From experiment results to a constraint hierarchy with the 'Rank Centrality' algorithm

Jennifer L. Smith
University of North Carolina at Chapel Hill
jlsmith@email.unc.edu

LSA 2020

1. The question: RC for phonology?

- **Rank Centrality** (**RC**) (Negahban et al. 2017): A rank-aggregation **algorithm** that computes a **total ranking** of elements from **noisy pairwise ranking** information
- *Question:* Given **forced-choice experiment results**, can RC model a speaker's **overall phonological grammar**?
  - How do RC results (*one-step algorithm*) **compare to the GLA** (*incremental learning*)?

2. Test-case experiment (Smith & Tashiro 2019)

Do Japanese speakers have a **productive markedness hierarchy** for loan nativizations?

- Two-alternative forced choice: Which nonce loanword adaptation is 'more natural'?
  - Each trial compares two constraints from { **No[si]**, **No[ti]**, **No[dd]**, **No[nt]**, **No[p]** } (all attested in Japanese)
  - Example: For the nonce loan ***siftant***... Choosing [ɕiɸɯtanto] over [siɸɯtando] ⇒ **No[si]** (satisfied) » **No[nt]** (violated)
- What is each participant's **overall hierarchy**?
- Structure of the response data
  - 40 native speakers of Japanese
  - 4 nonce loans for each constraint pair
  - Each participant's **score** for each constraint-pair order (e.g., **No[si]** » **No[nt]**) = **proportion** of compatible nonce-loan responses: ***0***, ***0.25***, ***0.5***, or ***1***
- How can we interpret the results?
  - Need to compute the **overall hierarchy**
  - Each pairwise comparison is **probabilistic**
- State-of-the-art: GLA (Boersma & Hayes 2001)
  - *Incremental* stochastic error-driven learning

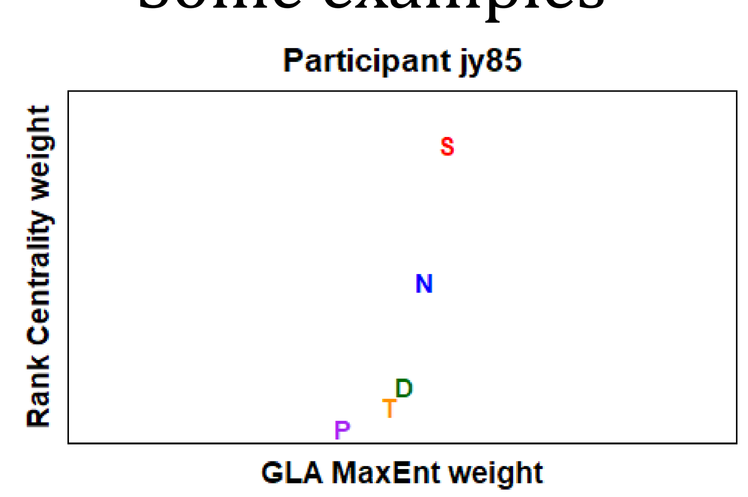
3. The RC algorithm

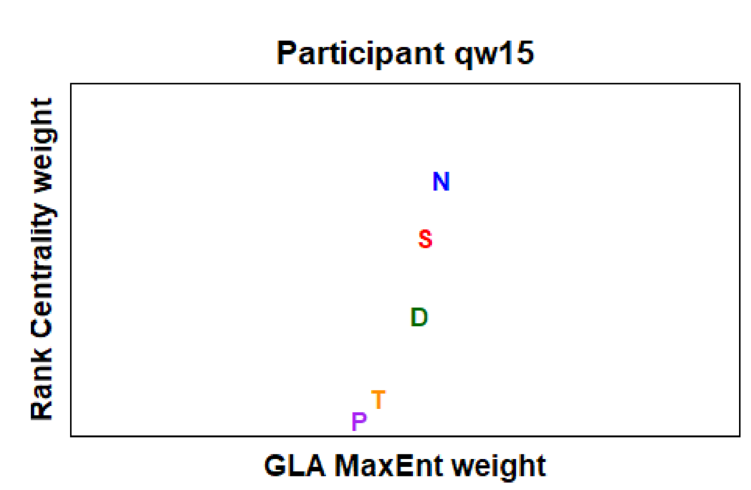
- Elements compared (here, **constraints**) are represented as nodes on a **directed graph**
  - Weight of edge E_{ij} corresponds to proportion of times node j is chosen over node i
- **Transition matrix** (random walk) is computed
  - **Stationary distribution** of graph (result of applying transition matrix repeatedly) = *largest left eigenvector* = **weights** of nodes
- RC thus assigns a value $0 \leq n \leq 1$ to each node
  - This represents both a **rank order** and a **distance** between elements in the graph
  - Resembles output of the Gradual Learning Algorithm (GLA) (Boersma & Hayes 2001)
  - → **Can RC replicate GLA in one step?**

