Vilnius
University

Phonotactic well-formedness is recombinant

19th Old World Conference on Phonology (OCP 19)

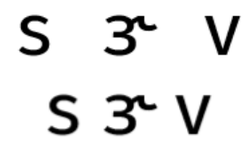
Ignas Rudaitis

28 January 2022

serve

S 3ˇ V

serve



serve

serve

verse

serve

verse

serve

verse

serve

a type of **unary** recombination

verse

a type of **ternary** recombination

spill

s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ

spill

s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ

spill

s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ

s p ɪ l

spill

s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ

s p ɪ l

spill

s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ

s p ɪ l

spill

s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ

p ɪ l oʊ

s p ɪ l

spill

s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ

p ɪ l oʊ

s p ɪ l

spill

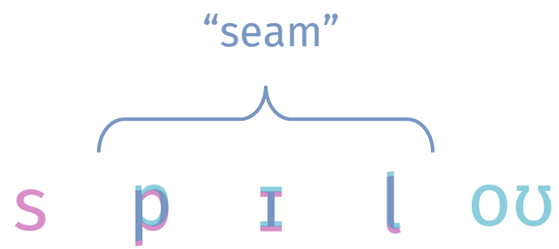
s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ



spill

s p ɪ l

pillow

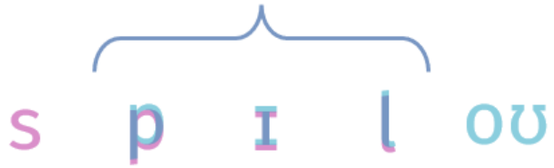
p ɪ l oʊ

so

s oʊ

s oʊ

“seam”



spill

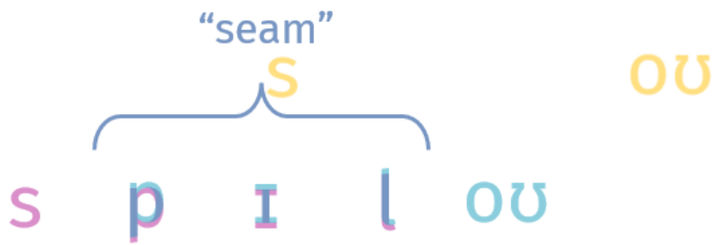
s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ



°*spillow*

spill

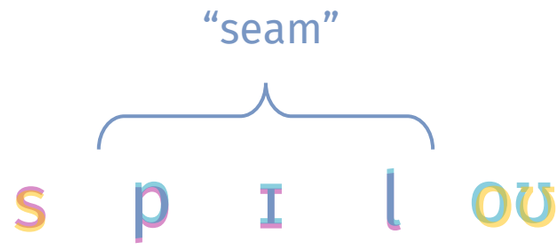
s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ



°*spillow*

well-formed

spill

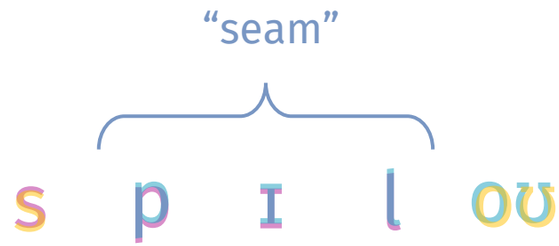
s p ɪ l

pillow

p ɪ l oʊ

so

s oʊ



°*spillow*

well-formed

spill

s p ɪ l

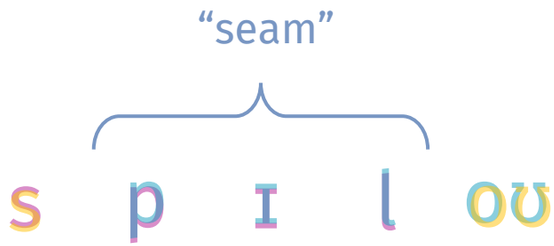
well-formed

pillow

p ɪ l oʊ

so

s oʊ



°*spillow*

well-formed

spill

s p ɪ l

well-formed

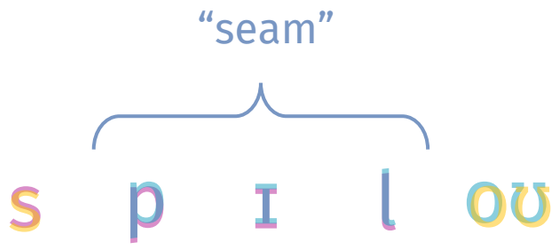
pillow

p ɪ l oʊ

well-formed

so

s oʊ



°*spillow*

well-formed

spill

s p ɪ l

well-formed

pillow

p ɪ l oʊ

well-formed

so

s oʊ

“seam”

s p ɪ l oʊ

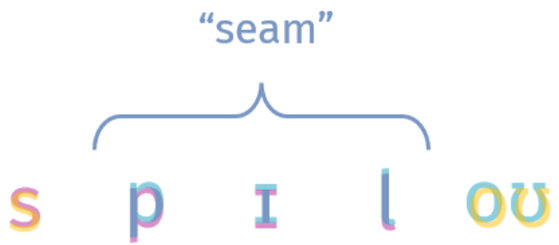
°*spillow*

well-formed

well-formed ∧ well-formed well-formed

spill *pillow* *so*

s p ɪ l p ɪ l oʊ s oʊ



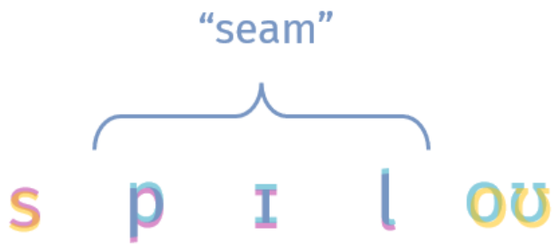
°*spillow*

well-formed

well-formed ∧ well-formed ∧ well-formed

spill *pillow* *so*

s p ɪ l p ɪ l oʊ s oʊ



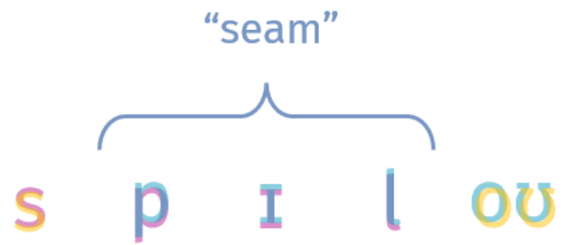
°*spillow*

well-formed

well-formed ∧ well-formed ∧ well-formed ⇒

spill *pillow* *so*

s p ɪ l p ɪ l oʊ s oʊ



°*spillow*

well-formed

well-formed ∧ well-formed ∧ well-formed ⇒

spill *pillow* *so*

s p ɪ l p ɪ l **well-formed** s oʊ

“seam”

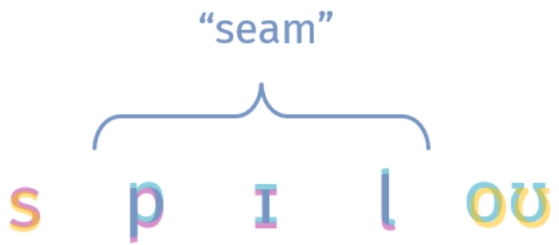
s p ɪ l oʊ

°*spillow*

well-formed ∧ well-formed ∧ well-formed ⇒ **well-formed**

spill *pillow* *so*

s p ɪ l p ɪ l oʊ s oʊ

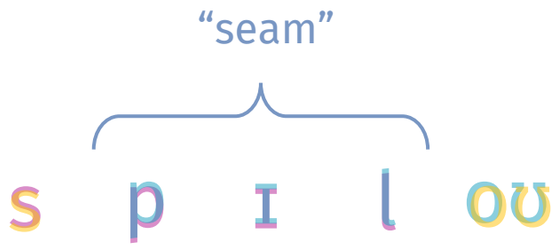


°*spillow*

well-formed ∧ well-formed ∧ well-formed ⇒ **well-formed**

spill *pillow* *so*

s p ɪ l p ɪ l oʊ s oʊ



°*spillow*

well-formed $\wedge$ well-formed $\wedge$ well-formed $\Rightarrow$ **well-formed**

XY well-formed ∧ YZ well-formed ∧ XZ well-formed ⇒ **XYZ well-formed**

XY well-formed $\wedge$ YZ well-formed $\wedge$ XZ well-formed $\Rightarrow$ **XYZ well-formed**

XY well-formed ∧ YZ well-formed ∧ XZ well-formed ⇒ **XYZ well-formed**

Hypothesis: this implication holds cross-linguistically
for nearly all segmental constraints

Coverage of the recombinant property

Bigram constraints	most local constraints
Trigram constraints	sonority sequencing
Final/initial constraints	final devoicing
Classical harmony	Finnish VH
Harmony, simple blocking	Imdlawn Tashlhiyt CH
Harmony, local bipartite blocking	Yaka CH
Harmony, distal bipartite blocking	Tutrugbu VH
Long-distance dissimilation	Latin /l ~ r/ at some stage (?)
Quantification	“at least one vowel”
Suprasegmentals	tone, stress
Any intersection of the above	multiple harmony

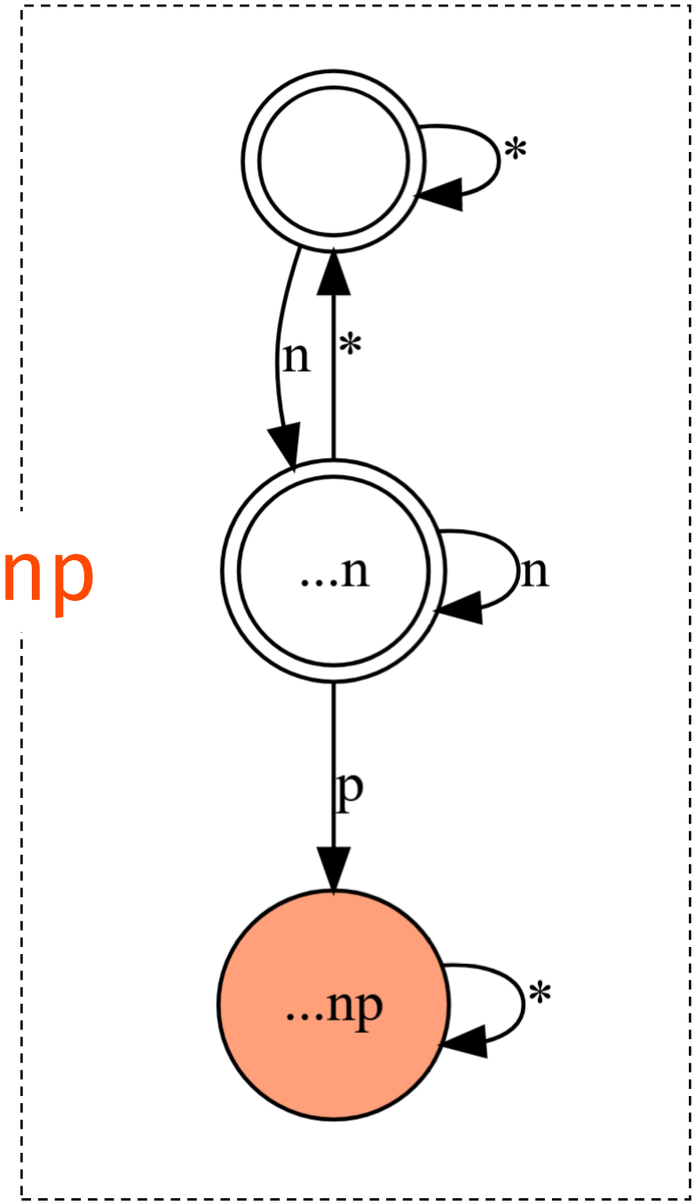
Outline

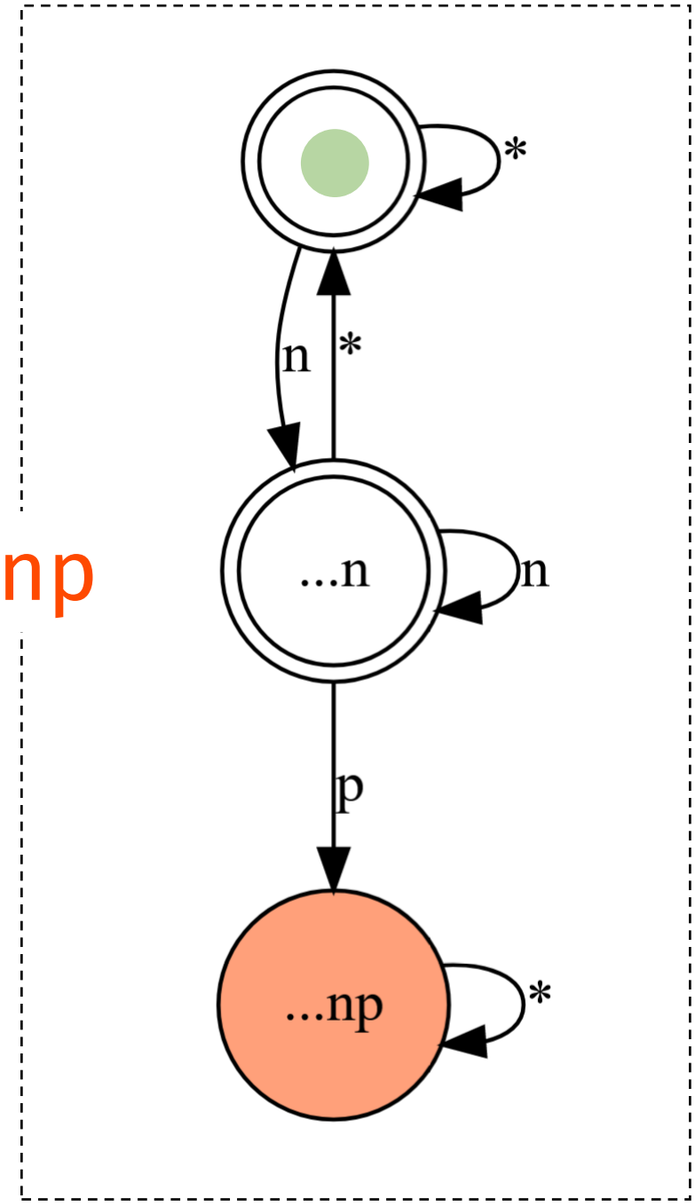
1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

