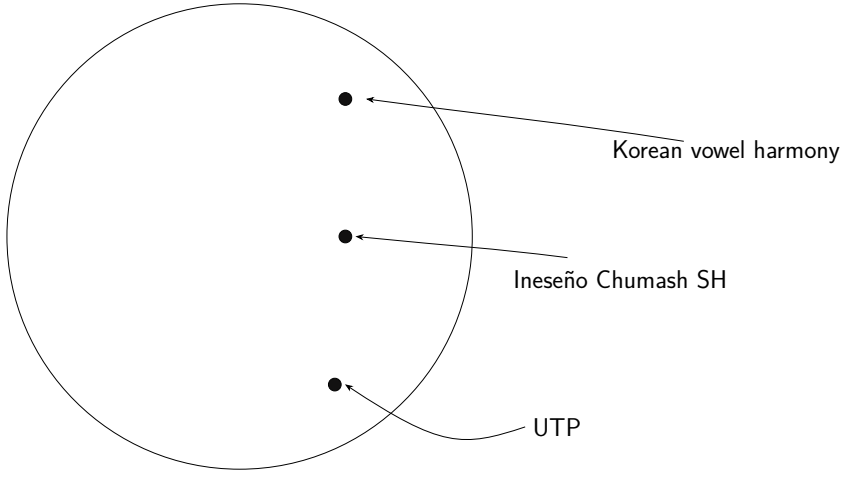
ITSL: Recap

Input-Sensitive TSL (ITSL; De Santo & Graf, 2019)

- n-local projection function
- strictly k-local constraints enforced on T.

- Generalization of TSL
- Covers a variety of patterns
- Gold learnable

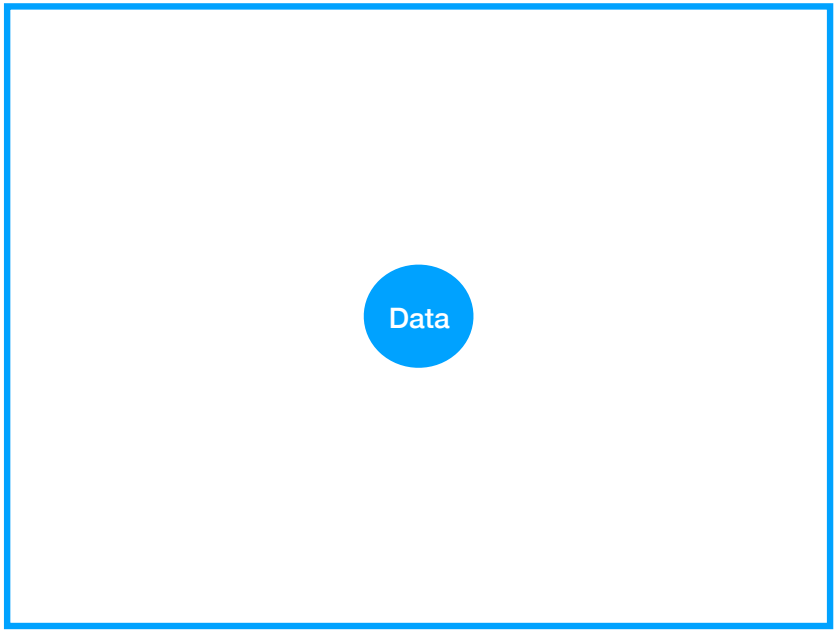
  Efficiently learnable?



Learnability and Formal Grammars

Problem:

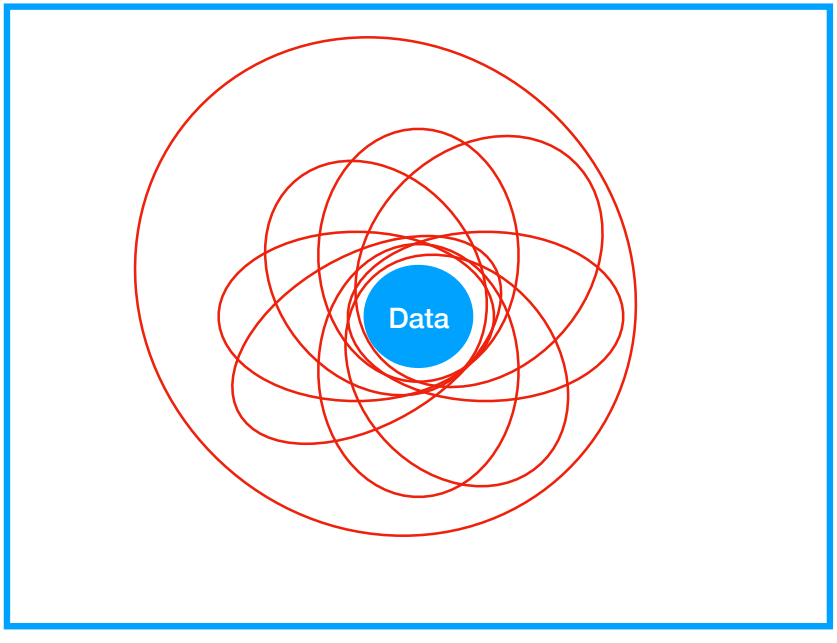
- Unrestricted Hypothesis Spaces



Learnability and Formal Grammars

Problem:

