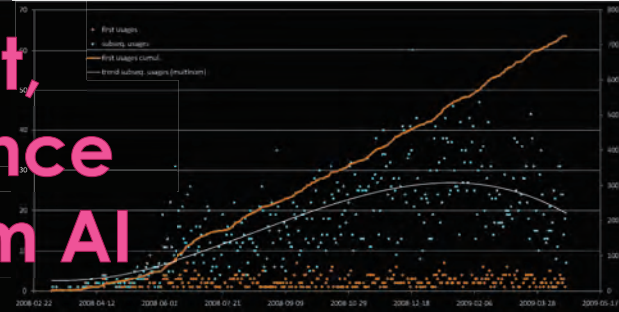


Follow the white rabbit, or how complexity science and linguistics can inform AI



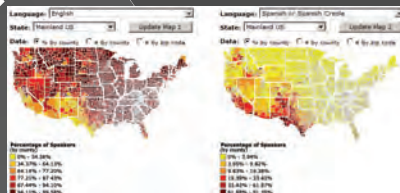
Michał B. Paradowski¹, Łukasz Jonak¹, Zoltán Kuscsik²
 1. University of Warsaw; 2. Université de Fribourg

Complex systems and language

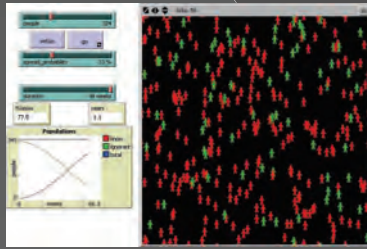
- variations and typology
- rise of new grammatical constructions
- language evolution
- spread and competition of both individual words and expressions, and entire languages...

Language(s) as (a) complex system(s)

- decisive factor – the impact of the structure of social networks (languages evolve and most often disappear due to human contacts or lack thereof) →
- formal modelling with the use of multi-agent simulations (microscopic models)



vs.



Inadequacy to the scenery of the 21st c.

- model takes into account only Euclidean relationships
- 'static'
- limited, identical number of 'neighbours' for every agent (4v8)
- identical perception of the prestige of a given individual by each of its neighbours
- invariant intensity of interactions between different agents
- no multilingual agents
- no memory effect
- zero noise

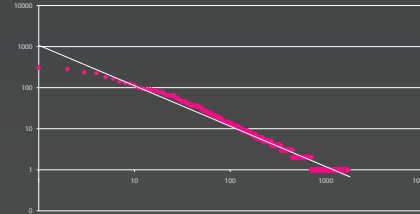
English on the Internet: a new word is born roughly every 98 minutes:

- alcolpop, Ardi, bangster (PL), brown state, choice mom, death panel, defriend, deleb, e-vampire, freemium, fundoo, funemployed, green washing, intoxiciated, jai ho, mobama, n00b, octomom, quendy-trendy, recessionista, seatmate of size, sexting, slumdog, teabagger, tramp stamp, wonderstar, zombie bank...

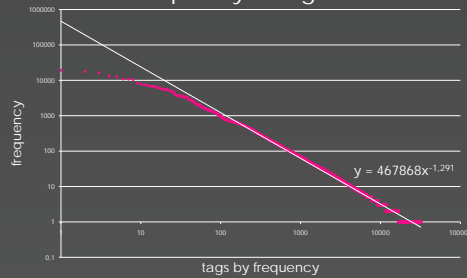


Diffusion of innovation is a function of the social structure: the chances for new content to spread are determined by the mutual connections between the participants in the communication network (Rogers 1962/2003; Watts 2007).

48-hr tag frequency distribution

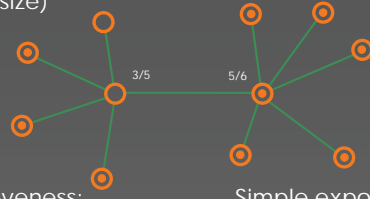


Overall frequency of tag occurrence



Social influence and diffusion:

- Mark Granovetter (1978) – threshold model of collective behaviour: an individual is likely to pursue a certain activity when s/he has observed it taken up by a certain number (proportion) of others
- Thomas W. Valente (1995) – network exposure: an individual only observes his/her network exposure (normalisation by ego-network size)

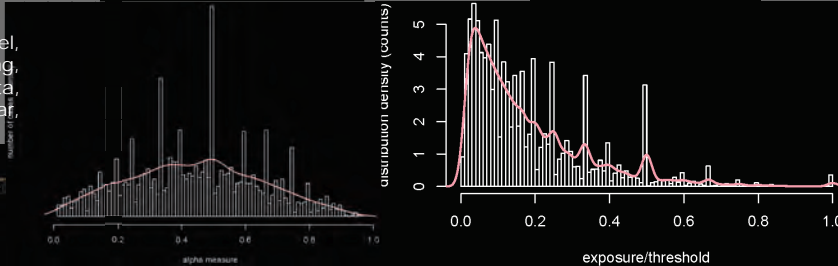


General innovativeness:

$$\alpha_u = \frac{\sum_{e_{u \rightarrow t} \in E(u)} a(e_{u \rightarrow t})}{|E(u)|}$$

Simple exposure:

$$\beta_u = \frac{\sum_{e_{u \rightarrow t} \in E(u)} \frac{A(e_{u \rightarrow t})}{H_i(u)}}{|E(u)|}$$



Implications:

- quantitative estimates on the spread of more general manifestations of culture;
- improvement of Web 2.0+ algorithms, social networking vortal applications, automatic text analysers and classifiers and meme aggregators;
- AI - delivery of parameters indispensable in the design of learning and interaction mechanisms (machine-human and machine-machine) on autonomous cognitive agents;