Outline

1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

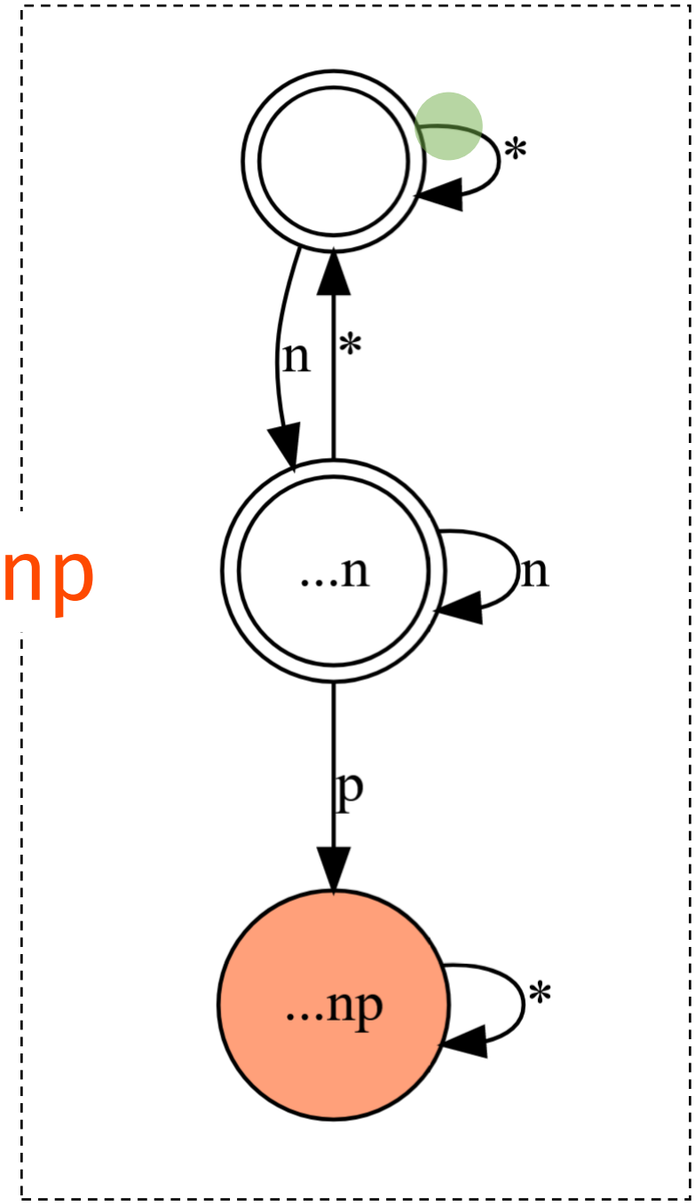
*np





*np

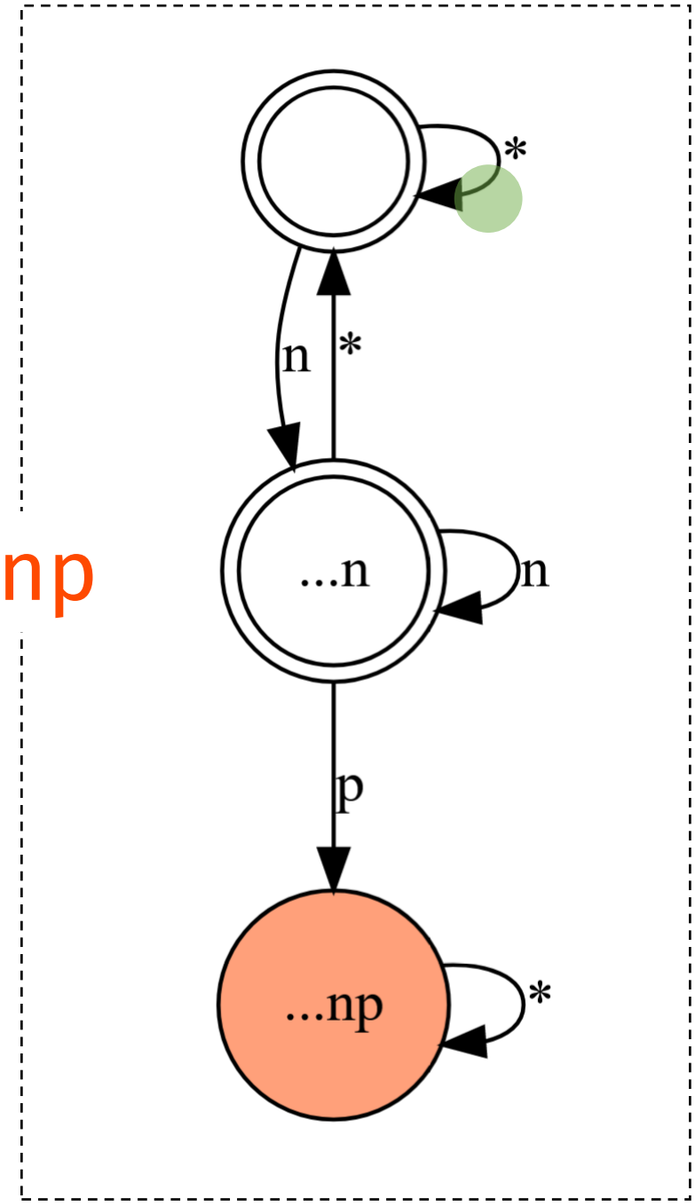
anpa



*np

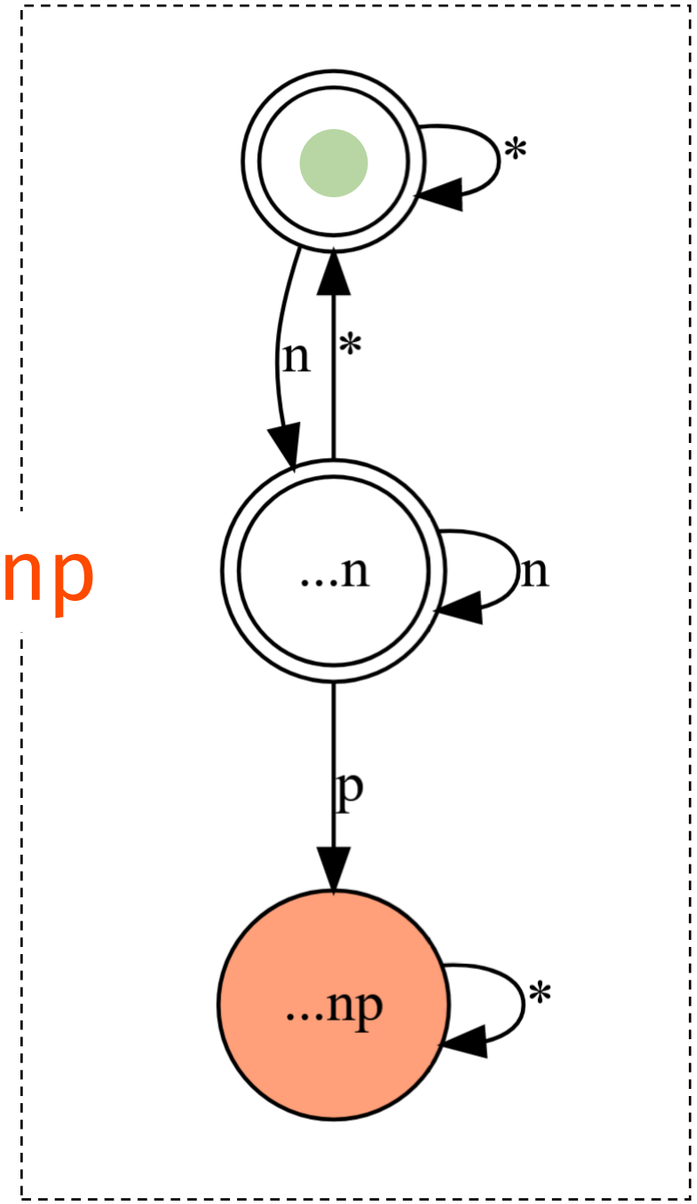
anpa

*np



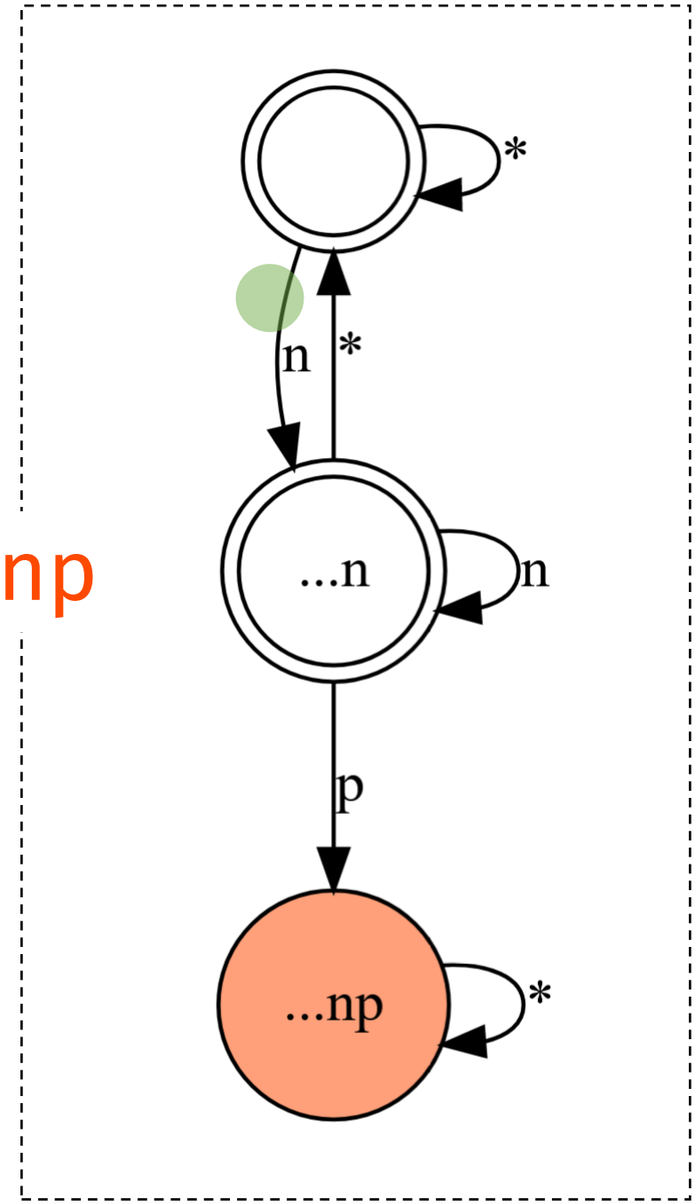
anpa

*np



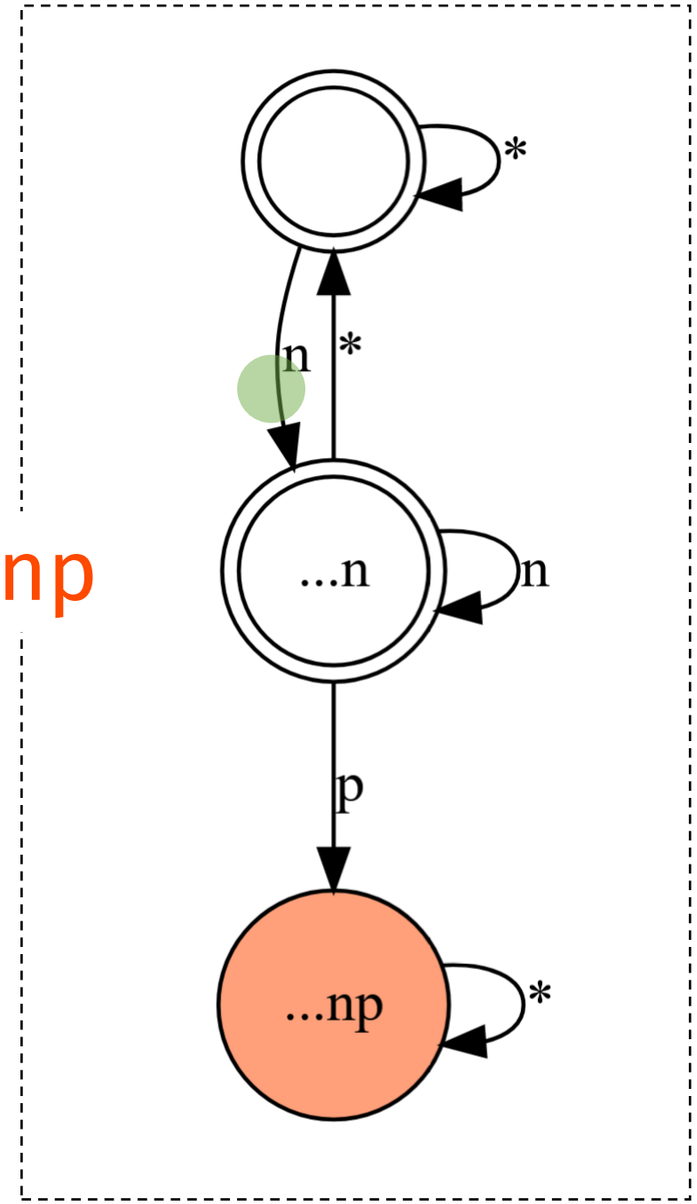
anpa

*np



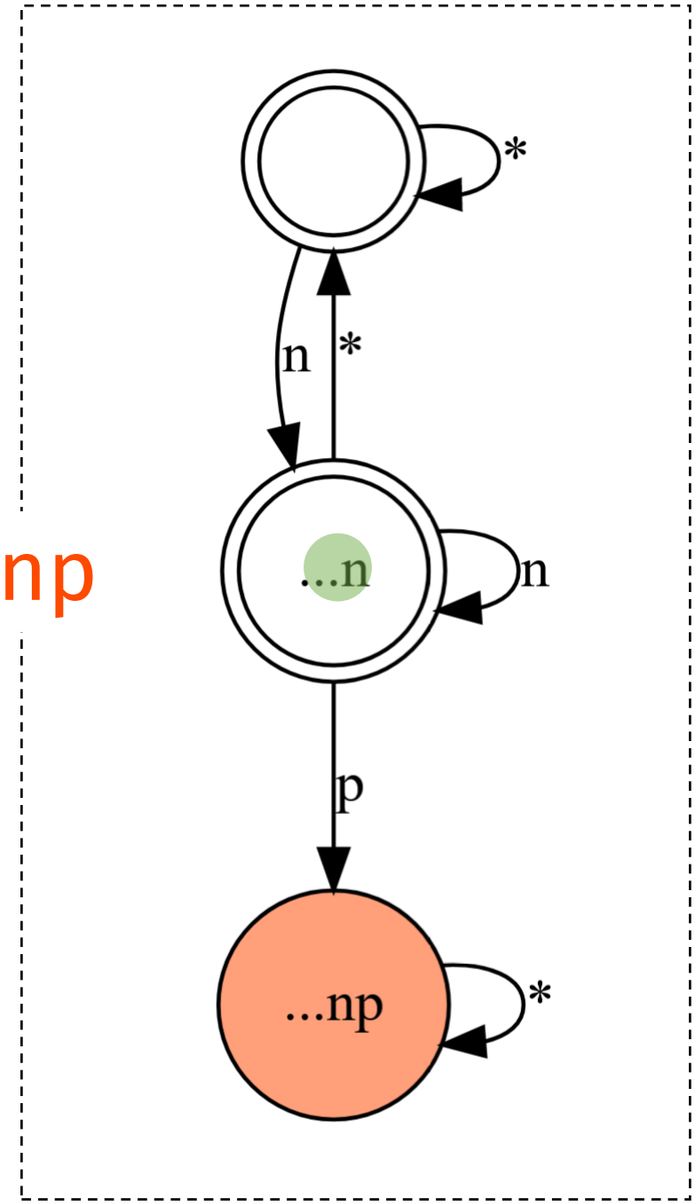
anpa

*np



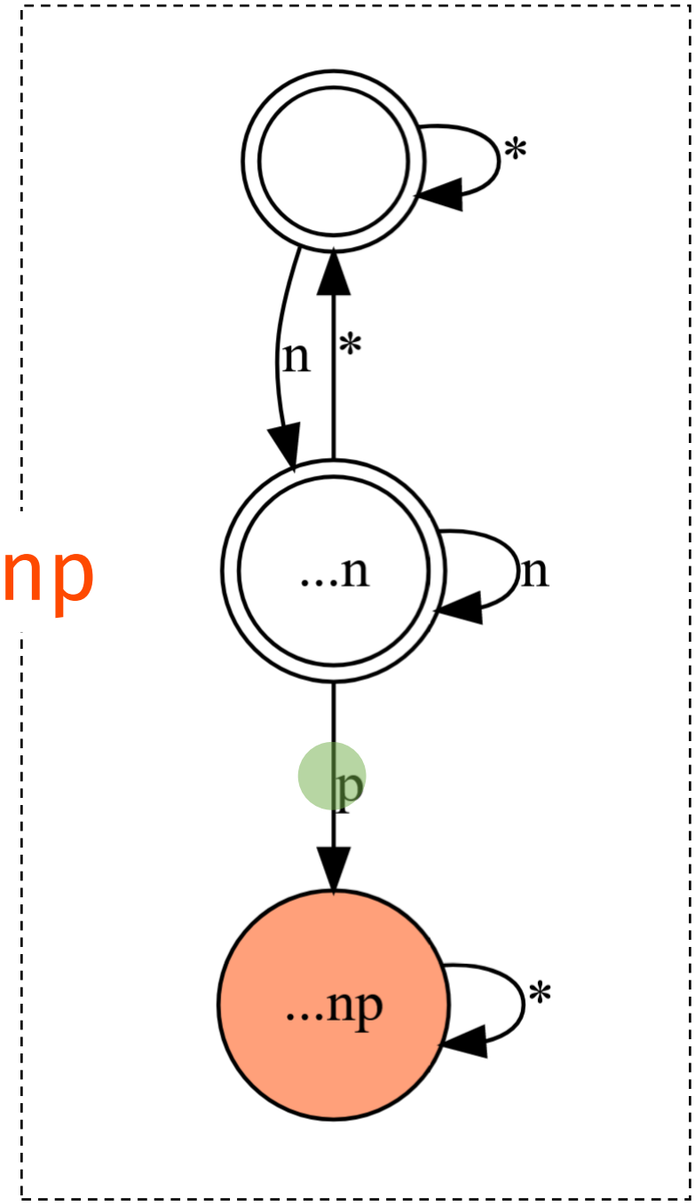
anpa

*np



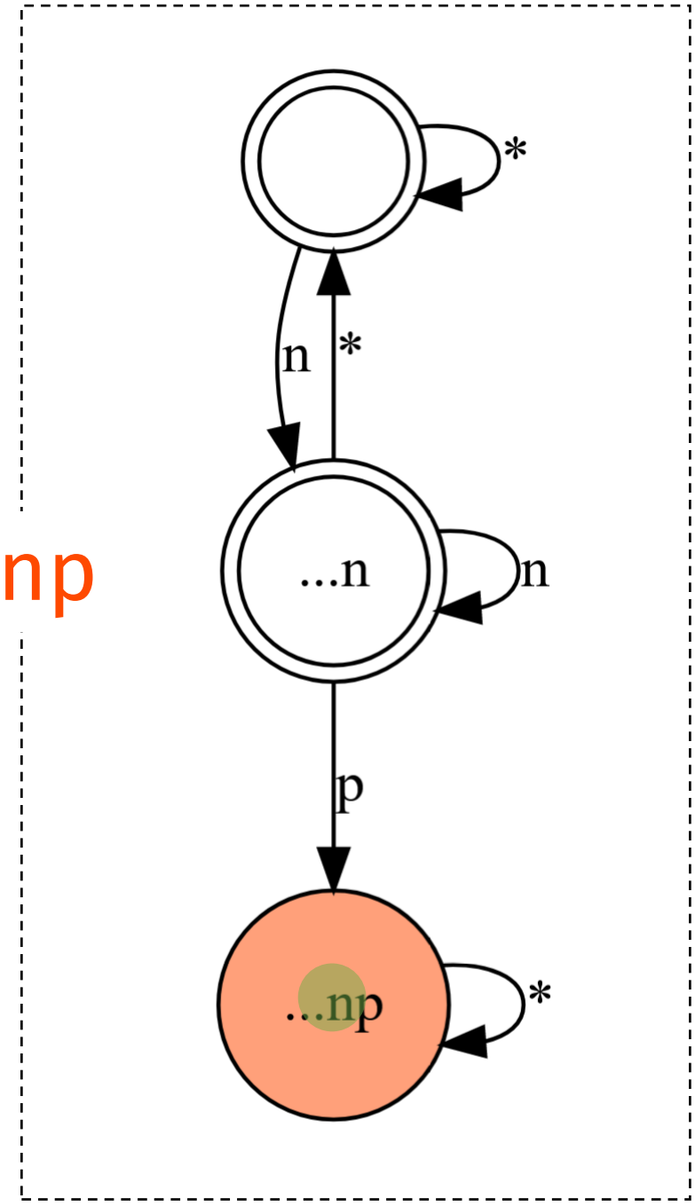
anpa

*np



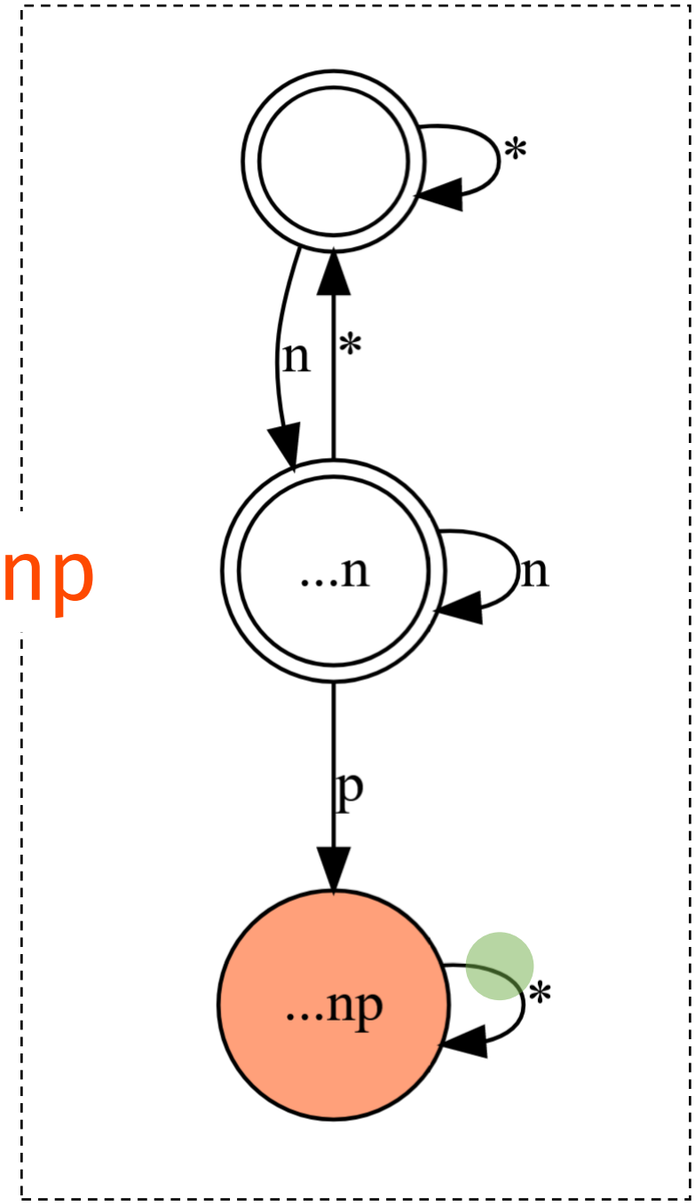
anpa

*np



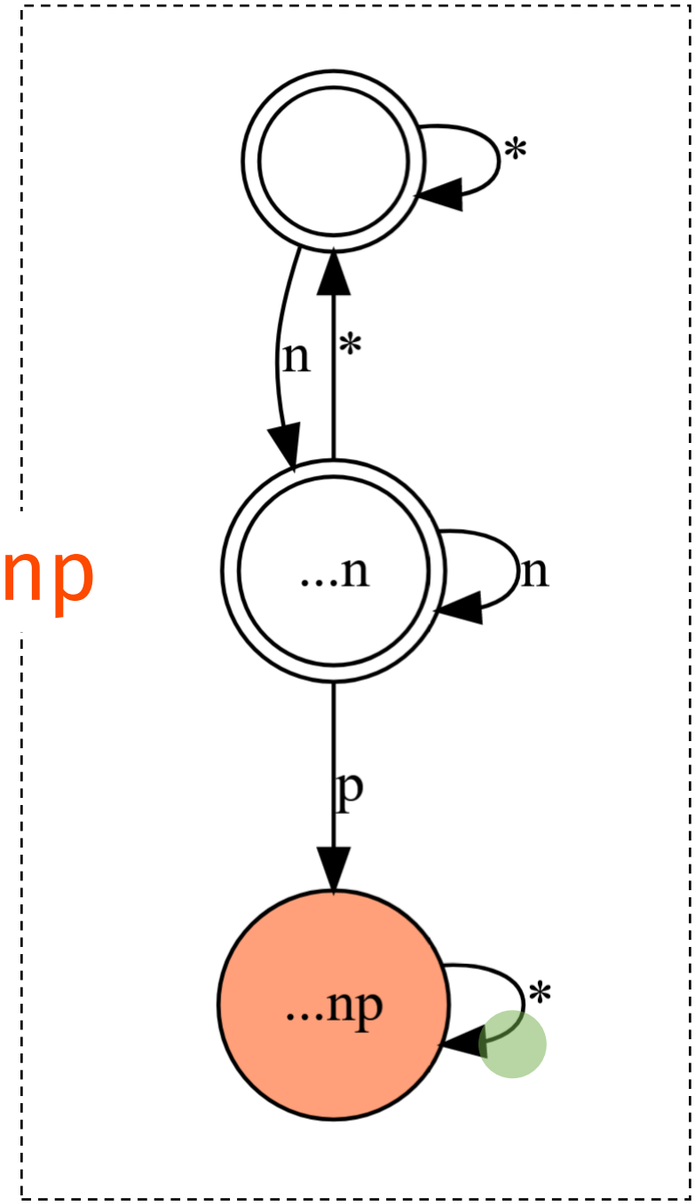
anpa

*np



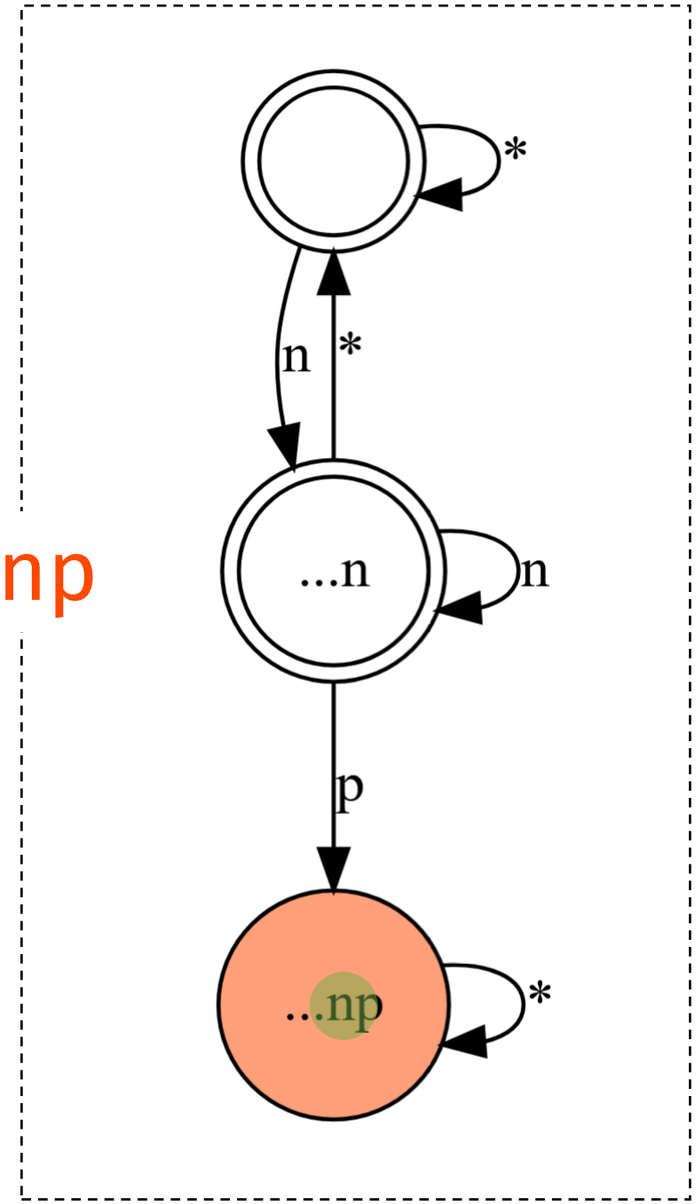
anpa

*np



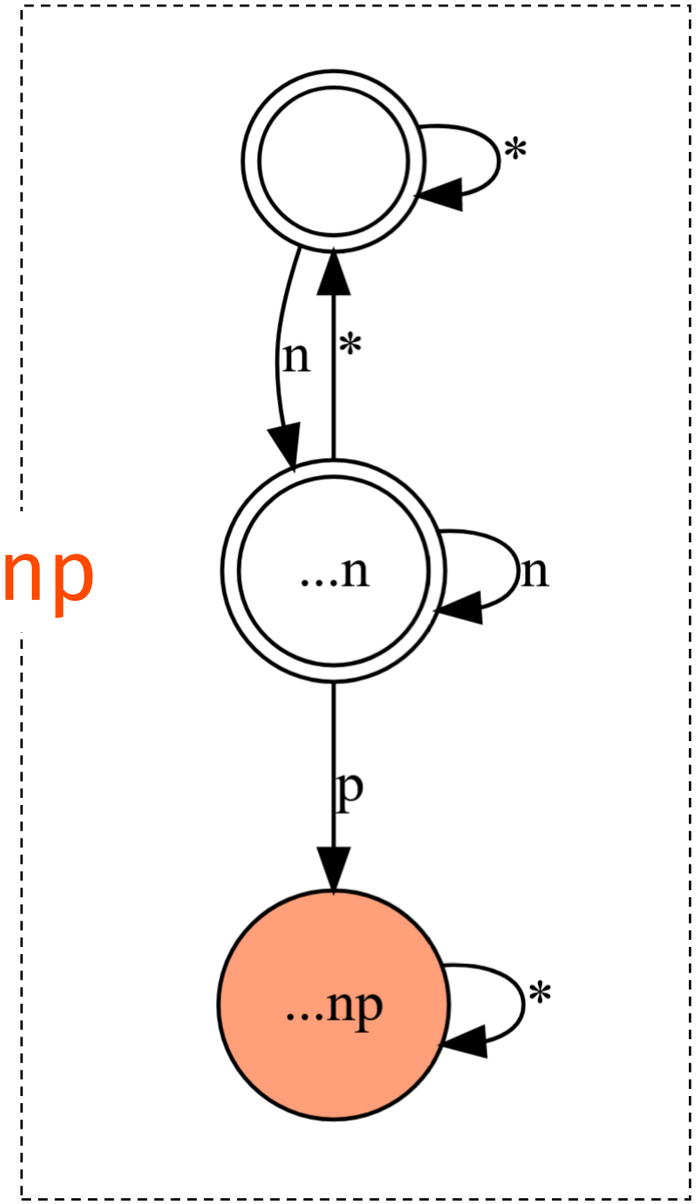
anpa

*np

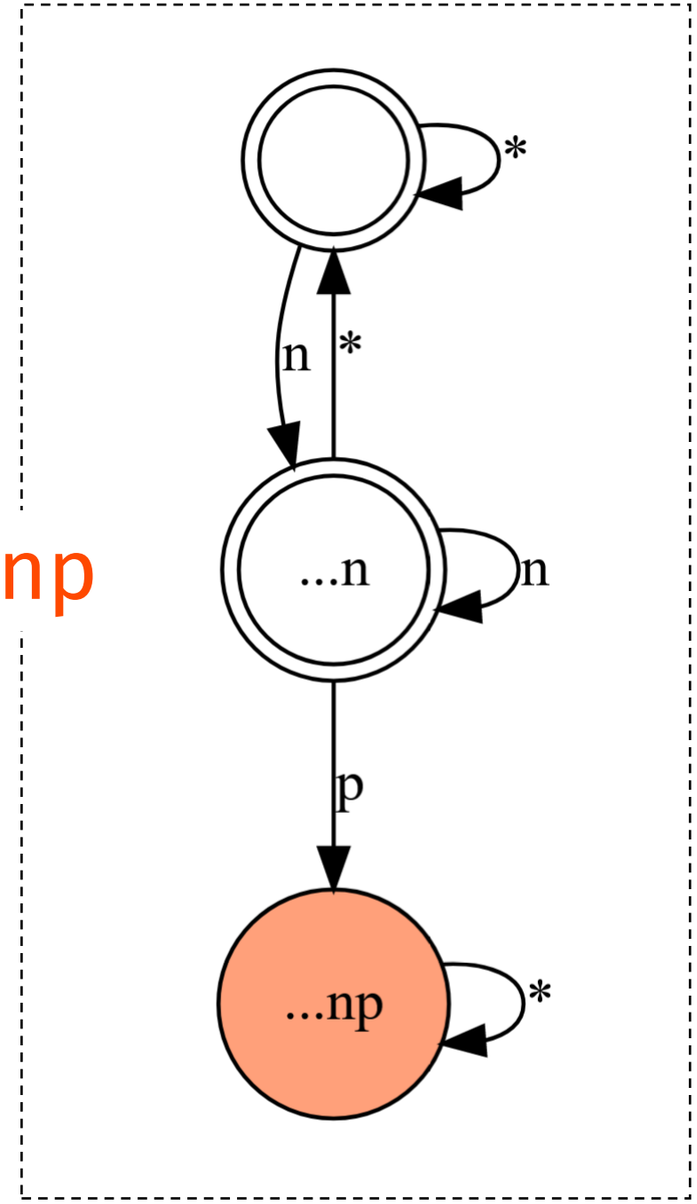


anpa

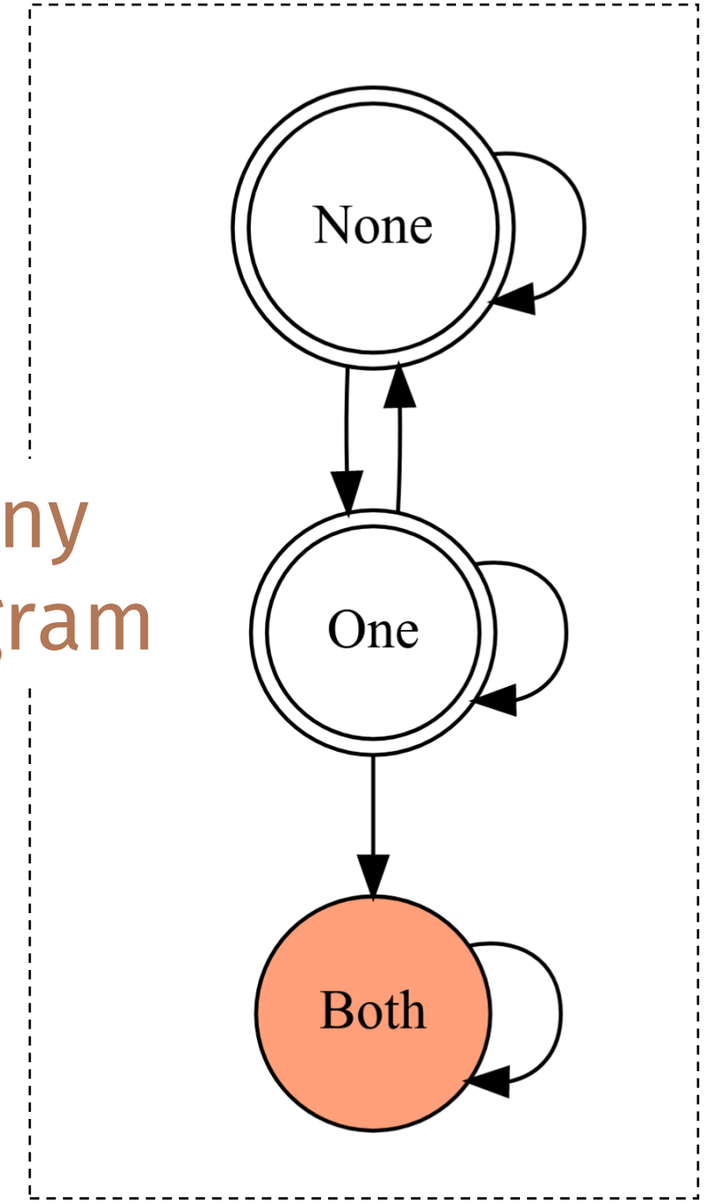
*np