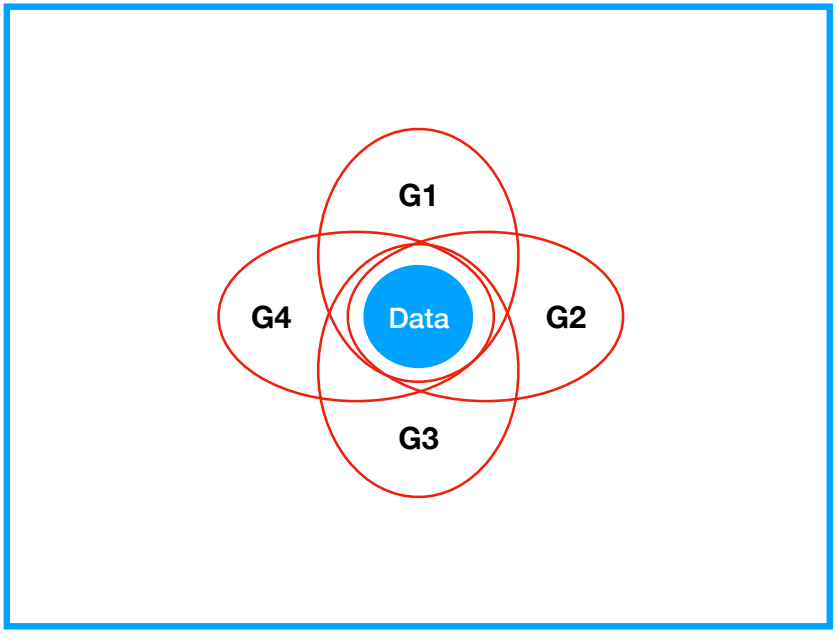
- Unrestricted Hypothesis Spaces



Learnability and Formal Grammars

Solution:

- Structural priors



Invariants

- SL: adjacency
- SP: precedence
- TSL: relativized adjacency
- ITSL: relativized adjacency + local contexts

Learning TSL and ITSL

Learning TSL_k Efficiently

- Batch learning:
  - Jardine & Heinz (2016); Jardine & McMullin (2017)
  - multiple TSL: McMullin, Akenova & De Santo (2019)
- Incrementally: Lambert (2021)

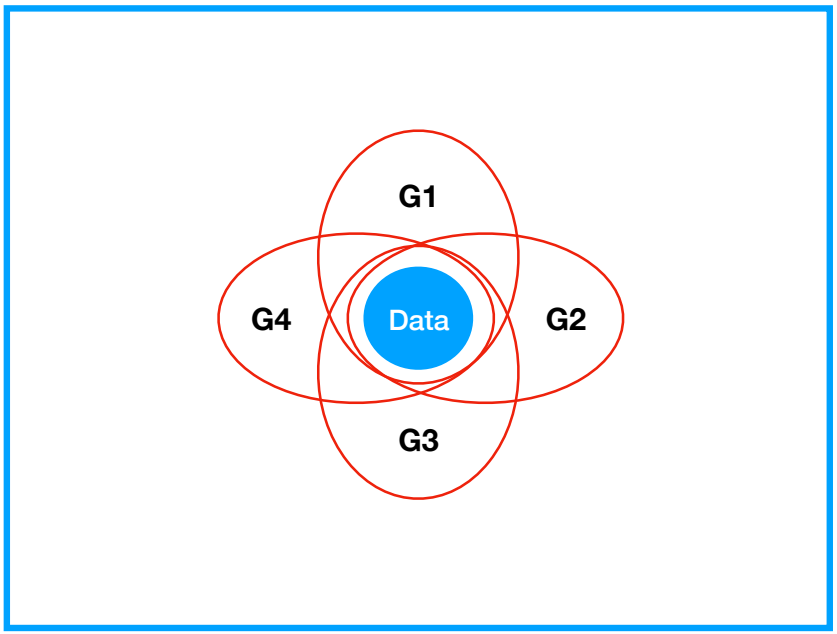
Learning ITSL_k^m Efficiently

- Batch learning:
  - De Santo & Aksënova (2021)
  - multiple ITSL:
    De Santo & Aksënova (2021), Johnson & De Santo (2023)
- Incrementally: Johnson & De Santo (2024)

Learning (M)ITSL Grammars [3]

Solution:

- Structural priors



De Santo & Aksënova (2021):

⇒ Assume relativized locality!

