

# What is the driving force behind the emergence of compositionality?

Willem H. Zuidema

Artificial Intelligence Laboratory

Vrije Universiteit Brussel

Pleinlaan 2, B-1050 Brussels, Belgium

*jelle@arti.vub.ac.be* <http://arti.vub.ac.be/~jelle>

June 13, 2001

## Language games: learners that learn from learners

1. Representation bias
2. Search bias (procedural bias)
3. Collective dynamics

## Language games: learners that learn from learners

1. **Representation bias**

*functional semantics (e.g. Steels, 2000), construction grammar*

2. **Search bias** (procedural bias)

*selectionism, on-line learning (e.g. Batali, 2000)*

3. **Collective dynamics**

*cultural evolution (e.g. Kirby, 2000), “selforganization” (e.g. Hashimoto & Ikegami, 1996, Zuidema & Hogeweg, 2000)*

How to compare these different models?

- Horizontal vs. vertical transmission
- Genetic vs. cultural transmission
- Learning modelled or not modelled
- Supervised, non-supervised or reinforcement learning

Formalization!

The mathematical model of Nowak et al. (2001):

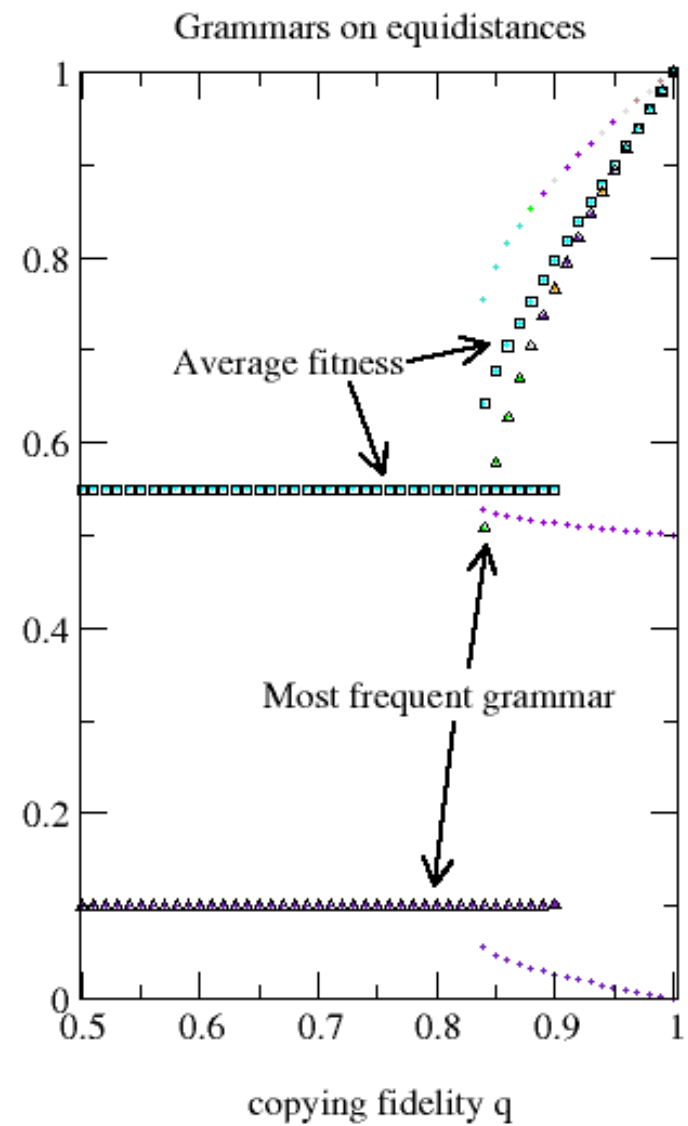
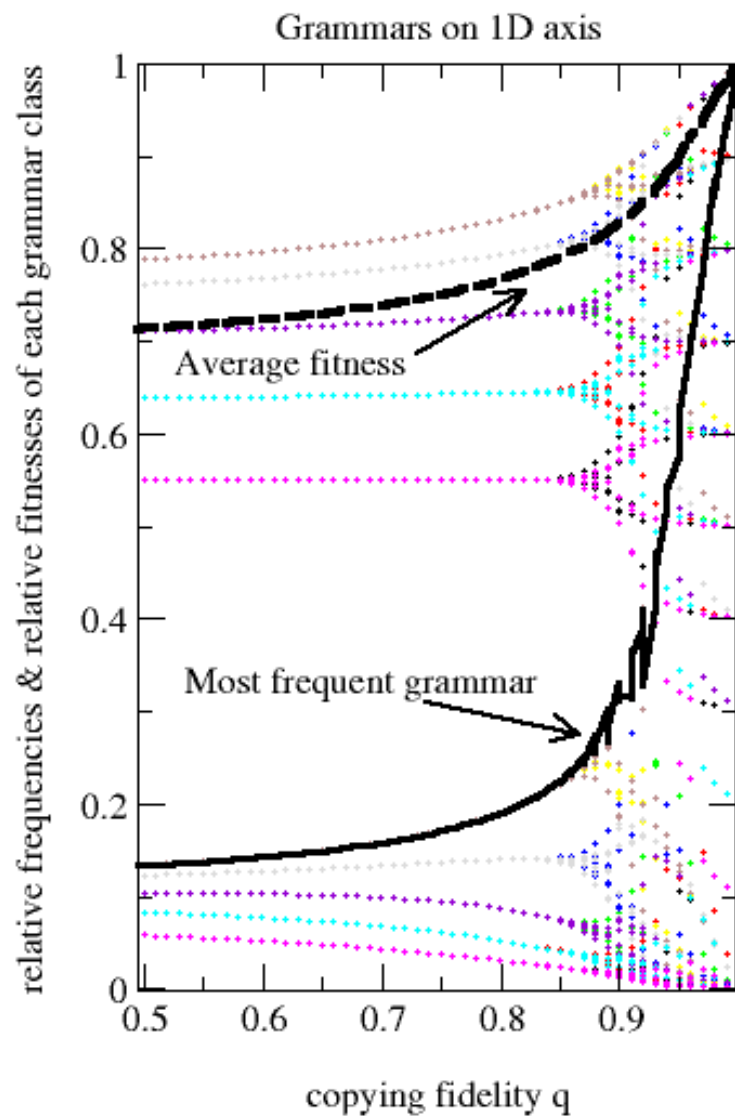
$$\dot{x}_i = \sum_j^N x_j f_j Q_{ji} - \phi x_i \quad (1)$$