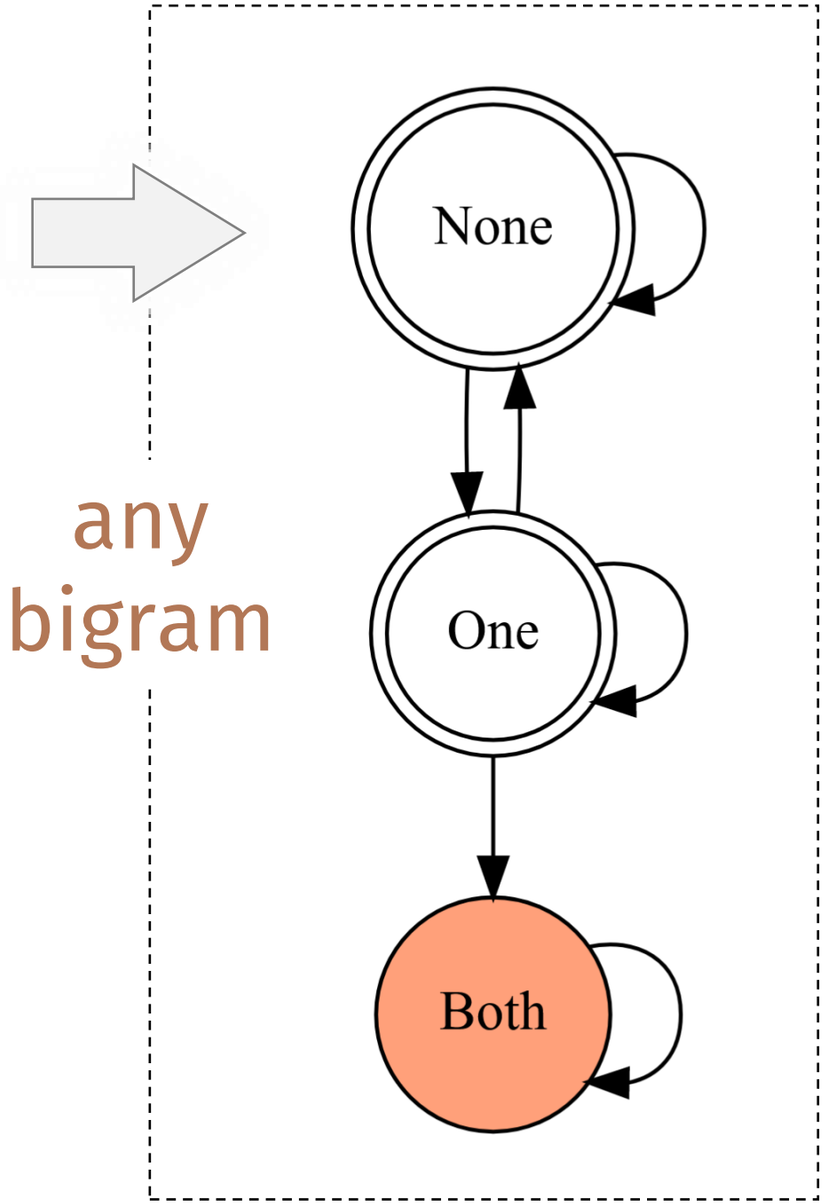
*np



any bigram



We prove the recombinant property for this topology of automata in general.



Initial and final constraints

- Let us call XY, YZ, and XZ the (recombinant) inputs, and XYZ the output
- The output XYZ **begins** with the same bigram, as the input **XY** already did
- The output XYZ **ends** with the same bigram, as the input **YZ** already did

Outline

1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

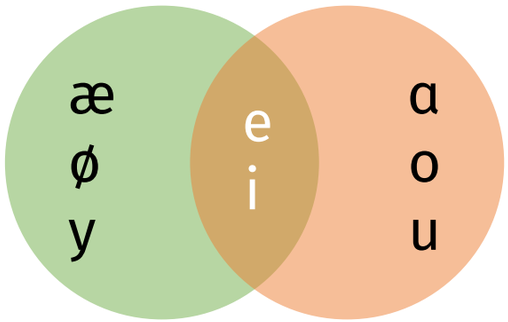
Outline

1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

Finnish vowel harmony

Some numerals:

kɑksi	‘2’	kɑhdeksɑn	‘8’
yksi	‘1’	yhdeksæn	‘9’



Rule: do not mix green and orange vowels in one word

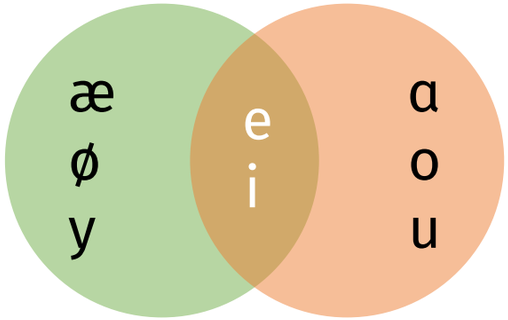
Finnish vowel harmony

Some numerals:

kɑksi	‘2’	kɑhdeksɑn	‘8’