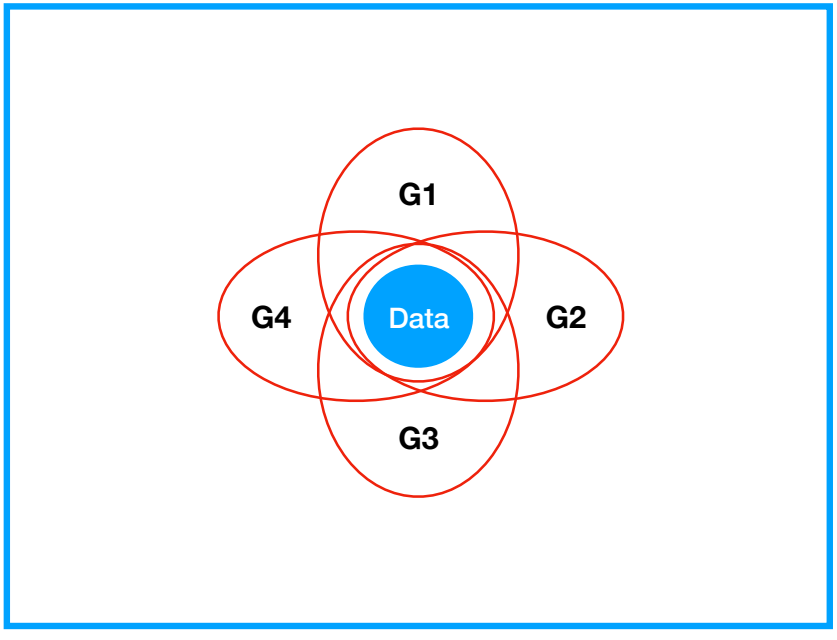
- tiers (but not their content)
- contextual tier-projection
- local tier constraints
- characteristic sample!

[3] McMullin, Aksënova; De Santo (2020), De Santo & Aksënova (2021)

Learning (M)ITSL Grammars [3]

Solution:

- Structural priors



De Santo & Aksënova (2021):

⇒ Assume relativized locality!

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

Guarantees

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

[3] McMullin, Aksënova; De Santo (2020), De Santo & Aksënova (2021)

Evaluating Convergence in Real World Scenarios[4][5]

	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%

Get the Code!

[4] (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.

[5] Aksënova (2020), De Santo & Aksënova (2021), Johnson & De Santo (2023)

Evaluating Convergence in Real World Scenarios[4][5]

	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%

	MITSL
Unbounded tone plateauing	
T	✓
A	100%
Two locally-driven long-distance assimilations (ITSL restrictions)	
T	✓
A	100%
First-Last Harmony	
T	✓
A	100%

Get the Code!

[4] (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.

[5] Aksënova (2020), De Santo & Aksënova (2021), Johnson & De Santo (2023)

Evaluating Convergence in Real World Scenarios[4][5]

	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%

	MITSL
Unbounded tone plateauing	
T	✓
A	100%
Two locally-driven long-distance assimilations (ITSL restrictions)	
T	✓
A	100%
First-Last Harmony	
T	✓
A	100%

Get the Code!

[4] (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.

[5] Aksënova (2020), De Santo & Aksënova (2021), Johnson & De Santo (2023)

Attested vs. Unattested (?) Patterns

Attested: Unbounded Sibilant Harmony

- Every sibilant needs to harmonize

* $ 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 $

Unattested (?): First-Last Harmony

- Harmony only holds between initial and final segments

^{ok} $ h a s x i n t i l a w ʃ $

* $ s a t x i n t i l a w ʃ $

Learnable vs. Unlearnable (?) Patterns

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 & Predictions

Table 4
Types of training items used in the Sibilant Harmony, First-Last, and control conditions. Vowels are omitted. (No training took place in the control condition.)

	Conditions	
Sibilant tier	Sibilant Harmony	First-Last
[s . . . s . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]
[ʃ . . . ʃ . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]
[s . . . ʃ . . . s]	None	[s . . . k . . . ʃ . . . s] [s . . . ʃ . . . k . . . s]
[ʃ . . . s . . . ʃ]	None	[ʃ . . . k . . . s . . . ʃ] [ʃ . . . s . . . k . . . ʃ]

Lai (2015): Stimuli & Predictions

Table 4
Types of training items used in the Sibilant Harmony, First-Last, and control conditions. Vowels are omitted. (No training took place in the control condition.)

	Conditions	
Sibilant tier	Sibilant Harmony	First-Last
[s . . . s . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]
[ʃ . . . ʃ . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]
[s . . . ʃ . . . s]	None	[s . . . k . . . ʃ . . . s] [s . . . ʃ . . . k . . . s]
[ʃ . . . s . . . ʃ]	None	[ʃ . . . k . . . s . . . ʃ] [ʃ . . . s . . . k . . . ʃ]

Table 5
Predicted preferences 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])	FL/SH vs. *FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . s . . . ʃ])	FL/SH vs. FL/*SH (e.g., [s . . . s . . . s] vs. [s . . . ʃ . . . s])
SH	No preference	[s . . . s . . . s] > [s . . . s . . . ʃ]	[s . . . s . . . s] > [s . . . ʃ . . . s]
FL	[s . . . ʃ . . . s] > [s . . . s . . . ʃ]	[s . . . s . . . s] > [s . . . s . . . ʃ]	No preference
Control	No preference	No preference	No preference

Lai (2015): Results[6]

