

NOTES

* distributing bias "nearby" doesn't work properly when the network is not connected.
*

WHAT IS IT?

This is a linguistics model about how a language change may (or may not) diffuse through a social network. The key research question that it is interested in investigating is this: How can initially rare grammar variants become dominant in a population, without any global bias in their favor? It is known that such changes can and do occur in the real world - but what conditions are necessary to produce this behavior? This model demonstrates that the behavior can be reproduced through the use of simple cognitively-motivated agent rules combined with a social network structure and certain distributions of heterogeneous bias in the population. The language cascade occurs despite the fact that all of the agents' biases sum to 0.

While the model was developed for linguistics, there are potentially useful lessons to be learned about the interactions of heterogeneous agents in a social network, which may perhaps be applied to other disciplines, such as epidemiology, or the diffusion of innovation in marketing.

HOW IT WORKS

In this model, there are two opposing grammar variants (G0 and G1) in the population. Each agent's grammar value lies on the range between 0.0 and 1.0. The value 0.0 means that the agent only speaks grammar variant G0, whereas 1.0 means that the agent only speaks grammar variant G1. For grammar values between 0.0 and 1.0, an agent may speak either G0 or G1, with some probability. The probability is determined by a "production function", the shape of which depends on the CATEGORICALNESS parameter, as well as a 'bias' which can vary between agents (this 'bias' may be distributed in various ways, as shall be discussed in more detail later). It is called a "production" function because it maps a tendency toward one grammar or another into a probability of producing a token for one grammar or the other. If CATEGORICALNESS = 0, the production function is linear, meaning that agents produce G1 tokens with probability given directly by their grammar value, and G0 tokens otherwise. If CATEGORICALNESS > 0 the production function is nonlinear (in particular, sigmoidal). The agent's bias determines the point at which the production function crosses the line $Y = X$, which may be considered a repelling point, because if the agent's grammar value is below this repelling point and the agent were talking only to itself, it would eventually end up with grammar value 0.0, but if the grammar value started above this point, it would eventually end up at grammar value 1.0. The larger the CATEGORICALNESS parameter, the closer the sigmoidal production function is to a step function, and at CATEGORICALNESS = 100, the production function actually becomes a step function. This means that if the agent's grammar value is above a point (determined by its bias) it will only speak G1, and if it is below that point, it will only speak G0. In this case, agents are completely categorical about their speech, and are unwilling to mix the usage of the two competing grammars.

Over time each agent updates the state of its internal grammar value based on the tokens it is hearing from neighboring agents in the social network. More specifically, in each tick, each agent first produces a single token probabilistically, based on their grammar state and their production function. Each agent then updates their grammar state to be closer to the mean grammar value that they heard from all of their neighbors. We use what is sometimes called "alpha learning", whereby the new grammar state is a weighted average of the old grammar state with the mean of all the tokens produced by the neighbors. Thus, high degree nodes (agents) in the network (which we refer to as "influentials") are considered to be "heard" by many more nodes than low-degree nodes. However, the LEARNING-RATE (rate of change from the current grammar state

toward the perceived grammar of the neighbors) of all of the nodes is the same.

As an example, an agent that start with grammar value 1.0 will certainly produce a G1 grammar token in the first tick of the model. After one tick, it may have heard G0 tokens from their neighbors, and have adjusted their grammar value downward, meaning that the probability of producing G1 is no longer 100%. However, if the LEARNING-RATE is not too large, the agent's grammar value will probably still be quite high, which corresponds to a high likelihood of producing a G1 token in the next tick. However, over time the grammar value may undergo significant changes.

HOW TO USE IT / MODEL PARAMETERS

While the basic mechanics of the model are described simply above, there are numerous parameters, and ways to initialize or setup the model, to address different questions.

Here is a brief explanation of each parameter of control for the model, and how they related to the initialization and running of the model.

The social network structure (NETWORK-TYPE) may be initialized in several ways:

- * "spatial" causes nearby agents (in Euclidean space) to be linked together
- * "random" refers to Erdos-Renyi random graphs
- * "preferential" refers to the Barabasi-Albert preferential attachment method of creating scale-free networks. The method has been extended slightly to handle the creation of networks with odd average degree, by probabilistically choosing to add either K or K+1 edges as each new node is attached to the network.
- * "two-communities" consists of two "preferential" networks loosely connected to each other by some number of links (specified by the INTERCOMMUNITYLINKS parameter).

The network is created with the specified NUMBER-OF-NODES and AVERAGE-NODE-DEGREE.

By default, nodes start with an internal grammar value of 0.0, meaning they have no chance of ever using variant G1. The NUM-START-WITH-G1 parameter, however, controls the number of nodes in the network that start with grammar value 1.0.

If START-TARGET = "none", the agents are randomly chosen to start with grammar value 1.0. But if START-TARGET = "influentials", then the 1.0 grammar value is assigned by starting with the START-TARGET-RANK most influential agent and going down in order. For instance, if START-TARGET-RANK = 9, and NUM-START-WITH-G1 = 3, then the 10th, 11th, and 12th most influential agents (highest-degree nodes) will be assigned grammar value 1.0.