yksi	‘1’	yhdeksæn	‘9’

kæksi ‘?’
uksi ‘?’

pseudo-words



Rule: do not mix green and orange vowels in one word

Finnish vowel harmony

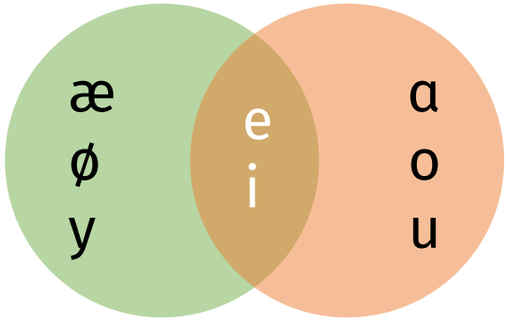
Some numerals:

kɑksi	‘2’	kɑhdeksɑn	‘8’
yksi	‘1’	yhdeksæn	‘9’

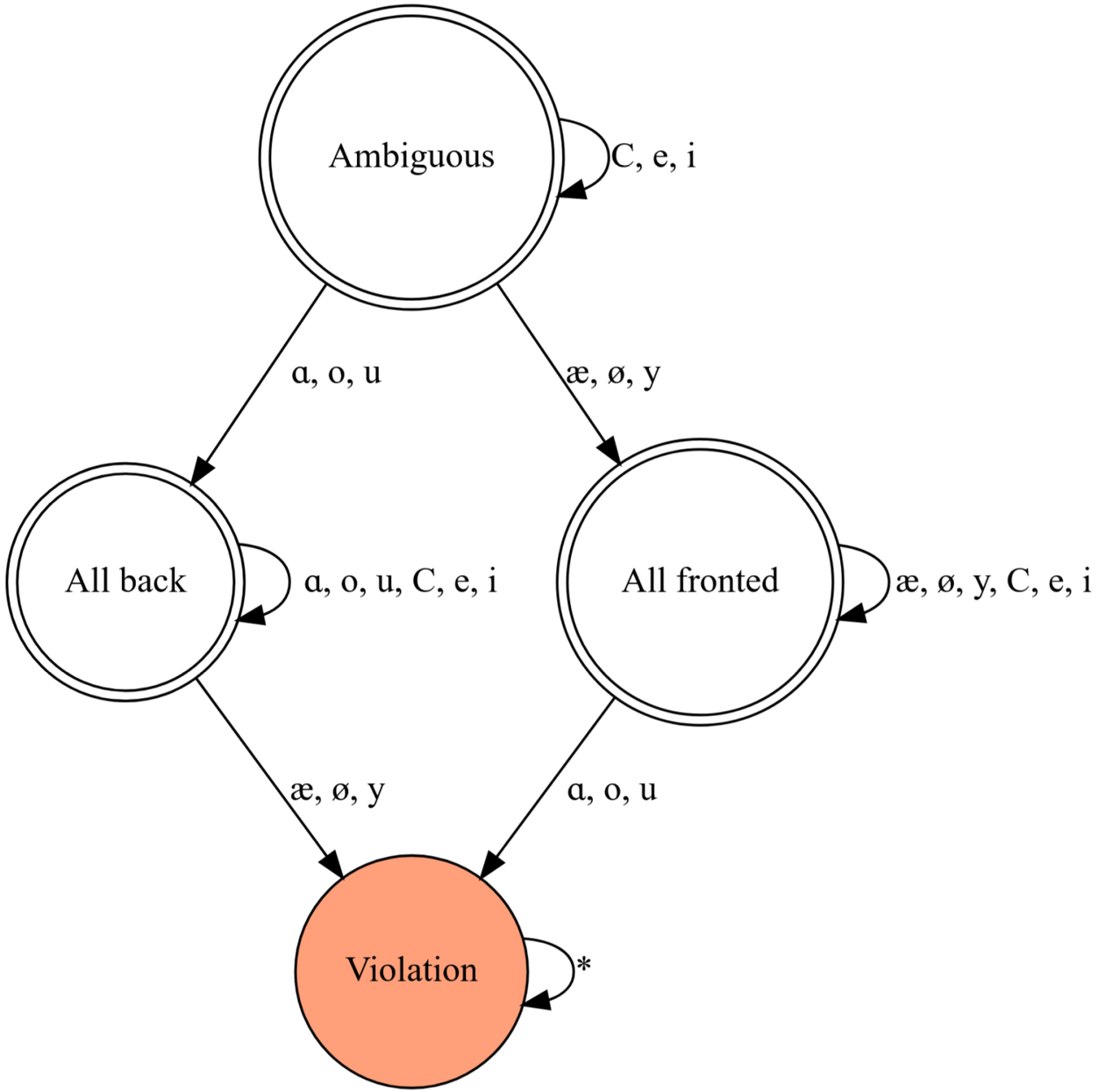
pseudo-words

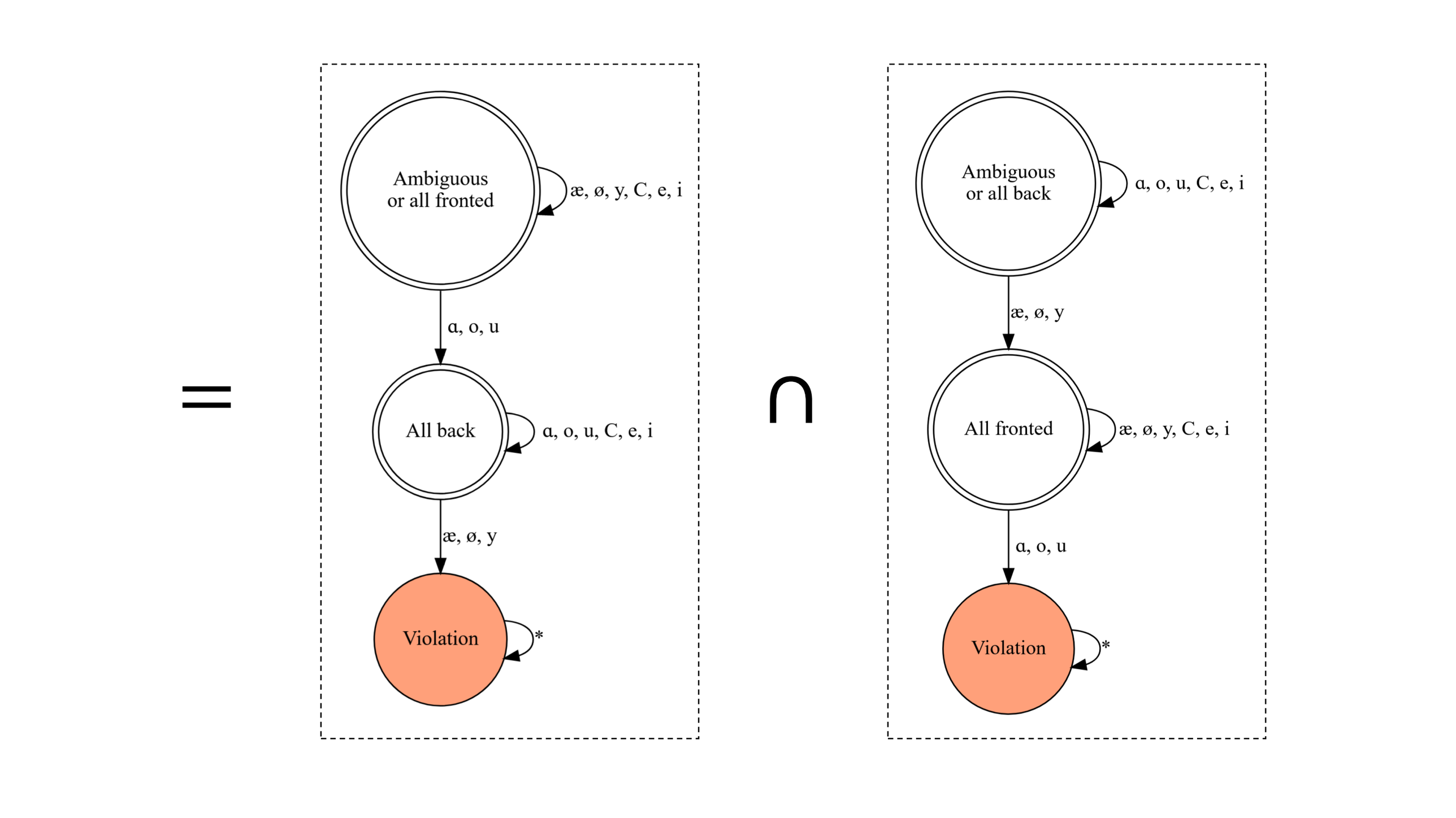
kæksi	‘?’	*kæhdeksɑn	–
uksi	‘?’	*uhdeksæn	–

ill-formed



Rule: do not mix green and orange vowels in one word


Ambiguous
C, e, i
ɑ, o, u
æ, ø, y
All back
ɑ, o, u, C, e, i
All fronted
æ, ø, y, C, e, i
æ, ø, y
ɑ, o, u
Violation
*


=
Ambiguous
or all fronted
æ, ø, y, C, e, i
ɑ, o, u
All back
ɑ, o, u, C, e, i
æ, ø, y
Violation
*
∩
Ambiguous
or all back
ɑ, o, u, C, e, i
æ, ø, y
All fronted
æ, ø, y, C, e, i
ɑ, o, u
Violation
*

The ternary recombinant property is closed under intersection.