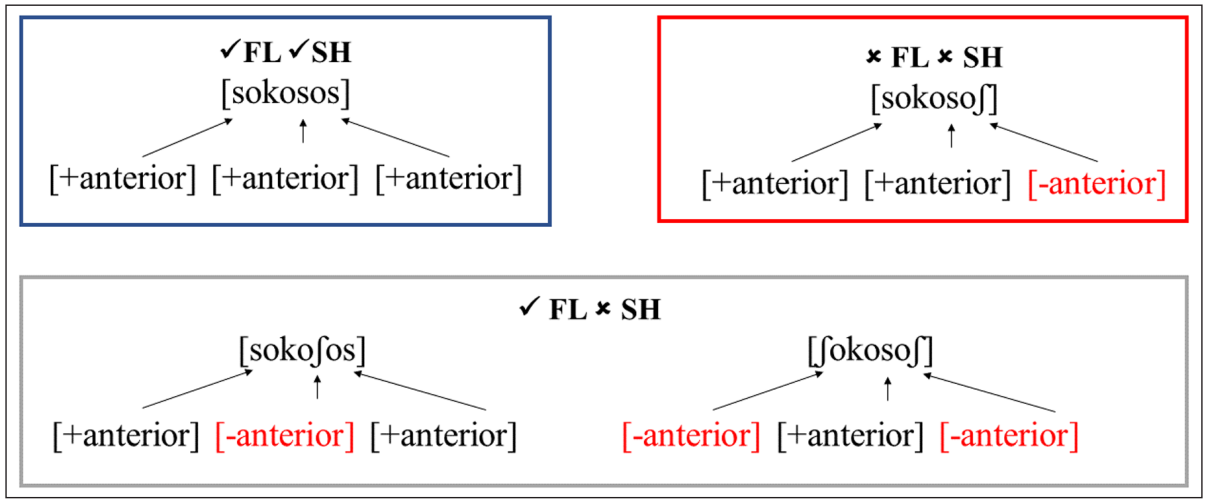


Figure 3: Comparison of SH and FL stimuli.

TL/DR

- SH participants perform in line with a SH grammar
- FL not leaned
- but FL participants perform in line with a SH-like grammar

[6] See Avcu & Hestvik (2020) for a partial replication.

Lai (2015) vs MITSLIA

Table 4
Types of training items used in the Sibilant Harmony, First-Last, and control conditions. Vowels are omitted. (No training took place in the control condition.)

	Conditions	
Sibilant tier	Sibilant Harmony	First-Last
[s . . . s . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]
[ʃ . . . ʃ . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]
[s . . . ʃ . . . s]	None	[s . . . k . . . ʃ . . . s] [s . . . ʃ . . . k . . . s]
[ʃ . . . s . . . ʃ]	None	[ʃ . . . k . . . s . . . ʃ] [ʃ . . . s . . . k . . . ʃ]

What if we train the (M)ITSL batch algorithm on Lai's stimuli?[7]

- is the SH training a characteristic sample?
- is the FL training a characteristic sample?

Get the Code!

[7] Johnson & De Santo (2023); De Santo, Johnson, Aksënova (in prep.)

Lai (2015) vs MITSLIA

Table 4
Types of training items used in the Sibilant Harmony, First-Last, and control conditions. Vowels are omitted. (No training took place in the control condition.)

	Conditions	
Sibilant tier	Sibilant Harmony	First-Last
[s . . . s . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]
[ʃ . . . ʃ . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]
[s . . . ʃ . . . s]	None	[s . . . k . . . ʃ . . . s] [s . . . ʃ . . . k . . . s]
[ʃ . . . s . . . ʃ]	None	[ʃ . . . k . . . s . . . ʃ] [ʃ . . . s . . . k . . . ʃ]

What if we train the (M)ITSL batch algorithm on Lai's stimuli?[7]

- is the SH training a characteristic sample? **No!**
- is the FL training a characteristic sample? **No!**

Get the Code!

[7] Johnson & De Santo (2023); De Santo, Johnson, Aksënova (in prep.)

Lai (2015) vs MITSLIA

Table 4
Types of training items used in the Sibilant Harmony, First-Last, and control conditions. Vowels are omitted. (No training took place in the control condition.)

	Conditions	
Sibilant tier	Sibilant Harmony	First-Last
[s . . . s . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]
[ʃ . . . ʃ . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]
[s . . . ʃ . . . s]	None	[s . . . k . . . ʃ . . . s] [s . . . ʃ . . . k . . . s]
[ʃ . . . s . . . ʃ]	None	[ʃ . . . k . . . s . . . ʃ] [ʃ . . . s . . . k . . . ʃ]

What if we train the (M)ITSL batch algorithm on Lai's stimuli?[7]

- is the SH training a characteristic sample? **No!**
- is the FL training a characteristic sample? **No!**
- what if: information about natural classes?
  E.g. vowels

Get the Code!

[7] Johnson & De Santo (2023); De Santo, Johnson, Aksënova (in prep.)

