Computation, learning, and typology

Class 4: Converging grammars and typological frequency

Adam Albright and Eric Baković

CreteLing 2023 — July 2023

creteling2023.phonology.party

Converging grammars

Predicting typological frequency

- Over the last few sessions, we have introduced some intuitive ideas about how formal properties of grammars may correlate with typological frequency
  - Simplicity: scored more highly by evaluation metric = preferred by learners $\Rightarrow$ more common typologically?
  - Learning trajectory: fewer learning "updates" $\Rightarrow$ quicker to learn $\Rightarrow$ more common typologically?
- Another question that has arisen: if more grammars can generate a pattern, is it more frequent?

R-volume

The R-volume hypothesis Bane & Riggle (2008); Riggle (2010)

- Given a constraint set $\mathbb{C}$ with size $|\mathbb{C}|$, and a set of tableaux for various UR's,
  - The number of distinct languages (sets of winners) is typically less than the number of total rankings ($|\mathbb{C}|!$)
- Reason: not all constraints are able to conflict with each other, so some distinct total rankings converge on the same language
- Hypothesis: more rankings that generate a language $\Rightarrow$ greater typological frequency

Illustration: Bane & Riggle (2008)

- Set of 12 constraints for quantity-insensitive (QI) stress, from Gordon (2002)

Align(σ́/σ̀)-Edge	Assign one * for each unstressed σ at the edge of a
Align(σ́)-L, Align(σ́)-R	Assign one * for each σ separating stress from th edge of the word
Align(σ́/σ̀)-L, Align(σ́/σ̀)-R	Assign one * for each σ separating stress from th edge of the word
NonFinality	Assign one * for stress on the final σ
*Clash	Assign one * for each sequence of two stressed σ's
*Lapse	Assign one * for each sequence of two stressless σ
*Lapse-L/R	Assign one * if neither of the initial/final two σ stressed
*ExtendLapse	Assign one * for each sequence of three stressless
*ExtLapse-R	Assign one * if none of the final three σ's is stresse

The predicted typology

- Gordon also assumes one meta-ranking
  - Either ALIGN(ǿ)-L or ALIGN(ǿ)-R must be ranked lowest (or ‘inactive’/absent)
- Two sets of 11! total rankings = 79,833,600
- Reduce to 152 distinct patterns
- Predicts 24 attested patterns, 128 unattested patterns[1]
- Every predicted pattern can be derived by multiple total rankings
  - Smallest by 23,760; largest by 11,880,000

[1]Bane and Riggle note two attested patterns that were not known to Gordon, and are not generated by his constraint set.

Relative frequencies

- The “largest” ranking: 11,880,000
  - Fixed initial stress, no secondary stress
  - Example: Koromfé (Gur; Rennison, 1997)
- The smallest: unattested

Comparing to typological frequency

- Bane and Riggle compare r-volumes to empirical counts from the StressTyp database
- 306 QI stress systems
- Examples

Pattern	Rankings	r-vol	Lgs	Emp. prob.
Fixed initial	11,880,000	14.9%	69	22.5%
Penultimate	228,0960	2.9%	60	19.6%
Antepenultimate	285,120	0.4%	8	2.6%
Peninitial	31,6800	0.4%	12	3.9%
(Unattested)	23,760	0.3%	0	0%

- What is the relation between r-volume and empirical frequency?

Bane and Riggle's result

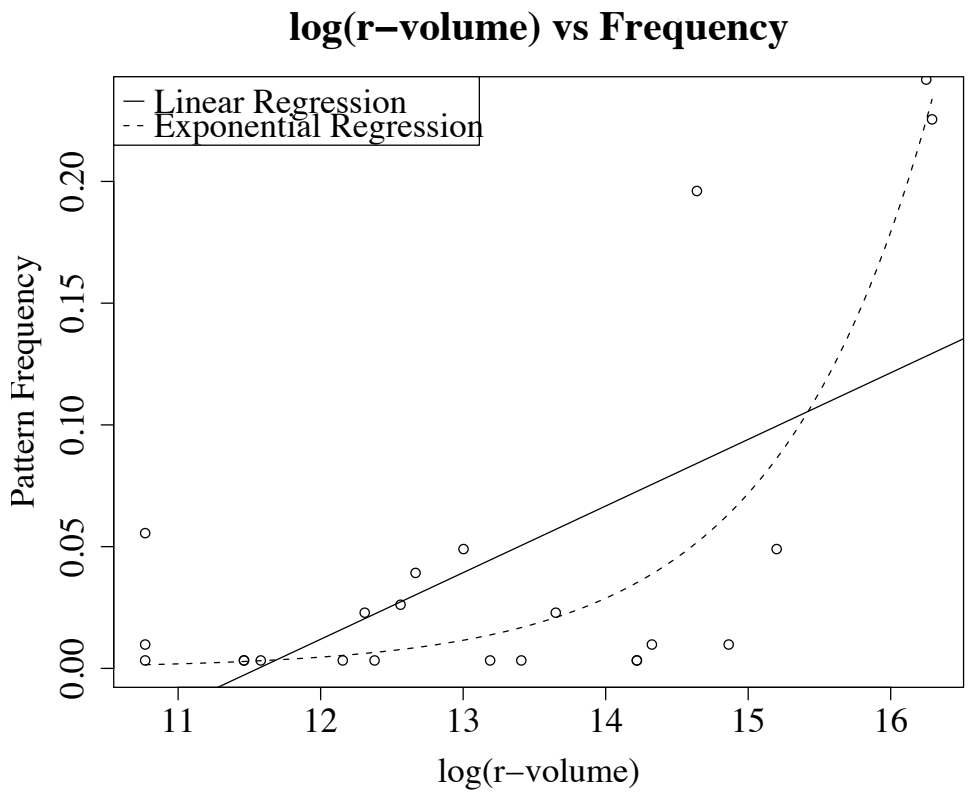

Figure 7: Linear and exponential regressions of typological frequency as a function of the natural logarithm of the pattern's r-volume.