Outline

1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

Outline

1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

Three harmonies in Imdlawn Tashlhiyt

s:-uga	‘CAUS-evacuate’
s-as:twa	‘CAUS-settle’
ʃ-fiaʃr	‘CAUS-be.full.of.straw’
z-bruz:a	‘CAUS-crumble’
ʒ-m:ʒdawl	‘CAUS-stumble’
s-ħuz	‘CAUS-annex’
s:-ukz	‘CAUS-recognize’
sˤ-rˤuˤfˤzˤ	‘CAUS-appear.resistant’
s-mχazaj	‘CAUS-loathe.each.other’
ʃ-quʒ:i	‘CAUS-be.dislocated’

(Elmedlaoui 1992; Hansson 2010),
reproduced here from (Aksënova et al. 2020)

Three harmonies in Imdlawn Tashlhiyt

s:-uga	‘CAUS-evacuate’
s-as:twa	‘CAUS-settle’
ʃ-fiaʃr	‘CAUS-be.full.of.straw’
z-bruz:a	‘CAUS-crumble’
ʒ-m:ʒdawl	‘CAUS-stumble’
s-ħuz	‘CAUS-annex’
s:-ukz	‘CAUS-recognize’
sˁ-rˁuˁfˁzˁ	‘CAUS-appear.resistant’
s-mχazaj	‘CAUS-loathe.each.other’
ʃ-quʒ:i	‘CAUS-be.dislocated’

(Elmedlaoui 1992; Hansson 2010),
reproduced here from (Aksënova et al. 2020)

Three harmonies in Imdlawn Tashlhiyt

s:-uga	‘CAUS-evacuate’
s-as:twa	‘CAUS-settle’
ʃ-fiaʃr	‘CAUS-be.full.of.straw’
z-bruz:a	‘CAUS-crumble’
ʒ-m:ʒdawl	‘CAUS-stumble’
s-ħuz	‘CAUS-annex’
s:-ukz	‘CAUS-recognize’
sˤ-rˤuˤfˤzˤ	‘CAUS-appear.resistant’
s-mχazaj	‘CAUS-loathe.each.other’
ʃ-quʒ:i	‘CAUS-be.dislocated’

emphasis (pharyngealization) harmony

(Elmedlaoui 1992; Hansson 2010), reproduced here from (Aksënova et al. 2020)

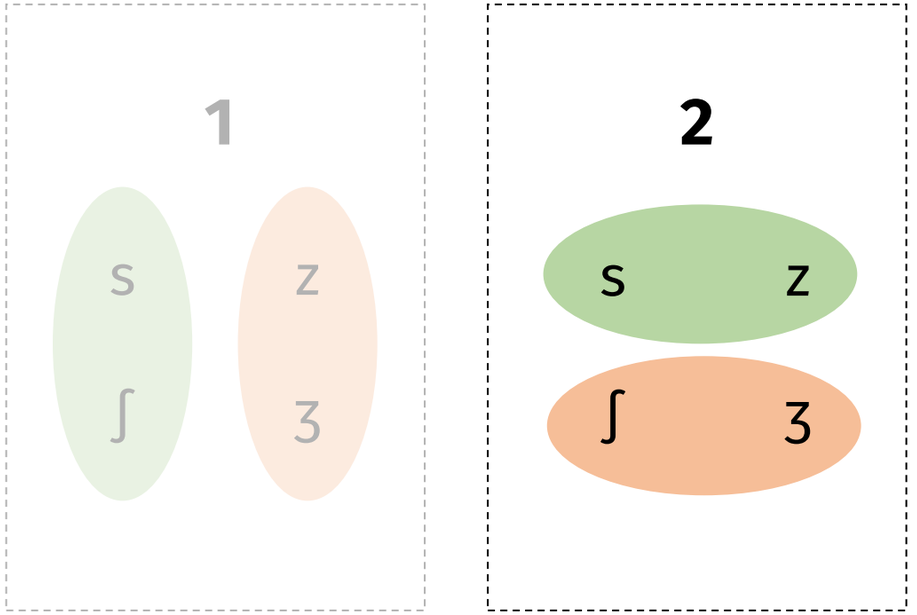
Two harmonies in Imdlawn Tashlhiyt

s:-uga	‘CAUS-evacuate’
s-as:twa	‘CAUS-settle’
ʃ-fiaʃr	‘CAUS-be.full.of.straw’
z-bruz:a	‘CAUS-crumble’
ʒ-m:ʒdawl	‘CAUS-stumble’
s-ħuz	‘CAUS-annex’
s:-ukz	‘CAUS-recognize’
s -r u f z	‘CAUS-appear.resistant’
s-mχazaj	‘CAUS-loathe.each.other’
ʃ-quʒ:i	‘CAUS-be.dislocated’

(Elmedlaoui 1992; Hansson 2010),
reproduced here from (Aksënova et al. 2020)

Two harmonies in Imdlawn Tashlhiyt

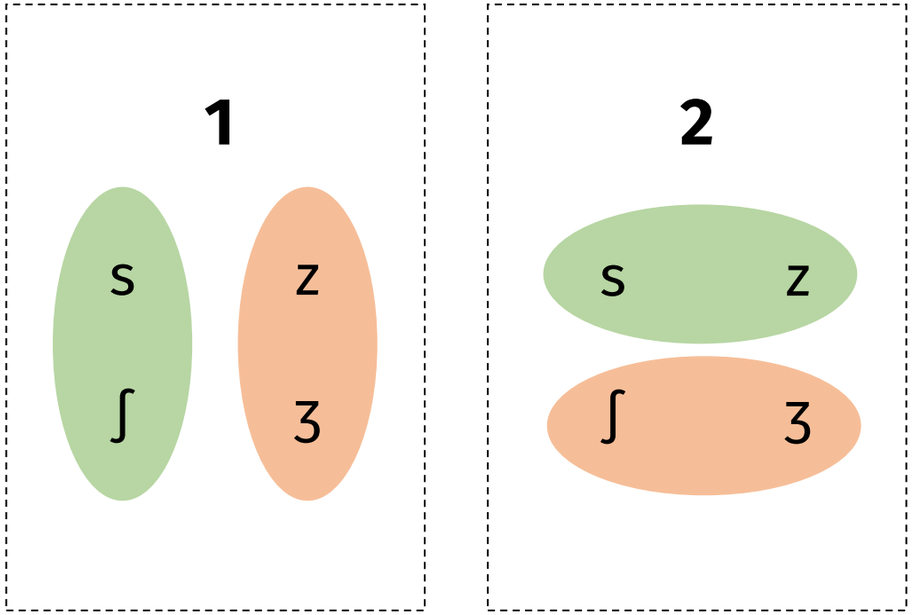
s:-uga	‘CAUS-evacuate’
s-as:twa	‘CAUS-settle’
ʃ-fiaʃr	‘CAUS-be.full.of.straw’
z-bruz:a	‘CAUS-crumble’
ʒ-m:ʒdawl	‘CAUS-stumble’
s-ħuz	‘CAUS-annex’
s:-ukz	‘CAUS-recognize’
s -r u f z	‘CAUS-appear.resistant’
s-mχazaj	‘CAUS-loathe.each.other’
ʃ-quʒ:i	‘CAUS-be.dislocated’



(Elmedlaoui 1992; Hansson 2010), reproduced here from (Aksënova et al. 2020)

Two harmonies in Imdlawn Tashlhiyt

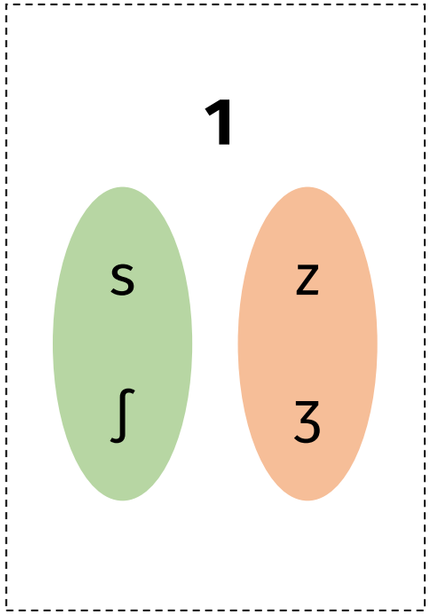
s:-uga	‘CAUS-evacuate’
s-as:twa	‘CAUS-settle’
ʃ-fiaʃr	‘CAUS-be.full.of.straw’
z-bruz:a	‘CAUS-crumble’
ʒ-m:ʒdawl	‘CAUS-stumble’
s-ħuz	‘CAUS-annex’
s:-ukz	‘CAUS-recognize’
s -r u f z	‘CAUS-appear.resistant’
s-mχazaj	‘CAUS-loathe.each.other’
ʃ-quʒ:i	‘CAUS-be.dislocated’



(Elmedlaoui 1992; Hansson 2010),
reproduced here from (Aksënova et al. 2020)

One harmony in Imdlawn Tashlhiyt

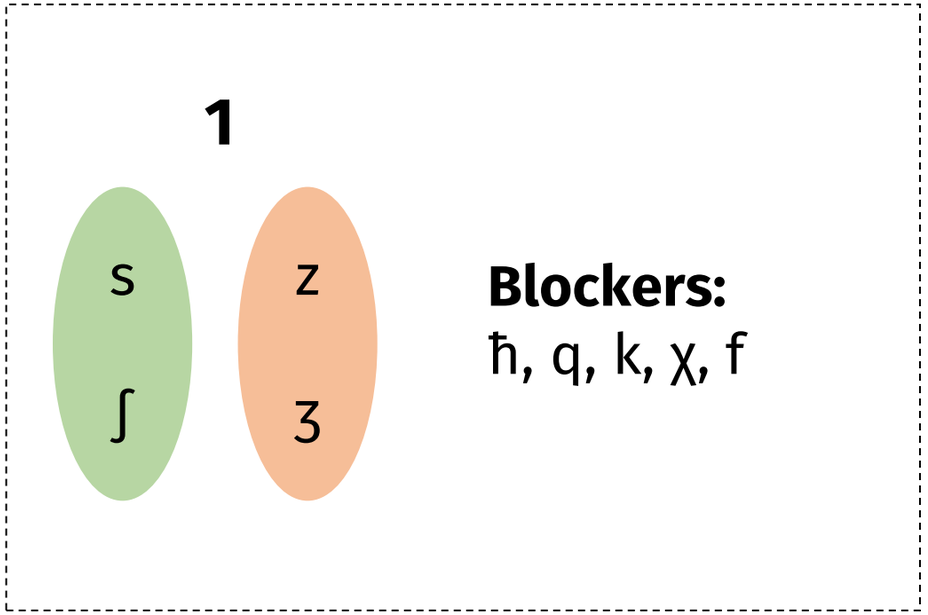
s:-uga	‘CAUS-evacuate’
s-as:twa	‘CAUS-settle’
ʃ-fiaʃr	‘CAUS-be.full.of.straw’
z-bruz:a	‘CAUS-crumble’
ʒ-m:ʒdawl	‘CAUS-stumble’
s-ħuz	‘CAUS-annex’
s:-ukz	‘CAUS-recognize’
s -r u f z	‘CAUS-appear.resistant’
s-mχazaj	‘CAUS-loathe.each.other’
ʃ-quʒ:i	‘CAUS-be.dislocated’



(Elmedlaoui 1992; Hansson 2010),
reproduced here from (Aksënova et al. 2020)

One harmony in Imdlawn Tashlhiyt

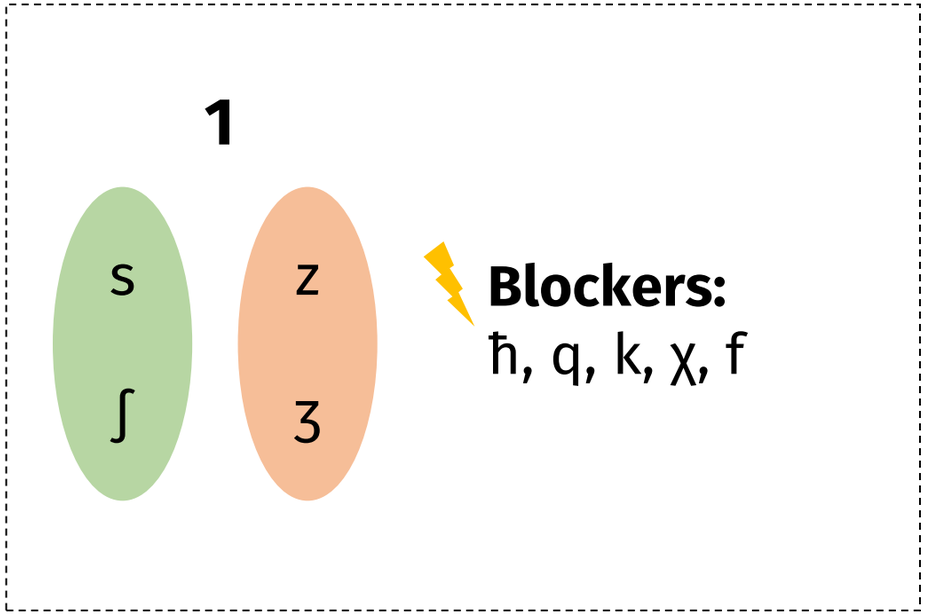
s:-uga	‘CAUS-evacuate’
s-as:twa	‘CAUS-settle’
ʃ-fiaʃr	‘CAUS-be.full.of.straw’
z-bruz:a	‘CAUS-crumble’
ʒ-m:ʒdawl	‘CAUS-stumble’
s-ħuz	‘CAUS-annex’
s:-ukz	‘CAUS-recognize’
s -r u f z	‘CAUS-appear.resistant’
s-mχazaj	‘CAUS-loathe.each.other’
ʃ-quʒ:i	‘CAUS-be.dislocated’



(Elmedlaoui 1992; Hansson 2010), reproduced here from (Aksënova et al. 2020)

One harmony in Imdlawn Tashlhiyt

s:-uga	‘CAUS-evacuate’
s-as:twa	‘CAUS-settle’
ʃ-fiaʃr	‘CAUS-be.full.of.straw’
z-bruz:a	‘CAUS-crumble’
ʒ-m:ʒdawl	‘CAUS-stumble’
s-ħuz	‘CAUS-annex’
s:-ukz	‘CAUS-recognize’
s -r u f z	‘CAUS-appear.resistant’
s-mχazaj	‘CAUS-loathe.each.other’
ʃ-quʒ:i	‘CAUS-be.dislocated’



(Elmedlaoui 1992; Hansson 2010), reproduced here from (Aksënova et al. 2020)