Lai (2015) vs. MITSLIA: Masked SH Training

Performance of ITSL Batch Learner:

- trained on Lai (2015)'s SH input
  $\rightarrow$ 40 words, masking vowels;
- tested on acceptance of 96 individual test strings.

```
Evaluating g_sh
What should this accept? (highlighted yellow)                    svkvSvs SvkvSvs svkvsvs svkv
svS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svkv
Svs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svkv
Svs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svkv
svS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svkv
svs svSvkvs svsvkvs Svsvkvs SvsvkvS SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS SvkvsvS Svsvkvs svsv
kvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svsv
kvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvsvkvS SvSvkvS svkvsvS SvkvsvS svkvsvS svkv
svs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
What does this accept?   (highlighted magenta)                   svkvSvs SvkvSvs svkvsvs svkv
svS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svkv
Svs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svkv
Svs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svkv
svS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svkv
svs svSvkvs svsvkvs SvsvkvS SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS SvkvsvS Svsvkvs svsv
kvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svsv
kvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvsvkvS SvSvkvS svkvsvS SvkvsvS svkvsvS svkv
svs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
Does the grammar make the same judgement as the sh acceptor? 100.00% svkvSvs SvkvSvs svkvsvs svk
vsvS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svk
vSvs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svk
vSvs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svk
vsvS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svk
vsvs svSvkvs svsvkvs SvsvkvS SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS SvkvsvS Svsvkvs svs
vkvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svs
vkvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvsvkvS SvSvkvS svkvsvS SvkvsvS svkvsvS svk
vsvs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
```

Lai (2015) vs. MITSLIA: Masked FL Training

Performance of ITSL Batch Learner:

- trained on Lai (2015)'s FL input
  $\rightarrow$ 40 words, masking vowels;
- tested on acceptance of 96 individual test strings.

```
Evaluating g_fl
What should this accept? (highlighted yellow)                    svkvSvs SvkvSvs svkvsvs svkv
svS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svkv
Svs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svkv
Svs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svkv
svS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svkv
svs svSvkvs svsvkvs Svsvkvs SvsvkvS SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS SvkvsvS Svsvkvs svsv
kvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svsv
kvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvsvkvS SvSvkvS svkvsvS SvkvsvS svkvsvS svkv
svs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
What does this accept?   (highlighted magenta)                   svkvSvs SvkvSvs svkvsvs svkv
svS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svkv
Svs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svkv
Svs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svkv
svS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svkv
svs svSvkvs svsvkvs SvsvkvS Svsvkvs SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS SvkvsvS Svsvkvs svsv
kvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svsv
kvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvsvkvS SvSvkvS svkvsvS SvkvsvS svkvsvS svkv
svs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
Does the grammar make the same judgement as the fl acceptor? 100.00% svkvSvs SvkvSvs svkvsvs svk
vsvS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svk
vSvs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svk
vSvs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svk
vsvS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svk
vsvs svSvkvs svsvkvs SvsvkvS Svsvkvs SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS SvkvsvS Svsvkvs svs
vkvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svs
vkvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvsvkvS SvSvkvS svkvsvS SvkvsvS svkvsvS svk
vsvs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
```

Lai (2015): Intensive FL Training

Table 4
Types of training items used in the Sibilant Harmony, First-Last, and control conditions. Vowels are omitted. (No training took place in the control condition.)

	Conditions	
Sibilant tier	Sibilant Harmony	First-Last
[s . . . s . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]
[ʃ . . . ʃ . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]
[s . . . ʃ . . . s]	None	[s . . . k . . . ʃ . . . s] [s . . . ʃ . . . k . . . s]
[ʃ . . . s . . . ʃ]	None	[ʃ . . . k . . . s . . . ʃ] [ʃ . . . s . . . k . . . ʃ]

Lai (2015): Intensive FL Training

Table 4
Types of training items used in the Sibilant Harmony, First-Last, and control conditions. Vowels are omitted. (No training took place in the control condition.)

	Conditions	
Sibilant tier	Sibilant Harmony	First-Last Intensive
[s . . . s . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]
[ʃ . . . ʃ . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]
[s . . . ʃ . . . s]	None	[s . . . k . . . ʃ . . . s] [s . . . ʃ . . . k . . . s]
[ʃ . . . s . . . ʃ]	None	[ʃ . . . k . . . s . . . ʃ] [ʃ . . . s . . . k . . . ʃ]

Lai (2015): Intensive FL Training

Table 4
Types of training items used in the Sibilant Harmony, First-Last, and control conditions. Vowels are omitted. (No training took place in the control condition.)

	Conditions	
Sibilant tier	Sibilant Harmony	First-Last Intensive
[s . . . s . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]	[s . . . k . . . s . . . s] [s . . . s . . . k . . . s]
[ʃ . . . ʃ . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]	[ʃ . . . k . . . ʃ . . . ʃ] [ʃ . . . ʃ . . . k . . . ʃ]
[s . . . ʃ . . . s]	None	[s . . . k . . . ʃ . . . s] [s . . . ʃ . . . k . . . s]
[ʃ . . . s . . . ʃ]	None	[ʃ . . . k . . . s . . . ʃ] [ʃ . . . s . . . k . . . ʃ]

TL/DR

- FL not leaned but IFL participants exhibit regularities...
- ... in line with a sibilant disharmony rule

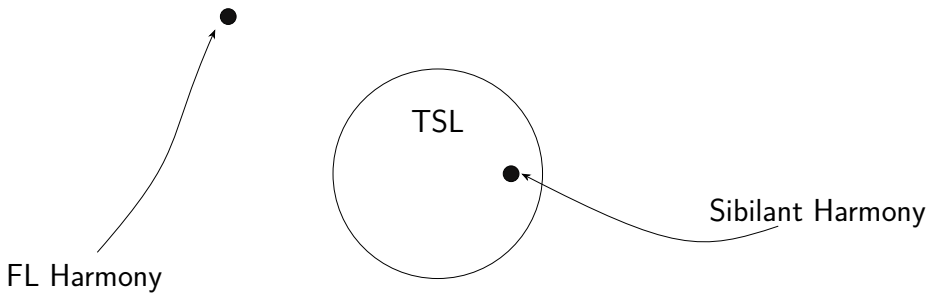
Lai (2015) vs. MITSLIA: Masked IFL Training