Summary

- The very most popular patterns generally have higher r-volumes
  - Final .242, initial .225, penultimate .196, initial/alternating/non-fin .056, initial/alternating .049, penultimate alternating .049, peninitial .039, antepenultimate .026
- Significant, but also not especially great predictor

Rankings or parses

Rankings, or parses?

- Recall Kager (2012): modeled stress windows at word edges, using various foot-based constraint sets
- Feet introduce “hidden” (silent) structure: determine a phonetic property (position of stress), but are not themselves present phonetically
- Multiple footings can yield the same output
  - σ(σ́σ) vs. (σσ́)<σ>
  - These correspond to different rankings, each with their own r-volume
- Kager’s hypothesis: stress patterns with more parses are more frequent typologically

Kager's observation: parses and attestation

N	Ex.
16	Modern Greek
16	Spanish
11	Kobon
10	—
9	Comanche
9	—
8	Aguaruna
8	Kashaya
8	Azkoitia Basque
8	—
6	Gidabal
6	Korafe
6	Hopi
6	—
6	Terêna
6	—
6	Pirahã
5	—
5	—
5	—
5	Latin

N	Ex.
5	Ctr, W. Macedonian
4	—
4	—
4	—
4	—
4	—
3	Ossetic
3	—
3	Choguita Rarámuri
3	—
3	—
3	—
3	—
3	—
3	Kunjen
3	Yapese
3	—
3	—
3	—
3	—
2	—

N	Ex.
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
2	—
1	—
1	—
1	—
1	—

N	Ex.
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—
1	—

Kager's observation: parses and attestation

- Patterns are more likely to be attested if they can be produced in multiple ways (more parses)
- Kager does *not* consider the relation between number of parses and number of rankings
  - If a pattern can be produced with multiple parses, by necessity it can be produced by multiple rankings
  - However, it is also possible for a pattern to correspond to exactly one parse, but supported by many rankings
- Question: is Kager's observation really about r-volume, and not the number of parses?

A toy example to test this

Peninitial vs. penultimate stress

- Penultimate stress: two parses
  - Final trochee: σσ(σ́σ)
  - Final iamb + final σ extrametricality: σ(σσ́)⟨σ⟩
- Peninitial: just one parse
  - Initial iamb: (σσ́)σσ
  - Assumption: no such thing as initial syllable extrametricality
- Empirically: penultimate ≈5 times more frequent
  - 60 vs. 12 lgs in StressTyp

Rankings for penultimate vs. peninitial stress

- Peninitial
  - Iambs:

    All-Ft-L ≫ All-Ft-R

    Iamb ≫ Trochee
- Penultimate
  - Iambs:

    All-Ft-R ≫ All-Ft-L

    Iamb ≫ Trochee

    NonFin ≫ All-Ft-R
  - Trochees:

    All-Ft-R ≫ All-Ft-L

    Trochee ≫ Iamb

    All-Ft-R ≫ NonFin
- Although there are two parses for penultimate stress, both involve more crucial rankings (smaller r-volume?)
- Question: does frequency of penultimate stress

The r-volume of Kager's parses

- Constraint set
  - All-Ft-L, All-Ft-R
  - Iamb, Trochee
  - NonFin
- Candidate set
  - Syllables parsed or unparsed into feet
  - Feet: iamb (σσ́), trochee (σ́σ), degenerate (σ́)
- Introduces one additional parse for penultimate stress: ⟨σσ⟩(σ́)⟨σ⟩ (non-final degenerate trochees)
- Also one additional mostly-peninitial language: initial in disyllables, peninitial otherwise

The typology

- Out of 120 (=5!) logically possible languages, only 11 distinct languages/patterns
- Focus here on “penultimate” and “peninitial” patterns
- R-vol of penultimate = .225, peninitial = .058/.125

	Parse	r-Vol
Penultimate	⟨σσ⟩(σ́σ)	0.092
	⟨σ⟩(σσ́)⟨σ⟩	0.033
	⟨σσ⟩(σ́)⟨σ⟩	0.100
Peninitial	(σσ́)⟨σσ⟩	0.058
	(σσ́)⟨σσ⟩ + nonfin	0.067

More generally…

Pattern	Parses	R-Vol	StressTyp2
initial	3	0.3750	101
final	2	0.2417	104
penult	3	0.2250	75
peninit	1	0.0583	16
peninit+nonfin	1	0.0667	1
antepenult	1	0.0333	11

Summary of this demonstration

- Turns out that r-volume and number of parses are correlated
- Other cases may yet show inverse predictions for r-vol and number of parses
- Both favor the typologically more common pattern, to some extent
    - Strength of preference less than observed asymmetry

References

Bane, Max, and Jason Riggle. 2008. Three correlates of the typological frequency of quantity-insensitive stress systems. *Proceedings of the tenth meeting of acl special interest group on computational morphology and phonology*, SigMorPhon '08, 29–38. Stroudsburg, PA, USA: Association for Computational Linguistics. Online: http://dl.acm.org/citation.cfm?id=1626324.1626330.

Gordon, Matthew. 2002. A factorial typology of quantity-insensitive stress. *Natural Language & Linguistic Theory* 20.491–552.

Kager, René. 2012. Stress in windows: Language typology and factorial typology. *Lingua* 122.1454–1493.

References (*cont.*)

Rennison, John. 1997. *Koromfe*. New York: Routledge.

Riggle, Jason. 2010. Sampling rankings. University of Chicago ms.