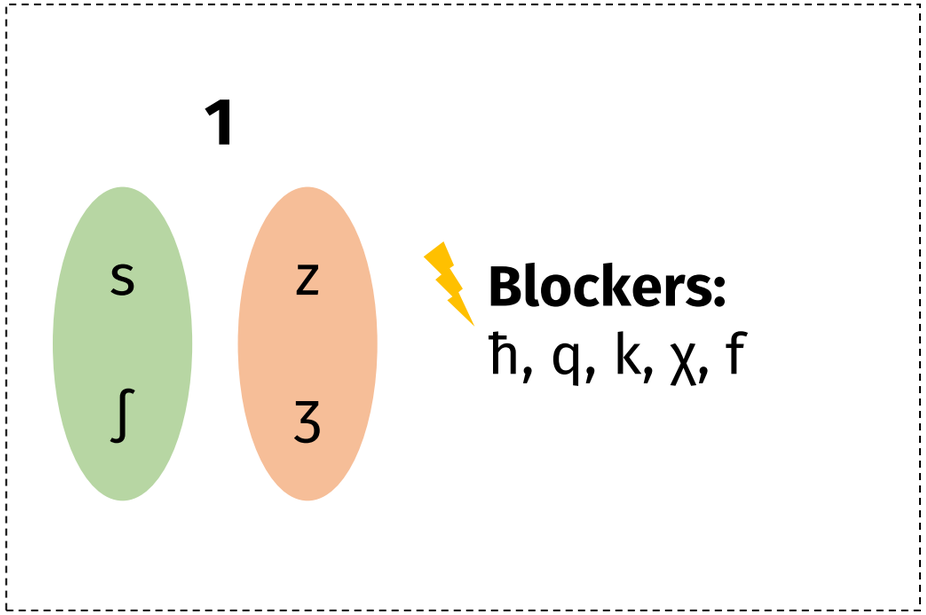
One harmony in Imdlawn Tashlhiyt

s a s: t w a

ʒ m: ʒ d a w l

s m χ a z a j

* s m a z a j



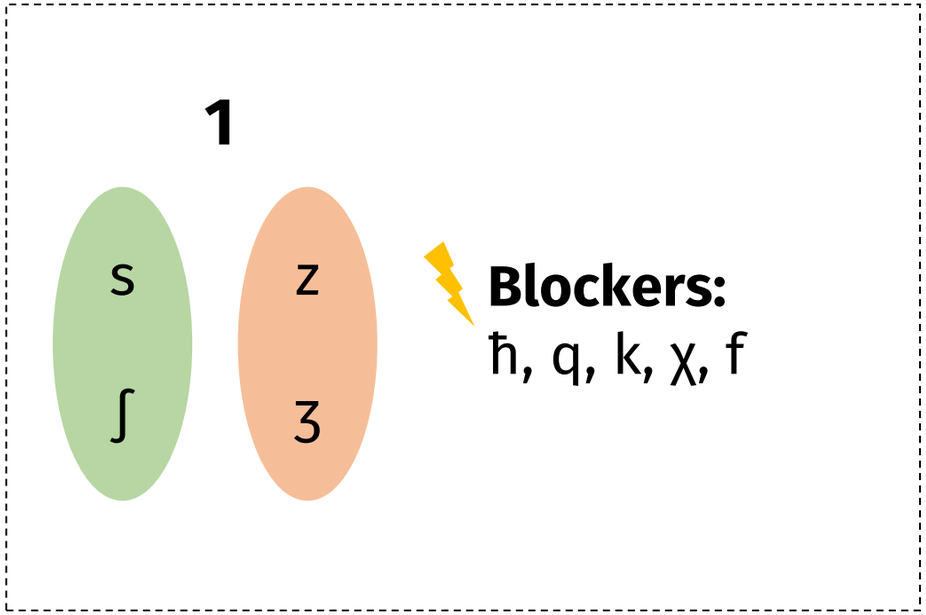
One harmony in Imdlawn Tashlhiyt

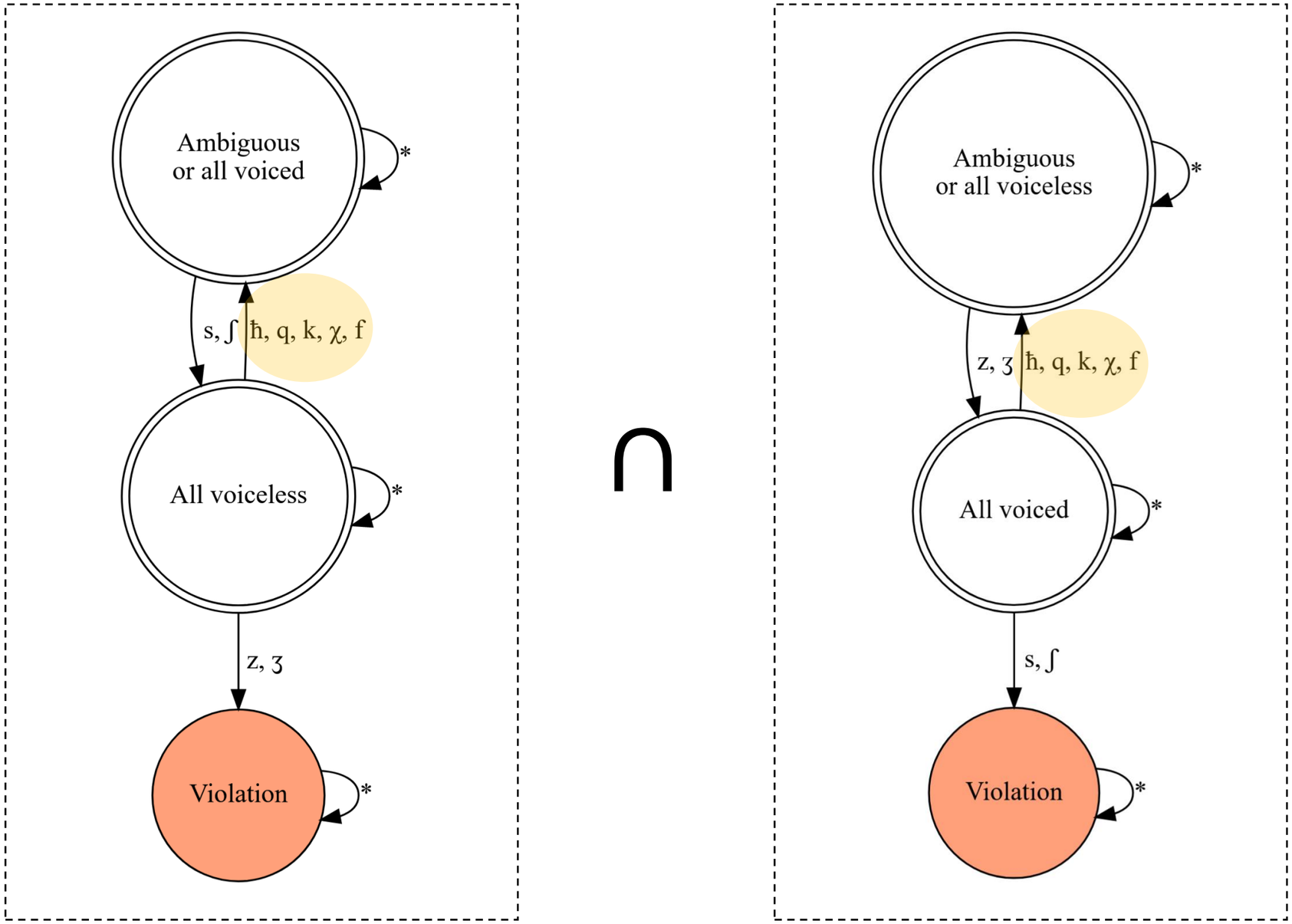
s a s: t w a

ʒ m: ʒ d a w l

s m χ a z a j

* s m a z a j




Ambiguous
or all voiced
*
s, ʃ
ħ, q, k, χ, f
All voiceless
*
z, ʒ
Violation
*
∩
Ambiguous
or all voiceless
*
z, ʒ
ħ, q, k, χ, f
All voiced
*
s, ʃ
Violation
*

?!

Phonotactics always constitutes a regular language
(Karttunen et al. 1987; Kaplan & Kay 1994).

Phonotactics always constitutes a regular language
(Karttunen et al. 1987; Kaplan & Kay 1994).

Moreover, Aksënova et al. (2020) have already shown that the **MTSL$_2$** (multiple tier-based strictly 2-local) subclass of regular languages suffices to account for a broad range of harmony phenomena.

The Computational Power of Harmony

Alëna Aksënova, Jonathan Rawski, Thomas Graf, Jeffrey Heinz

July 1, 2020

1 Computation and Phonology

Some of the most stunning scientific advances of the twentieth century came from the theory of computation. It is hard to think of an area in science which has not been enriched by studying a subject computationally. At the core of the theory of computation is the ability to formally define problems and processes, and ask how (or whether) they can be solved computationally, and if so, with which resources. For linguistics, the science of language, this means characterizing the cognitive processes involved in the human capacity for lan-

Phonotactics always constitutes a regular language
(Karttunen et al. 1987; Kaplan & Kay 1994).

Moreover, Aksënova et al. (2020) have already shown that the **MTSL$_2$** (multiple tier-based strictly 2-local) subclass of regular languages suffices to account for a broad range of harmony phenomena.

On the other hand,

Phonotactics always constitutes a regular language
(Karttunen et al. 1987; Kaplan & Kay 1994).

Moreover, Aksënova et al. (2020) have already shown that the **MTSL$_2$** (multiple tier-based strictly 2-local) subclass of regular languages suffices to account for a broad range of harmony phenomena.

On the other hand,

MTSL$_2$ =  .

Phonotactics always constitutes a regular language
(Karttunen et al. 1987; Kaplan & Kay 1994).

Moreover, Aksënova et al. (2020) have already shown that the **MTSL$_2$** (multiple tier-based strictly 2-local) subclass of regular languages suffices to account for a broad range of harmony phenomena.

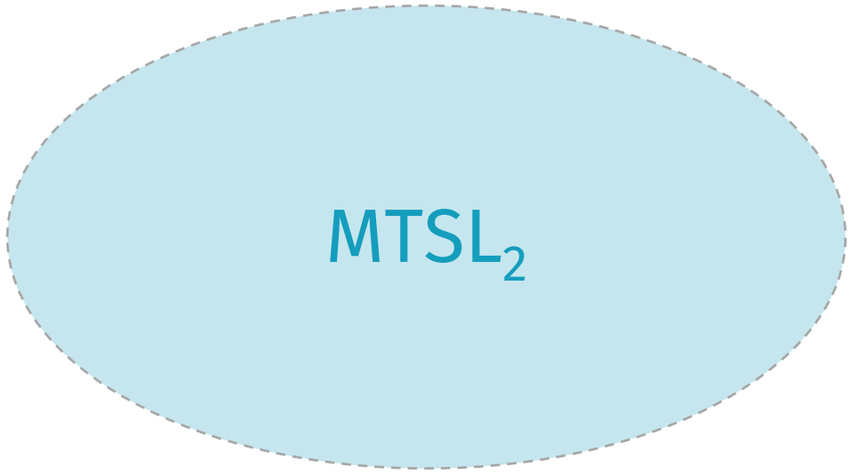
On the other hand,

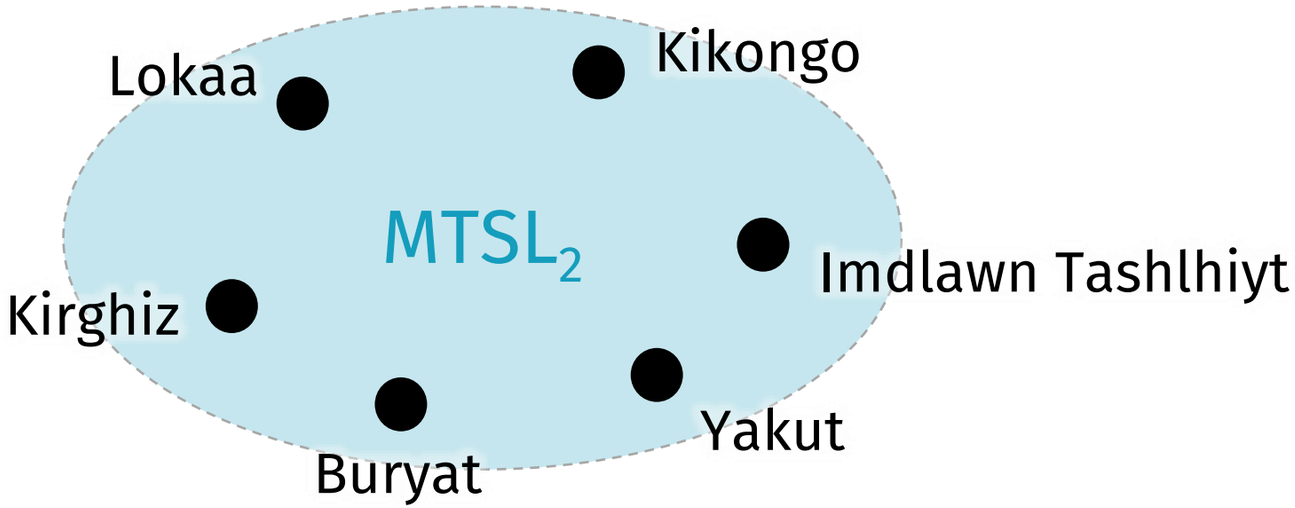
MTSL$_2$ =  .

Therefore, the entire MTSL$_2$ subclass has the ternary recombinant property.

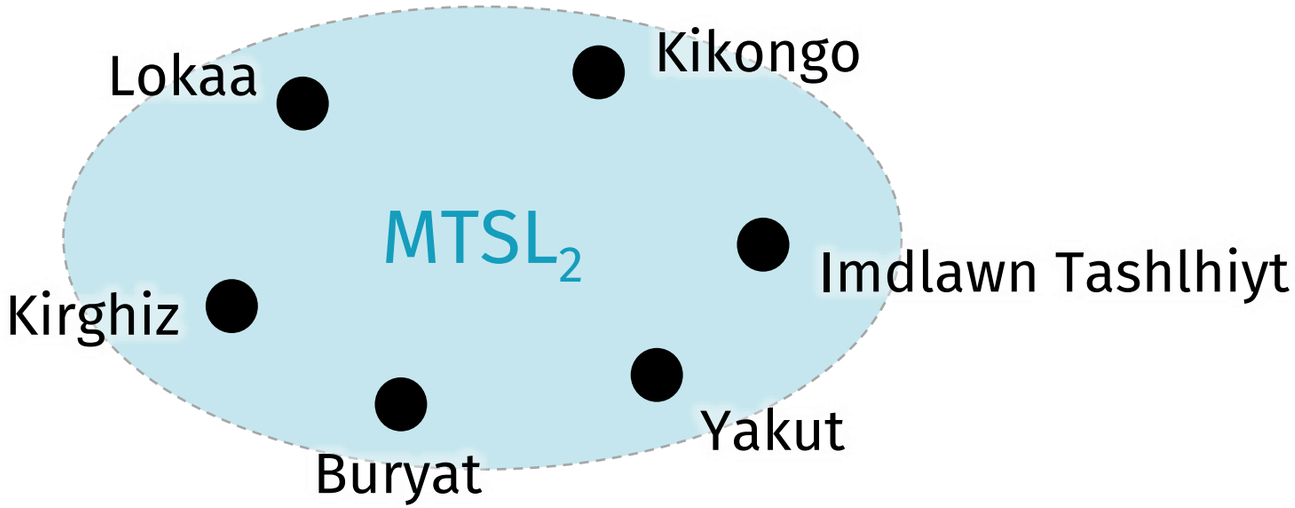
Outline

1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion


MTSL2


Lokaa
Kikongo
MTSL2
Imdlawn Tashlhiyt
Kirghiz
Yakut
Buryat

Yaka (Aksënova & De Santo 2021)



Tutrugbu (McCollum et al. 2020)

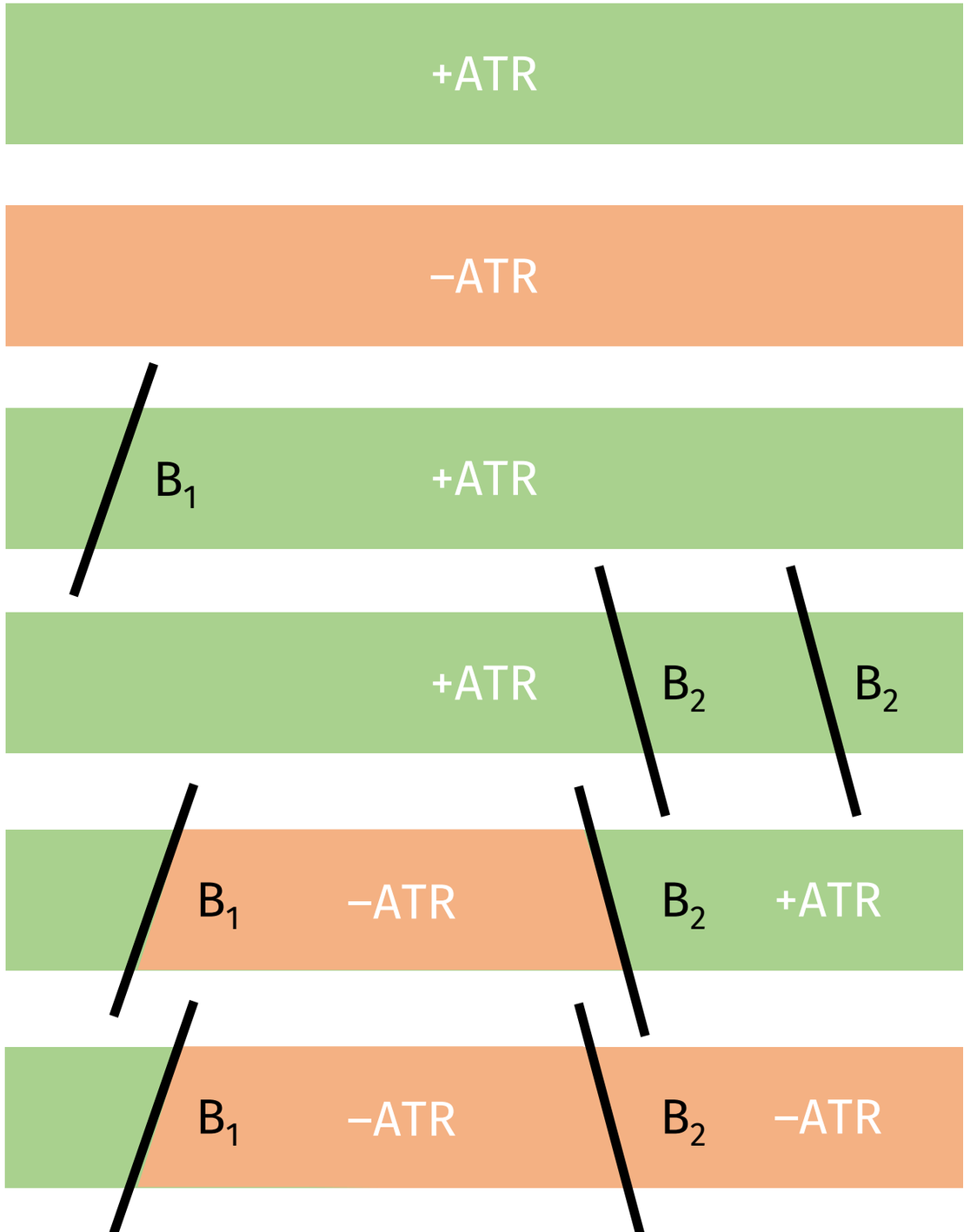
Tutrugbu (Atlantic-Congo)

ATR harmony, as per McCollum et al. (2020):