Performance of ITSL Batch Learner:

- trained on Lai (2015)'s Intensive FL input

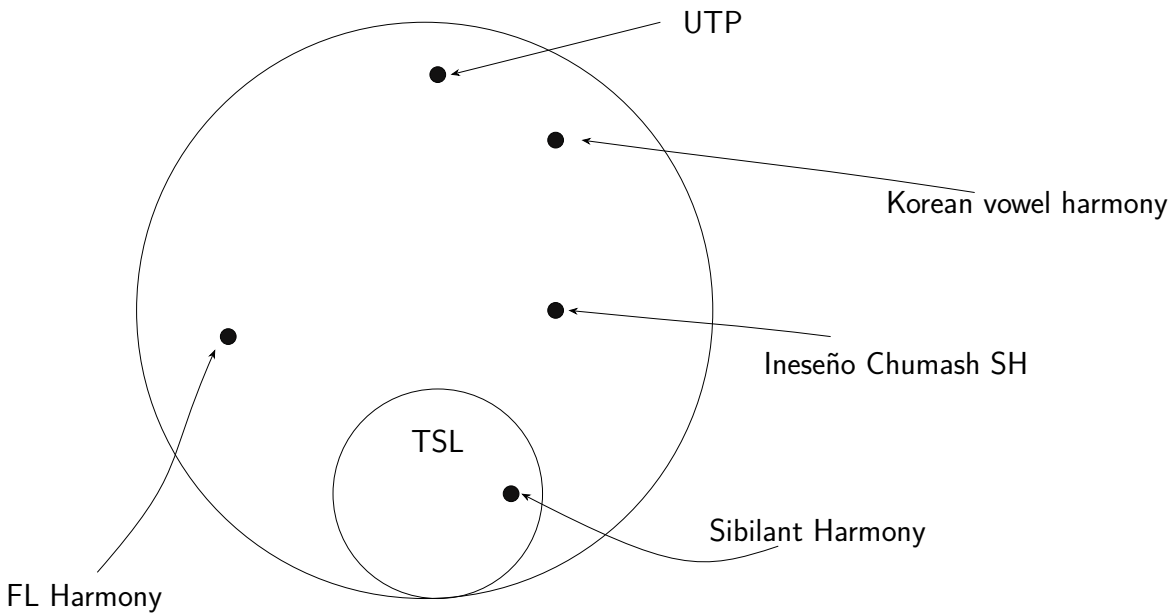
```
Evaluating g_ifl
What should this accept? (highlighted yellow)                    svkvSvs SvkvSvs svkvsvs svkv
svS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svkv
Svs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svkv
Svs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svkv
svS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svkv
svs svSvkvs svsvkvs Svsvkvs SvsvkvS SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS Svsvkvs svsv
kvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svsv
kvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS svkvsvS SvkvsvS svkvsvS svkv
svs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
What does this accept?   (highlighted magenta)                   svkvSvs SvkvSvs svkvsvs svkv
svS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svkv
Svs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svkv
Svs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svkv
svS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svkv
svs svSvkvs svsvkvs Svsvkvs SvsvkvS SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS SvkvsvS Svsvkvs svsv
kvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svsv
kvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS svkvsvS SvkvsvS svkvsvS svkv
svs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
Does the grammar make the same judgement as the sh acceptor? 33.33% svkvSvs SvkvSvs svkvsvs svkv
svS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svkv
Svs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svkv
Svs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svkv
svS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svkv
svs svSvkvs svsvkvs Svsvkvs SvsvkvS SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS SvkvsvS Svsvkvs svsv
kvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svsv
kvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS svkvsvS SvkvsvS svkvsvS svkv
svs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
Does the grammar make the same judgement as the fl acceptor? 66.67% svkvSvs SvkvSvs svkvsvs svkv
svS SvSvkvS SvsvkvS SvsvkvS Svsvkvs SvSvkvS svSvkvS svkvsvs svkvSvs svSvkvs svSvkvS SvkvSvS Svkv
Svs svsvkvs svSvkvs SvkvsvS svkvsvS svsvkvs Svsvkvs SvkvSvS SvkvsvS svSvkvs svSvkvS SvkvSvS Svkv
Svs SvkvSvS SvkvsvS SvkvsvS svkvsvS SvSvkvS svSvkvS svsvkvs svSvkvs svkvSvs SvkvSvs svkvsvs svkv
svS svkvsvs svkvSvs SvsvkvS Svsvkvs svsvkvs Svsvkvs SvSvkvS SvsvkvS SvkvSvs svkvSvs svkvsvS svkv
svs svSvkvs svsvkvs Svsvkvs SvsvkvS SvkvSvs SvkvSvS SvsvkvS SvSvkvS svkvsvS SvkvsvS Svsvkvs svsv
kvs svkvSvs svkvsvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS Svsvkvs SvsvkvS Svsvkvs svsv
kvs svSvkvs svsvkvs svSvkvS svSvkvs svSvkvS SvSvkvS SvkvsvS SvkvSvS svkvsvS SvkvsvS svkvsvS svkv
svs svkvSvs svkvsvs SvkvSvs svkvSvs SvkvSvS SvkvsvS SvkvSvS
```