4. From pairwise results to full hierarchy: Comparing RC and GLA

Methodology: For each participant...

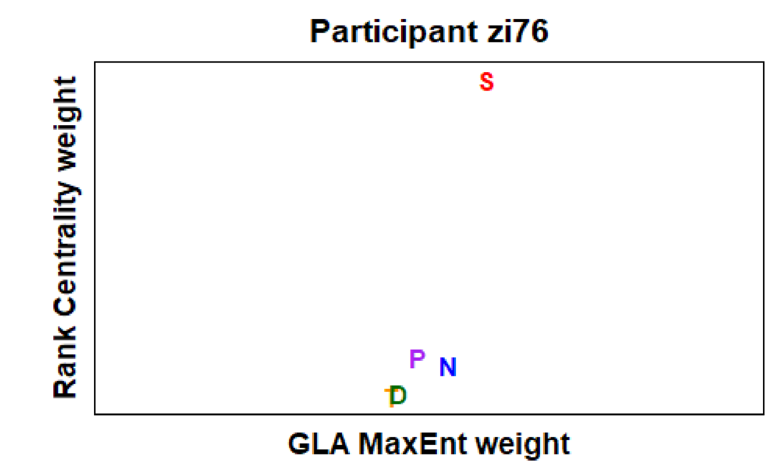
- **RC**: Apply RC algorithm to graph representing all C_i » C_j response proportions
  - Output: **weights per participant, 0≤*w*≤1**
- **GLA**: Apply MaxEnt learner (Jäger 2003) in Praat
  - *Initial State grammar* has all constraint weights at 50
  - *Pair Distribution* from Ci » Cj response proportions
  - Why MaxEnt? → RC based on multinomial logit model
  - Output: **weights per participant, 25≤*w*≤75**
- **Compare**: Plot RC results against GLA results
  - Some examples

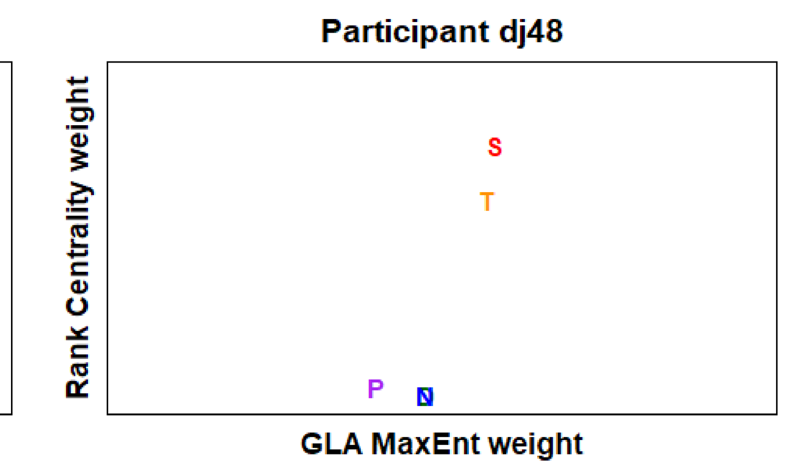




(1) Do RC and GLA derive the **same rank order for the constraints?** | **Mostly.**