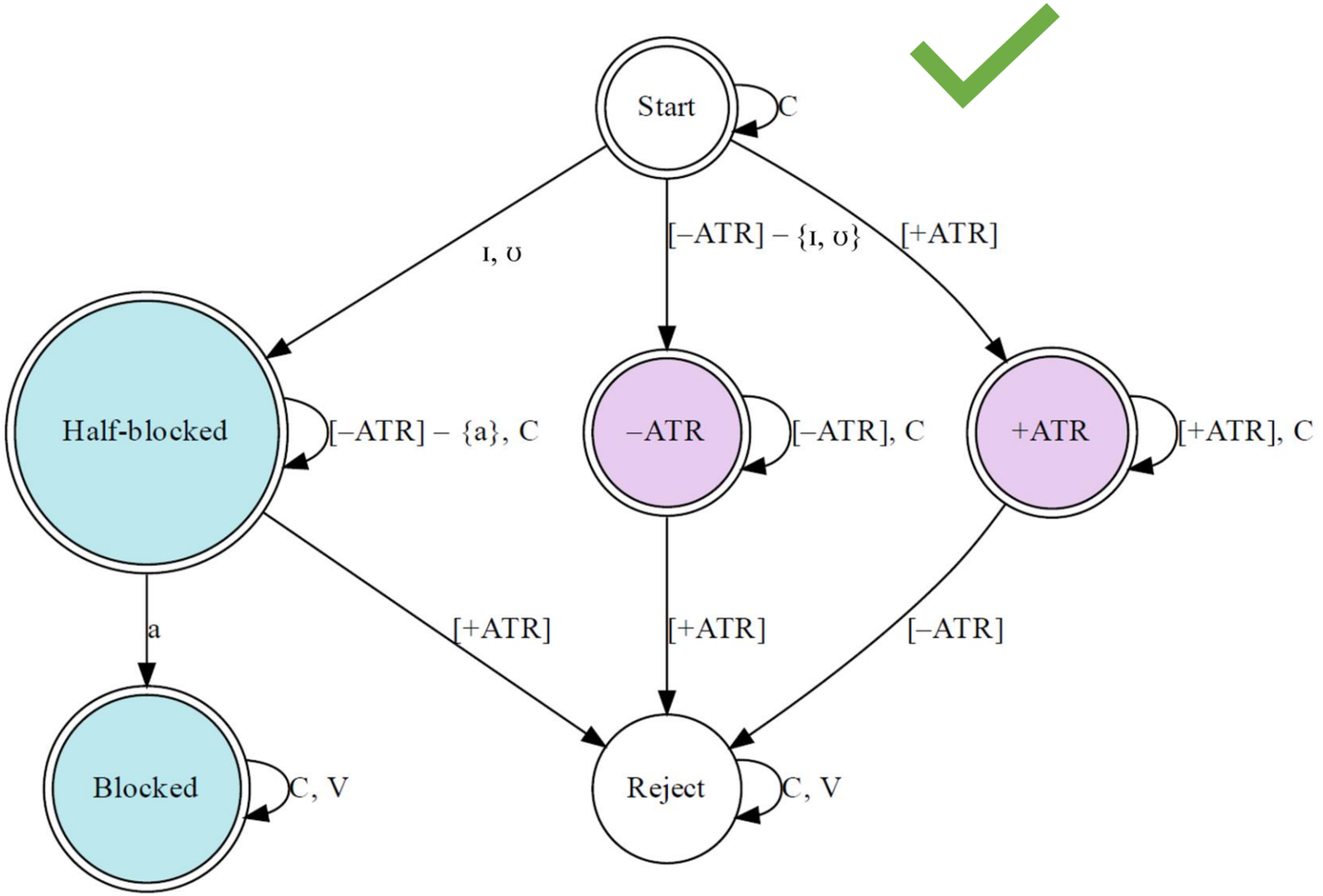
- [+ATR] = /i, e, u/
- [–ATR] = /ɪ, ɛ, a, ɔ, ʊ/

Blockers in ATR harmony:

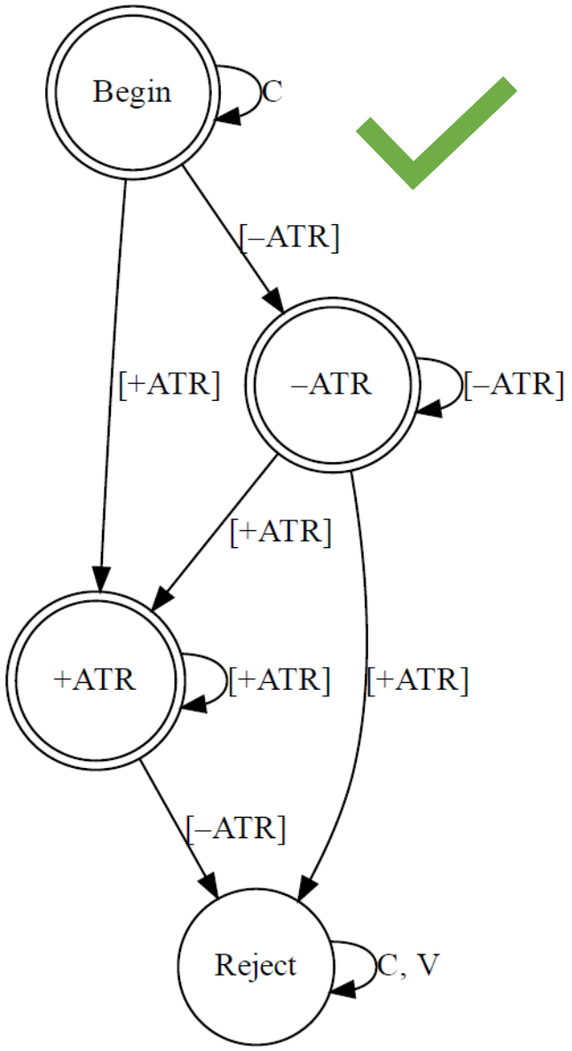
- First half: /ɪ, ʊ/ in word-initial syllable, that is, /#(C)ɪ/ or /#(C)ʊ/
- Second half: /a/


+ATR
–ATR
B1
+ATR
+ATR
B2
B2
B1
–ATR
B2
+ATR
B1
–ATR
B2
–ATR

	Transcription, harmonizing feature (±ATR), blockers (B_1, B_2)	**Gloss**
	ɪ d ɪ t ɔ – – – B_1	‘1S-ITIVE-cook’
	i d i w u + + +	‘1S-ITIVE-climb’
*	ɪ d ɪ w u – – + B_1	–
	ɪ b a w u – – + B_1 B_2	‘1S-FUT-climb’
*	ɔ b a w u – – + B_2	–
	ɪ t ɪ k a w u – – – + B_1 B_2	‘1S-NEG-PFV-climb’


Start
C
ɪ, ʊ
[–ATR] – {ɪ, ʊ}
[+ATR]
Half-blocked
[–ATR] – {a}, C
–ATR
[–ATR], C
+ATR
[+ATR], C
a
[+ATR]
[+ATR]
[–ATR]
Blocked
C, V
Reject
C, V

∩


Begin
C
[–ATR]
[+ATR]
–ATR
[–ATR]
[+ATR]
+ATR
[+ATR]
[+ATR]
[–ATR]
Reject
C, V

Outline

1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

*ksk in Finnish (Juho Pystynen, p. c.)

naks	onomatopoeia
skini	‘skinhead’
nakkini	‘frankfurter.1POSS’
***nakskini**	–

*ksk in Finnish (Juho Pystynen, p. c.)

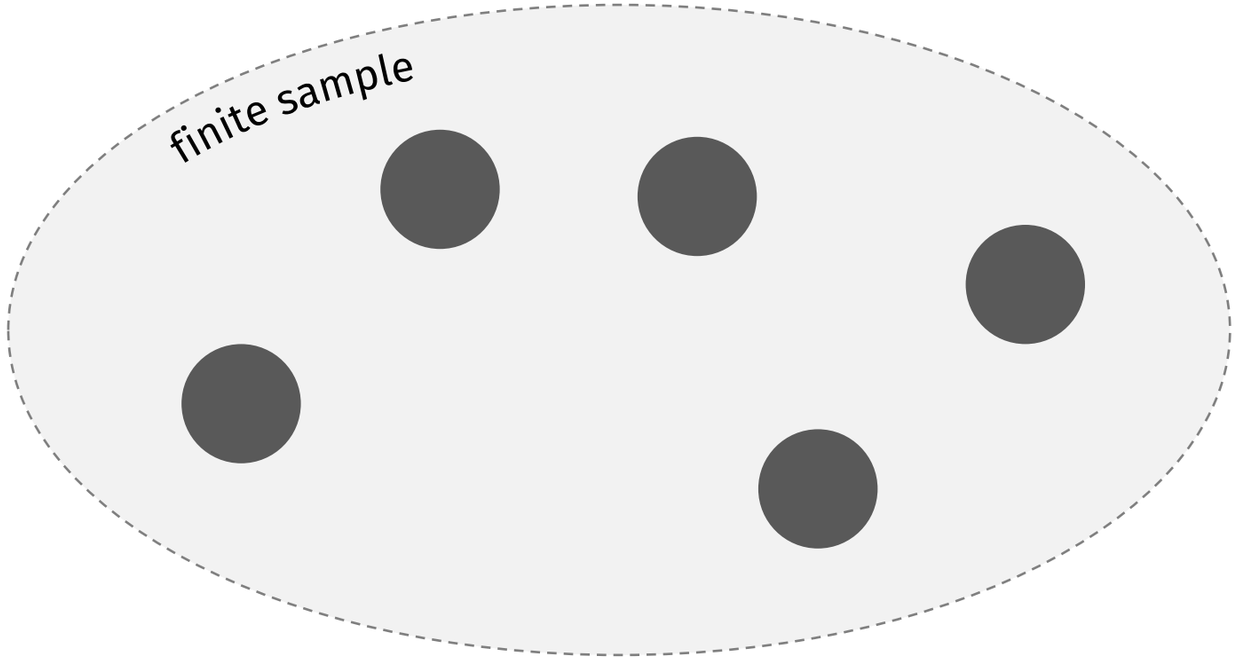
naks	onomatopoeia
skini	‘skinhead’
nakkini	‘frankfurter.1POSS’
***nakskini**	–

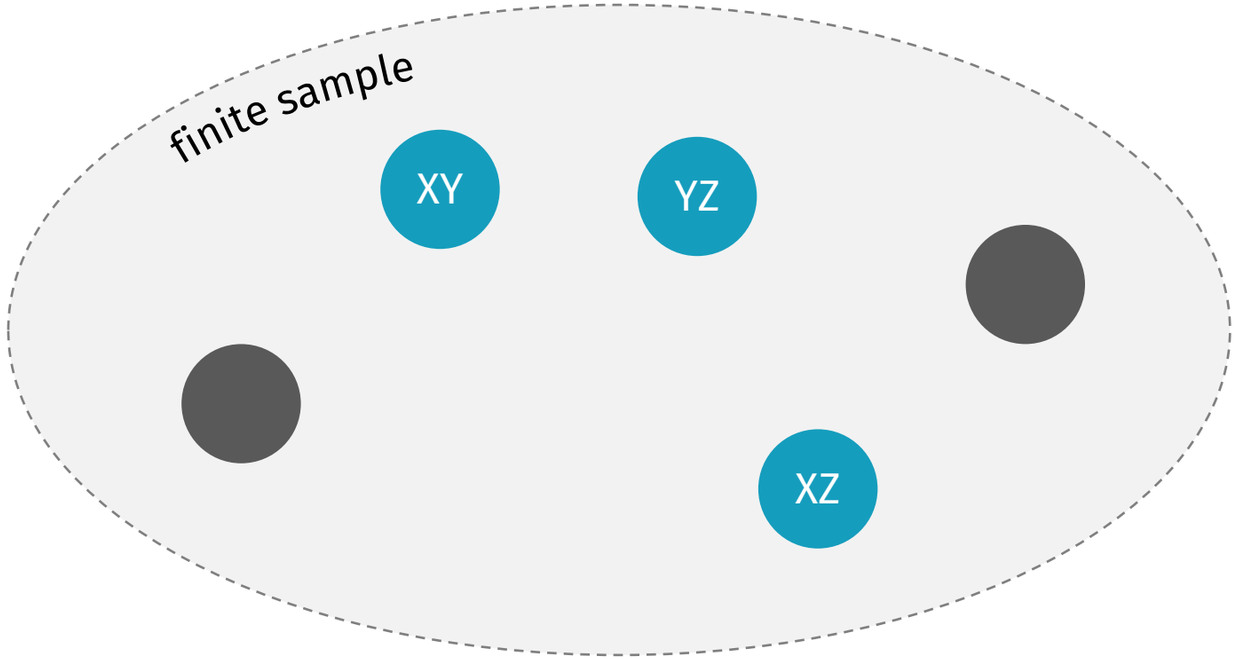
*krt in Russian (Vladimir Panov, p. c.)

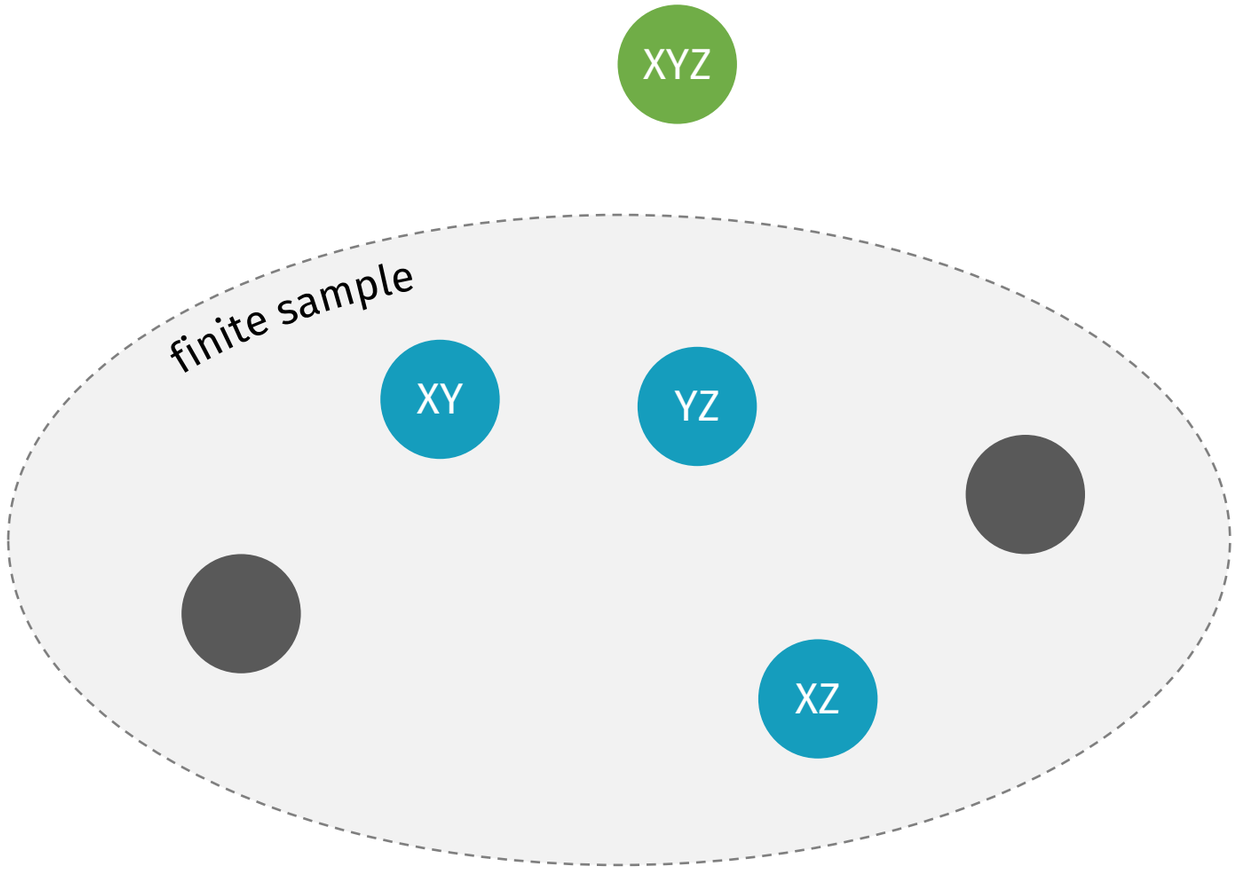
akr	‘acre’
rta	‘mouth.GEN’
akta	‘act.GEN’
***akrta**	–