Subregularity, Typology, & AGL in the TSL Neighborhood



- Comparing behavior within classes is essential
- Different characterizations and/or non-FLT explanations
  (Garret & Johnson 2011, Endress et al. 2005; Endress & Mehler 2010)
- How does this tie to cognitive resources (e.g. WM)?
  (Baddeley 2000, Pierce et al. 2017)

Subregularity, Typology, & AGL in the TSL Neighborhood



- Comparing behavior within classes is essential
- Different characterizations and/or non-FLT explanations
  (Garret & Johnson 2011, Endress et al. 2005; Endress & Mehler 2010)
- How does this tie to cognitive resources (e.g. WM)?
  (Baddeley 2000, Pierce et al. 2017)

Conclusion

FLT & Typology: Representational primitives!

- Do we look for restrictive theories?
- YES! But driven by invariants, not boundaries

AGL: Probe which primitives humans are sensitive to!

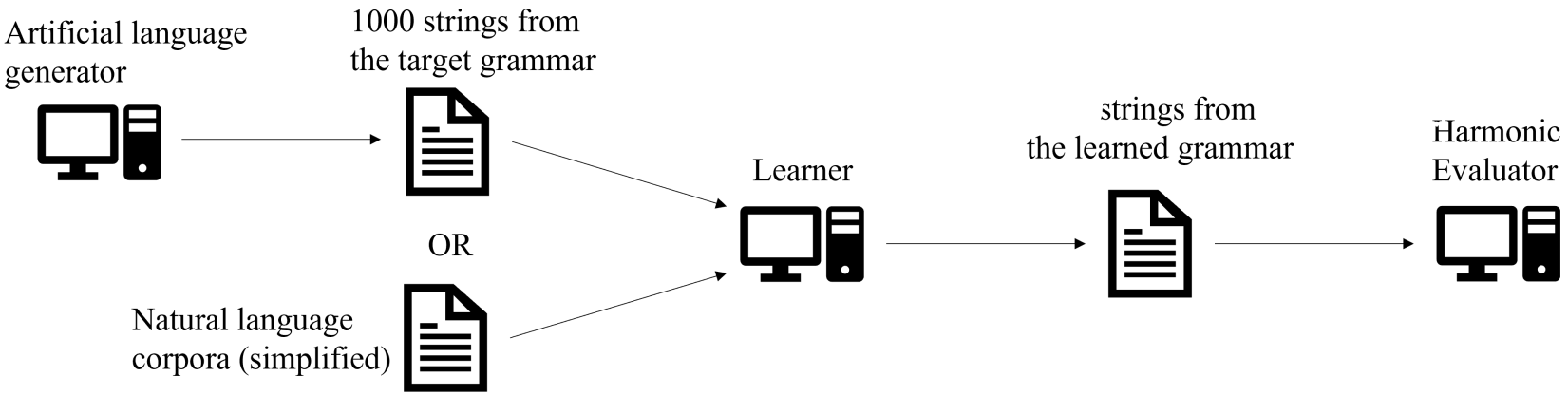
- Compare between classes and within classes!
- Formal Learners provide insights/ground truths
- Not just about complexity: TSL vs. SP?
- AGL Design: Not just FLT & Typology!
  - Working memory, attention, input source, etc.
  - Look at psycholinguists/laboratory phonologists/etc.

Thank You!

Special thanks to Jon Rawski, Caleb Belth, Yang Wang, Alëna Aksënova, Jacob Johnson, and members of the Utah Comp Ling Working Group for discussions on several aspects of this work!

Appendix

Evaluation[8]



- TSL and ITSL mplemented in Python 3 following requirements of SigmaPie
- Artificial datasets exemplifying different subregular classes
- 3 simplified natural language corpora (German, Finnish, Turkish)
- Proportion of first 5000 strings accepted by the learned grammar also accepted by the target grammar
- Learning + evaluation iterated 10 times

[8] Aksenova (2020); Johnson & De Santo (2023, 2024)

Online Learner: Evaluation

	TSL	ITSL
Word-final devoicing		
T	✓	✓
A	100%	100%
N_G	100%	100%
Single vowel harmony without blocking		
T	✓	✓
A	100%	100%
N_F	100%	100%
Single vowel harmony with blocking		
T	✓	✓
A	100%	100%
Several vowel harmonies without blocking		
T	✓	✓
A	100%	100%
Several vowel harmonies with blocking		
T	✓	✓
A	100%	100%
N_T	100%	100%

Online Learner: Evaluation

	TSL	ITSL
Word-final devoicing		
T	✓	✓
A	100%	100%
N_G	100%	100%
Single vowel harmony without blocking		
T	✓	✓
A	100%	100%
N_F	100%	100%
Single vowel harmony with blocking		
T	✓	✓
A	100%	100%
Several vowel harmonies without blocking		
T	✓	✓
A	100%	100%
Several vowel harmonies with blocking		
T	✓	✓
A	100%	100%
N_T	100%	100%

	TSL	ITSL
Unbounded tone plateauing		
T	✗	✓
A	9.97% (0.51%)	100%
First-Last Assimilation		
T	✗	✓
A	50.02%	100%
Locally-driven long-distance assimilation (ITSL restriction)		
T	✗	✓
A	94.88% (0.15%)	100%

Why Care?