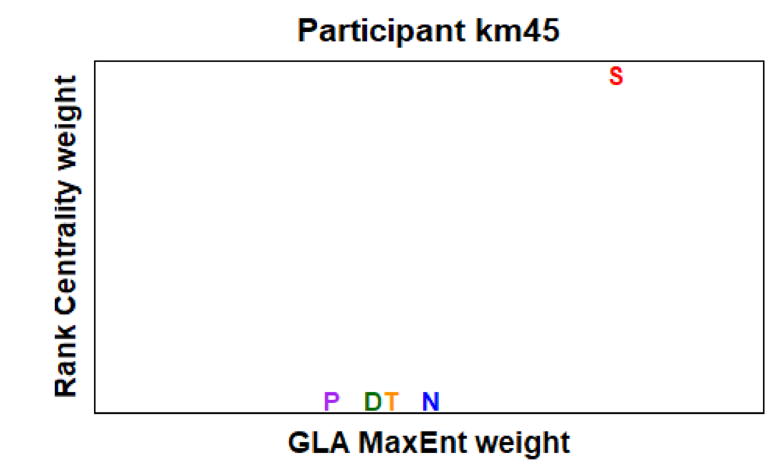
- 11/40 grammars: two constraints reverse
- But in all reversals, weights very close together
  - The two most extreme constraint reversals:

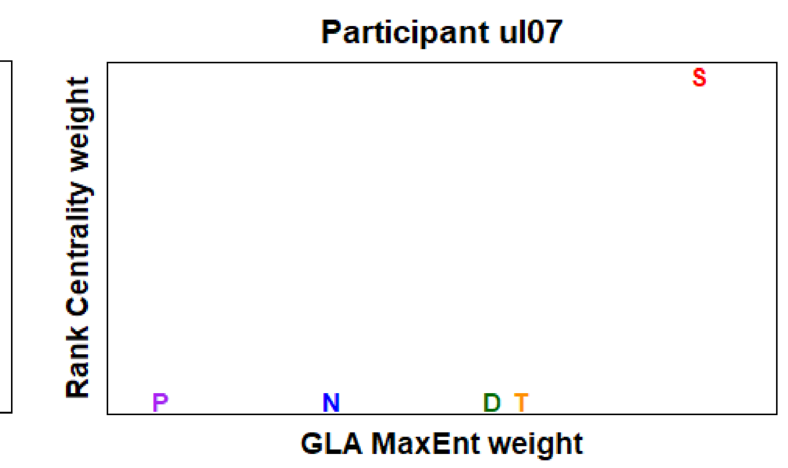




(2) Do RC and GLA produce the **same distances between constraints?** | **Not at the low end.**

- Many plots are S-shaped: RC makes the bottom of the range more **polarized** than GLA
  - 14 grammars have **No[si]** far above others: RC assigns 1 vs. 0, but GLA has more spread





5. Discussion and conclusions

- Results for **rank order** are very promising
- Results for **distance**: Why the low-end effect?
  - RC works best with large numbers of i–j comparisons—there may be too few here
  - *Regularized RC* includes a prior probability term to compensate for small numbers of observations: Try this next?

Acknowledgments

Many thanks to Elliott Moreton for comments, discussion, graph theory, and linear algebra.

References

Boersma, P., & B. Hayes. 2001. Empirical tests of the Gradual Learning Algorithm. *Linguistic Inquiry* 32: 45–86.
Jäger, G. 2007. Maximum entropy models and stochastic Optimality Theory. In A. Zaenen (ed.), *Architectures, rules, and preferences*. Stanford: CSLI, 467-479.
Negahban, S., et al. 2017. Rank Centrality: Ranking from pairwise comparisons. *Operations Research* 65: 266-287.
Smith, J.L., & Y. Tashiro. 2019. Nonce-loan judgments and impossible-nativization effects in Japanese. *LSA* 4: #26.