- $f_i$  is the *relative fitness* (quality) of grammars of type  $i$  and equals  $f_i = \sum_j x_j F_{ij}$ , where  $F_{ij}$  is the expected communicative success from an interaction between an individual of type  $i$  and an individual of type  $j$ . The relative fitness  $f$  of a grammar thus depends on the frequencies of all grammar types, hence it is *frequency dependent*. The proper way to choose  $F$  depends on the characteristics of **language use** (production and interpretation).
- $Q_{ij}$  is the probability that a child learning from a parent of type  $i$ , will end

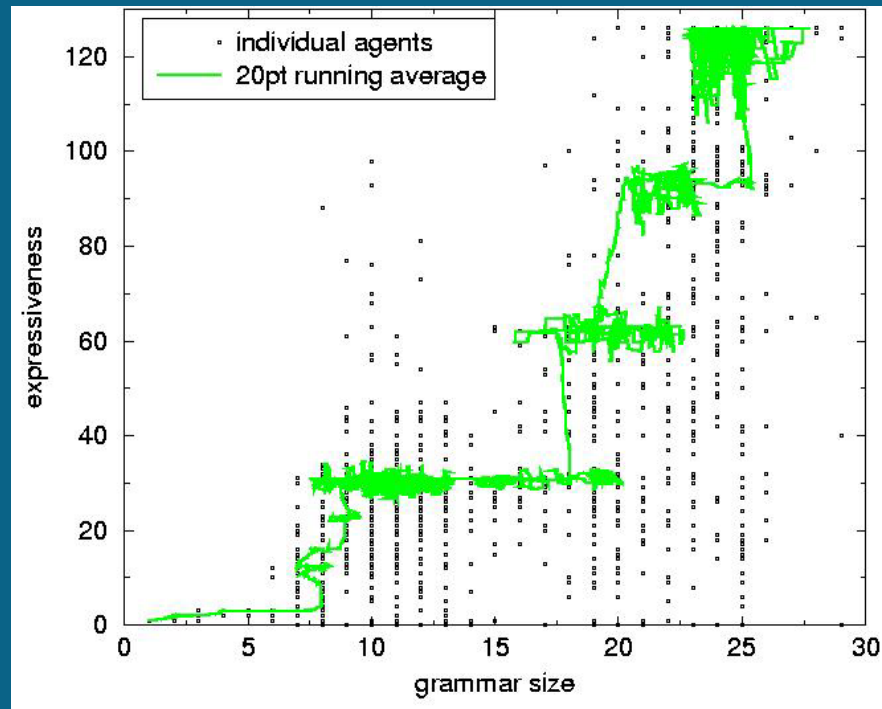
up with grammar of type  $j$ . The probability that the child ends up with the same grammar,  $Q_{ii}$ , is defined as  $q$ , the copying fidelity. The proper way to choose  $Q$  depends on the characteristics of **language acquisition** (learning and development).

- $\phi$  is the average fitness in the population and equals  $\phi = \sum_i x_i f_i$ . This term is needed to keep the sum of all fractions at 1.

Below the **coherence threshold** in the copying fidelity, a stable language cannot establish



Conforming to the group has a major conservative influence, if communication is beneficial for both speaker and hearer



**Language adapts to learner**, such that the language becomes easier to learn than one would expect a-priori

#### Learning algorithm

- Incorporation
- Compression
- Generalization

#### Goals:

- Replicate Kirby (2000)
- Replicate Nowak et al. (2001)
- Show quantitatively that children need less training samples than the lower bound that they derive

## Conclusions

- Formalization helps to relate and compare distinct models
- The coherence threshold is a **necessary condition** for the emergence of a stable language in the population
- **Conforming to the group** is a possible problem for innovation

- Three candidates for the driving force behind the emergence of compositionality:
  1. need for higher expressiveness (agents)
  2. language adapts to be better learnable (language)
  3. elements of language compete with each other (rules/words)
- Future work:
  - show quantitatively that Nowak is wrong
  - evaluate if Kirby is right
  - repeat experiments with more interesting, but more complicated cognitive grammar formalism