Regularity then provides a well-defined proposed computational universal that is sufficiently expressive while ruling out a great many non-phonological patterns.

(Chandlee 2024)

The larger goal of this work is to formally delimit the boundary between possible and impossible grammatical patterns, as evidenced by attested and unattested language patterns.

(McCollum et al.)

The Weak Subregular Hypothesis (Heinz, 2018:155) restricts phonotactic generalizations to strings over tiers [...] While it is not difficult to find apparent counterexamples to the Weak Subregular Hypothesis cross-linguistically, they are often explained by invoking standard phonological structures.

(Lamont, u.r.)

Lai (2015): Results[9]

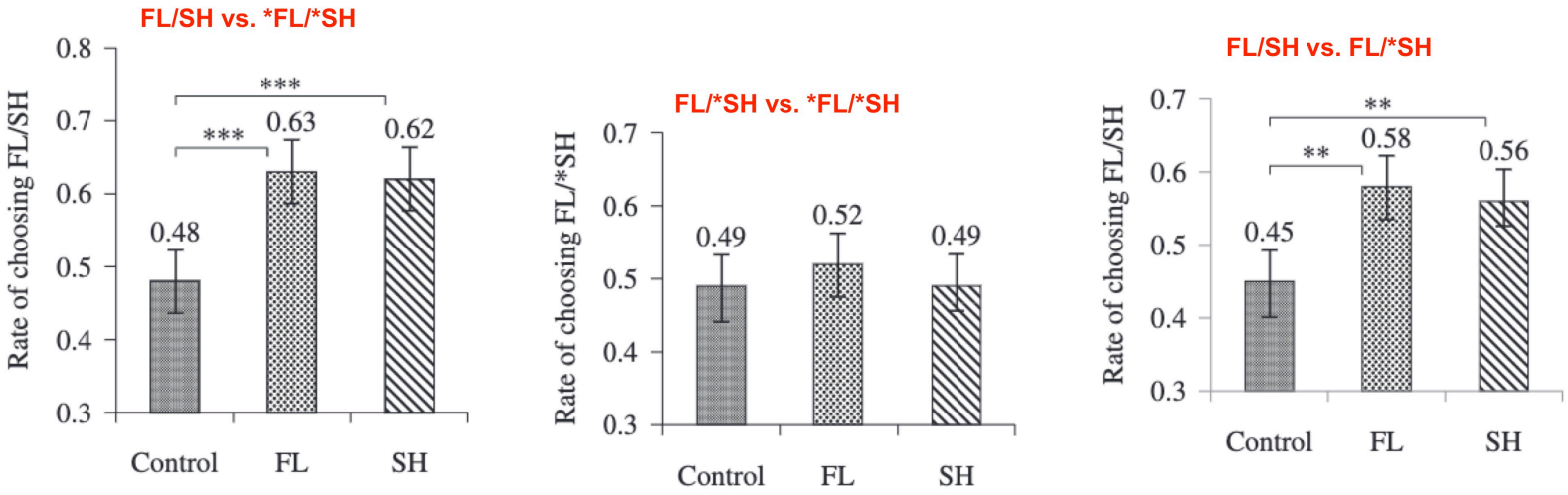


Table 6
Predicted results with respect to the control group for each test pairing if Sibilant Harmony and First-Last Assimilation grammars were internalized

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

[9] See Avcu & Hestevik (2020) for a partial replication.

Lai (2015): Results (Part 2)

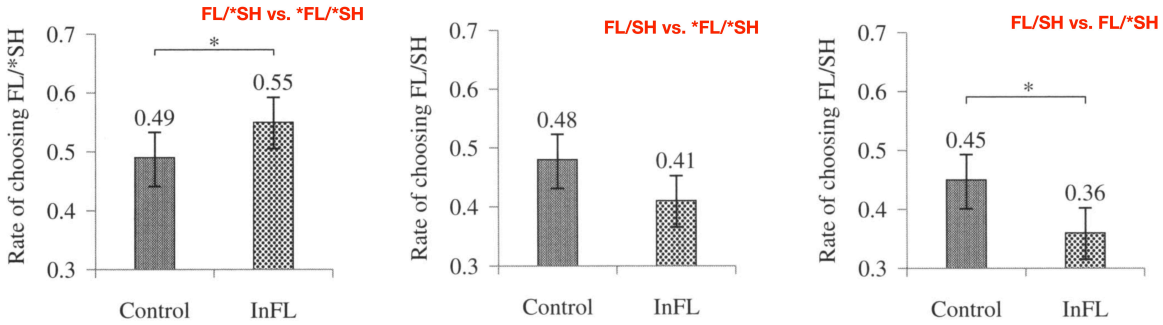


Table 6
Predicted results with respect to the control group for each test pairing if Sibilant Harmony and First-Last Assimilation grammars were internalized

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