Introduction

- Phonological inventories are the results of the interaction between internal (articulatory, perceptual and cognitive) and external (e.g. sociolinguistic) constraints ([1], [2], [3], [4], [5]).
- A system can be said to be coherent if it satisfies internal constraints [6].
- Internal constraints lead systems to diachronically stabilize around maximal coherence; evolution accommodating only for these constraints is deterministic.
- External constraints can disrupt this evolution toward less coherent systems; they bring "stochasticity".
- The synchronic frequency of distribution of an inventory reflects i) its response to internal constraints and ii) its positioning or not at the crossroad of evolutionary trajectories.

Goal

- Design an evolutionary model based on the synchronic distribution of phones in actual inventories.
- Analyse inventories in the light of notions of coherence and stability by considering frequencies of individual segments and pairs of segments.
- Assess the weight of external constraints.

Data & methodology

Data

- The genetically and geographically balanced UPSID database ([7], [8]).
- 451 languages, 833 phonemes, 96 articulatory features

Building a measure of coherence for inventories

- If inventories were randomly built, all segments would have a probability of occurrence of 0.5; the intrinsic strength of a segment relates to how much it deviates from randomness.
- If segments behaved independently, the probability of 2 segments to co-occur would be equal to the product of their respective frequencies; the interactional strength of a pair of segments relates to how much it deviates from independence
- The log of the binomial test quantities the previous deviations.
- For individual segment, the sign tells whether this segment is favored or disfavored in systems; for pairs of segments, it indicates whether these two elements 'attract' or 'repulse' each other.

Expected distribution | Actual distribution | Strengths
---|---|---
/a/ | 225 | /a/ | 992 | $\text{str}(\text{/a/}) = 140.4$
/A/ | 225 | /A/ | 83 | $\text{str}(\text{/A/}) = 100.0$
/a | 2 | /a/ | 2 | $\text{str}(\text{/a/a/}) = 2.2$
/A | 2 | /A | 2 | $\text{str}(\text{/A/A/}) = 8.3$

- For any given system, we obtain its intrinsic strength by summing the intrinsic strengths of all present segments AND absent segments.
- We compute the interactional strength of the system by adding the interactional strengths of all pairs of present AND absent segments.
- The (global) coherence of a system results in the combination of these intrinsic and interactional strengths. As it takes into account all possible segments and not just those present, it is independent of the size of the system.

Deriving stability and evolutionary trajectories from coherence

- An index of stochasticity S models deterministic vs. stochastic evolutions.
- From a given inventory, we can compute a set of new inventories differing by a few units and rank them according to their coherence.
- A deterministic model (high S) will tend to maximize coherence by choosing the highest ranked output as the next evolutionary step.
- A stochastic one (low S), incorporates 'external factors' by assigning to each potential new system a probability of being chosen.

Results

Coherence of real and artificial systems

- Figures, from left to right, represent the intrinsic, interactional and global coherences of UPSID and random systems (red and blue markers respectively).
- The third figure also indicates the most frequent systems in UPSID; they appear to be the ones with the highest coherence.

Stability of real and artificial systems

- Top figures: stability of UPSID and random systems (red and blue markers respectively); Bottom figures: global coherence of UPSID systems and of the systems they became after a 50-step evolution (blue and purple markers respectively).
- From left to right: 3 different S values - 0.008 (mixed model), 0.5 (deterministic) and 0.001 (stochastic).
- Both deterministic and stochastic models led to new systems that are not consistent with the actual ones in terms of global coherence; the intermediate model produced more consistent systems, at least for sizes up to 11 segments.
- The comparison of the 3 S values seems to indicate that the evolution of phonological inventories is primarily internally driven.
- Starting from specific systems, we studied the outcome of 500 50-step evolutions; results show that systems with high coherence tend to be preserved whereas systems with low coherence will systematically get transformed.

Conclusions & Perspectives

- The notions of coherence and stability used in our model seem to capture at least part of the forces driving evolution of phonological inventories.
- More qualitative studies of the outcomes of the model are needed to test the validity of the observed changes.
- The next step is to test predictions of the model with actual known evolutionary paths.

References