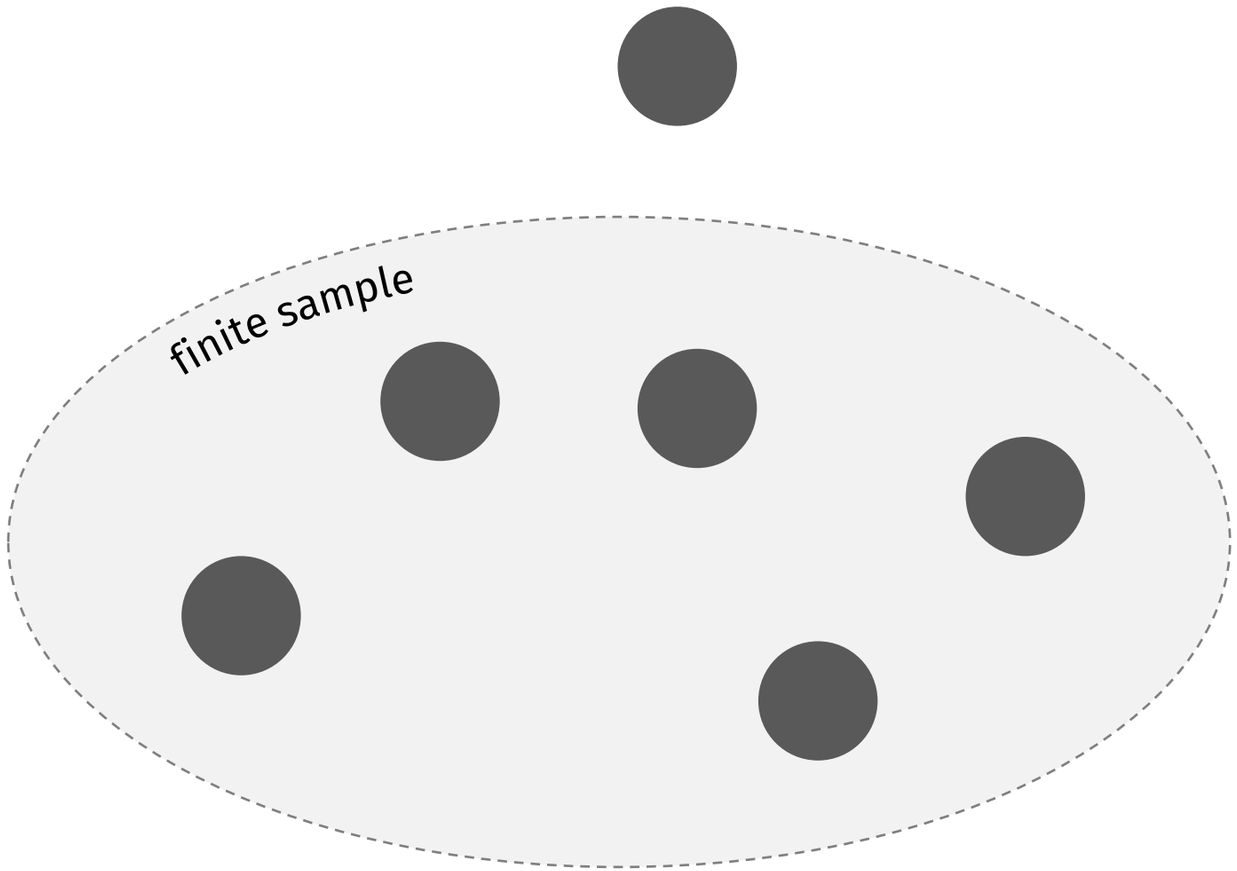
Outline

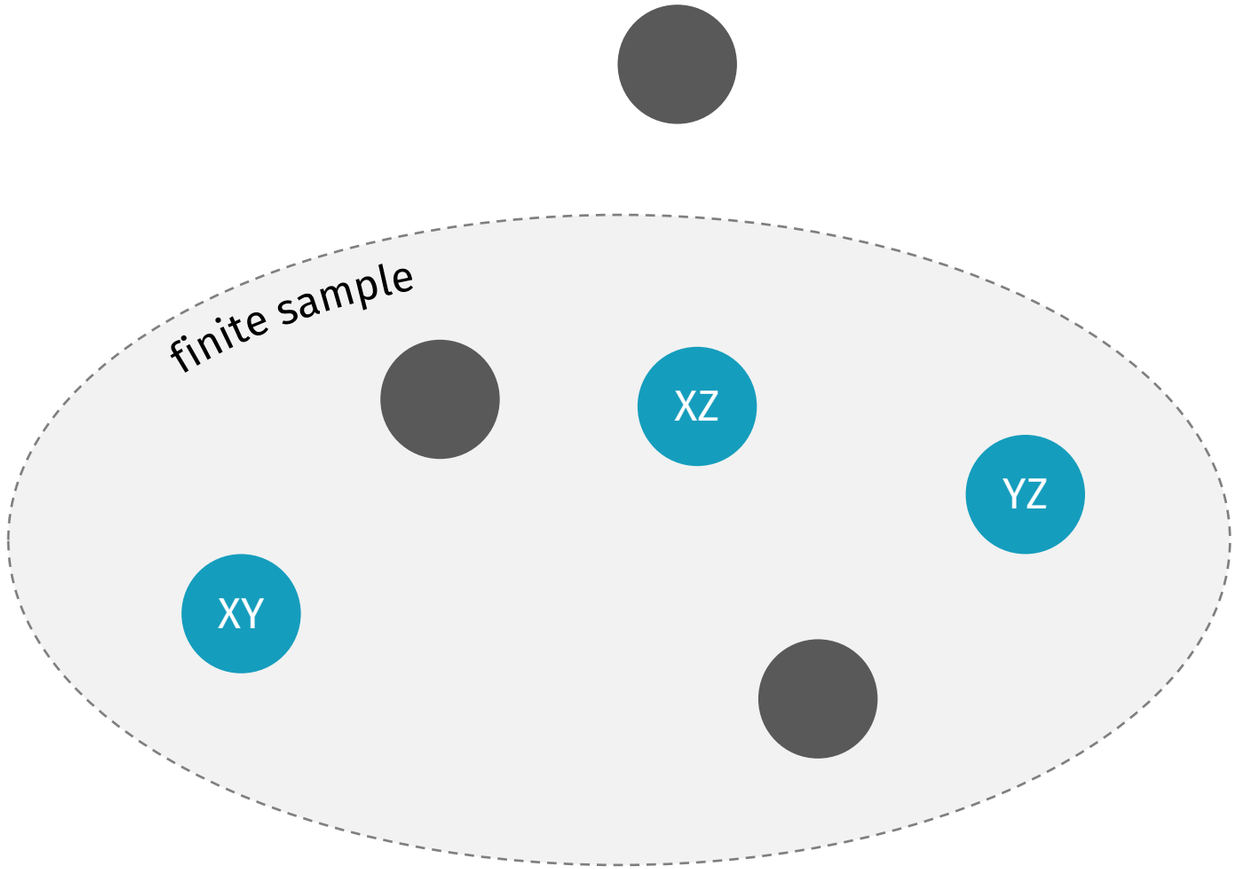
1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

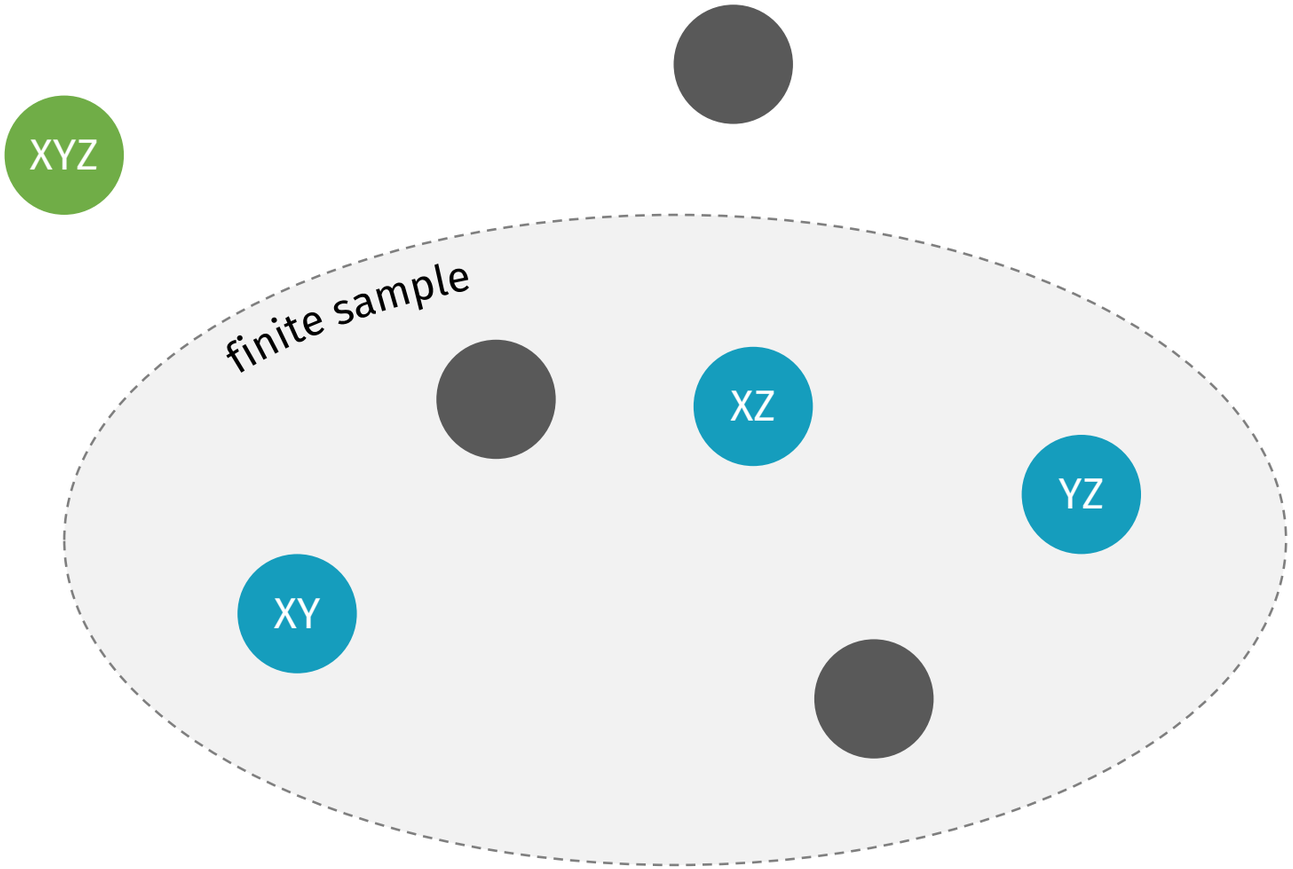

finite sample

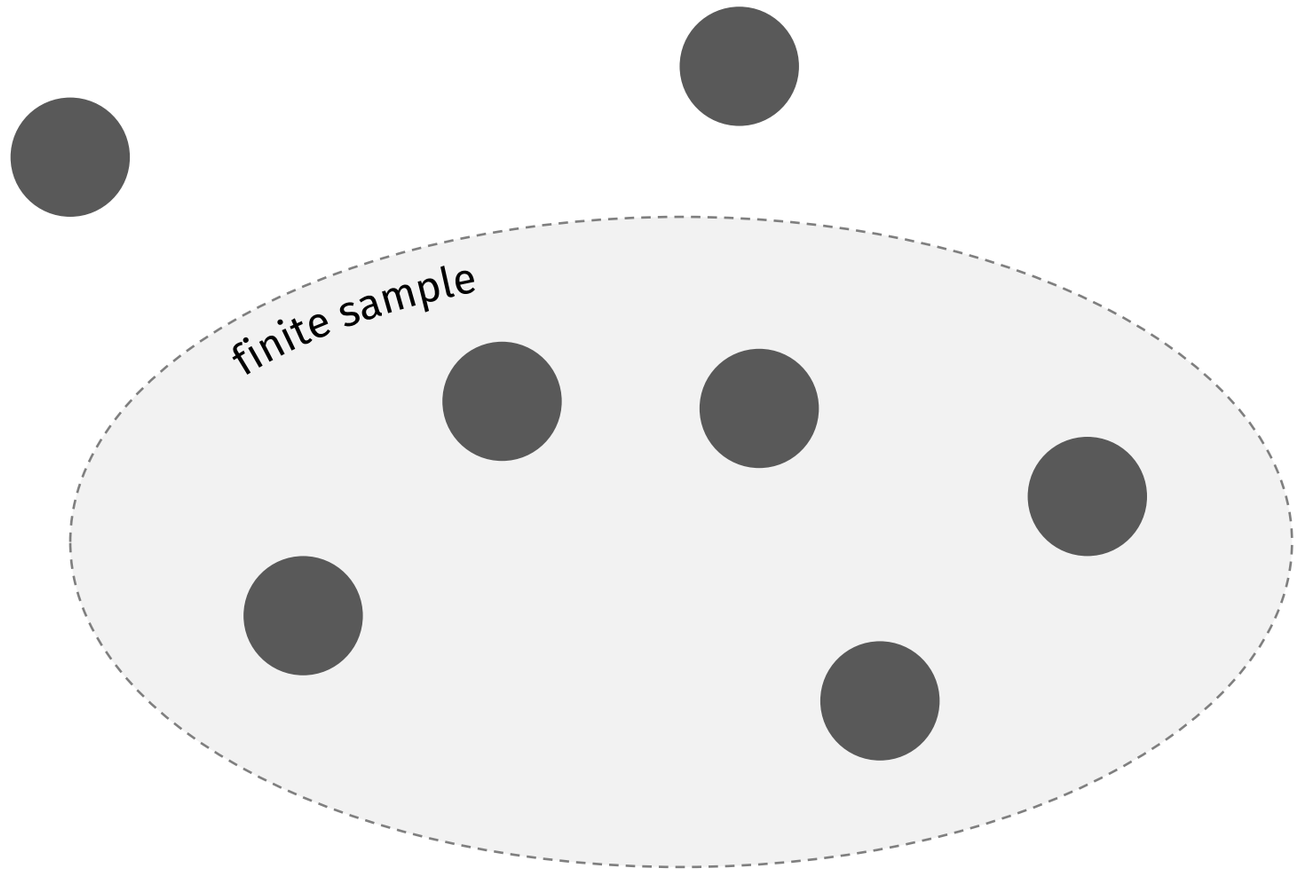

finite sample
XY
YZ
XZ

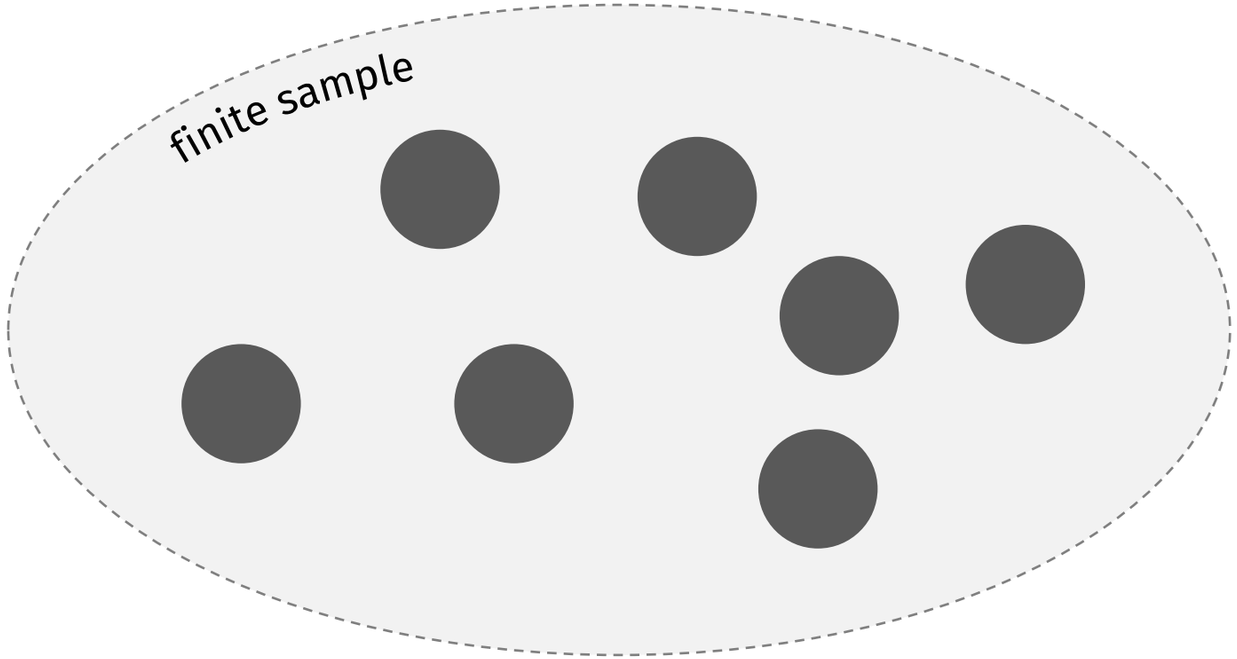

XYZ
finite sample
XY
YZ
XZ


finite sample


finite sample
XZ
YZ
XY


XYZ
finite sample
XZ
YZ
XY


finite sample


finite sample

iterate

Outline

1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

Outline

1. Local, initial, final constraints
2. Finnish vowel harmony (“classical” harmony)
3. Imdlawn Tashlhiyt consonant harmony (with simple blockers)
4. Tutrugbu vowel harmony (with distal bipartite blockers)
5. Trigrams
6. Induction procedure based on the implication
7. Conclusion

Conclusion

- So far, everything is accounted for, except trigram constraints
  - Although requiring |Y| ≥ 2 keeps trigrams at bay
- The shape of the recombinant property is very simple
  - Subjectively, simpler to describe than MTSL_2
  - Or any particular harmony system, really
- A trivial learning algorithm follows directly
  - Not very eager to generalize
  - Exemplar-driven by design
- Future test cases: Sanskrit *n*-retroflexion (Ryan 2017), limited-distance harmony, phonologization of harmony (?)
- Wordhood is an issue, and so is representing suprasegmentals

References

Aksënova, A., & De Santo, A. (2021). Learning Interactions of Local and Non-Local Phonotactic Constraints from Positive Input.

Aksënova, A., Rawski, J., Graf, T., & Heinz, J. (2020). The Computational Power of Harmony.

Elmedlaoui, M. (1992). Aspects des représentations phonologiques dans certaines langues chamito-sémitiques. *Doctorat d'état, Université Mohamed V, Rabat*.

Hansson, G. Ó. (2010). *Consonant harmony: Long-distance interactions in phonology* (Vol. 145). Univ of California Press.

Kaplan, R. M., & Kay, M. (1994). Regular models of phonological rule systems. *Computational linguistics*, *20*(3), 331-378.

Karttunen, L., Koskenniemi, K., & Kaplan, R. (1987). A compiler for two-level phonological rules. *Tools for morphological analysis*.

McCollum, A. G., Baković, E., Mai, A., & Meinhardt, E. (2020). Unbounded circumambient patterns in segmental phonology. *Phonology*, *37*(2), 215-255.

Ryan, K. (2017). Attenuated spreading in Sanskrit retroflex harmony. *Linguistic Inquiry*, *48*(2), 299-340.

Thank you!