Each agent is assigned a bias toward one grammar variant or the other. The bias can range from +0.5 (strongly supporting G1) to -0.5 (strongly supporting G0). If BIAS-DIST = "flat", then all agents are assigned the same bias. If BIAS-DIST = "uniform-symmetric", then the biases are chosen symmetrically in pairs (X and -X) from a uniform distribution between -0.5 and 0.5. If BIAS-DIST = "normal-symmetric", then the biases are chosen symmetrically in pairs (X and -X) from a normal distribution, centered around 0, and with the log (base 10) of the standard deviation given by BIAS-STDEV-LOG10 parameter. The distribution is truncated at -0.5 and 0.5 (if a value is out of range, we redraw from the distribution).

Additionally, all agents' biases are affected by the GLOBAL-BIAS parameter.

The BIAS-TARGET parameter controls how bias is distributed in the social network. If BIAS-TARGET = "none", then bias is randomly distributed. If BIAS-TARGET = "nearby", then bias is distributed in sorted order (positive bias down to negative) starting with the most influential agent, down to the least influential agent. If BIAS-TARGET = "nearby", then bias is distributed in sorted order outward from a random one of the agents that is starting with the G1 grammar. This last method has the effect of creating a very favorable initial audience for

this G1 speakers, and (from our experiments) appears to greatly improve the chances of a language cascade.

The preceding discussion is most relevant for the "spatial", "random", and "preferential" network types. The grammar states and biases for the "two-communities" network-type are initialized according to different rules. In this case, two "preferential" network communities are created - one consisting initially of all G0 speakers and the other consisting of all G1 speakers. The COMA-START and COMB-START parameters control whether the bias is distributed in such a way that the community is more ripe for a language cascade to occur, or more resistant against change to the status quo. More specifically, in each community, the biases are distributed outward from a random node in sorted order (either up, or down, depending). In Community A, if the bias is distributed outward starting with positive bias (supporting G1) down to negative bias, then the network will be more "ripe" for a G1 cascade. On the other hand, distributing bias from negative bias (supporting G0) outward to positive bias will create a configuration that is more resistant to change. For Community B (which starts with G1 prevalent) the situation is reversed, but otherwise exactly the same.

The links between these two communities are chosen based on the COMA-BRIDGE-BIAS and COMB-BRIDGE-BIAS parameters. If COMA-BRIDGE-BIAS = 0, then the agents in Community A that are most biased towards G0 will be chosen as "bridge" nodes - meaning they will be linked to the other community. If COMA-BRIDGE-BIAS = 1, then the agents most biased towards G1 will be bridge nodes. Similarly, COMB-BRIDGE-BIAS determines which nodes will be bridge nodes in Community B.

As mentioned above, the CATEGORICALNESS parameter affects the degree to which nodes are willing to speak the two grammar variants interchangeably, rather than having a stronger preference to speak consistently, or semi-categorically.

The LEARNING-RATE parameter controls the rate at which agents speaking affects other agents internal grammar values. The grammar value of each agent is updated by a weighted sum of the old grammar value and the mean heard grammar of its neighbors. The LEARNING-RATE is the weight given to new data, while $(1 - \text{LEARNING-RATE})$ is the weight given to the old grammar value.

The PROBABALISTIC-SPEECH? parameter controls whether agents always speak 0 or 1 tokens probabilistically (ON), or else speak the real-valued numbers between 0 and 1 produced by their production functions (OFF). The default is for PROBABALISTIC-SPEECH? to be ON. However, turning it OFF could correspond to a longer iterated batch learning process. In many ways, turning it OFF has the effect of removing some noise from the system, and causing faster convergence to an equilibrium. However, the noise *can* be crucial in certain situations, and the behavior will be different. There may be some interesting avenues for further research here...

The VISUALS? parameter turns on or off the visual display. Turning VISUALS? OFF can help speed up the runs when running experiments. It will not effect the outcome of the model at all.

The COLOR-BY-BIAS and COLOR-BY-GRAMMAR buttons affect the visualization of the network, scaling low values (0.0 grammar, or -0.5 bias) to black, and high values (1.0 grammar, 0.5 bias) to white.

The LAYOUT button can be used to try to improve the visual aesthetics of the network layout. Note that this only affects visualization, and does not affect the model itself.

The SETUP button initializes the model, and the GO button runs the simulation until it has converged (or forever, if it does not converge). The STEP button goes one tick at a time.

Various monitors show statistics (min, max, mean) for the grammar values or

grammar biases. The "GRAMMAR STATE" plot also plots the mean internal grammar value of the agents over time, as well as the mean spoken value.

THINGS TO NOTICE

THINGS TO TRY

EXTENDING THE MODEL

RELATED MODELS

Language Change (by Celina Troutman)

CREDITS AND REFERENCES

Written by Forrest Stonedahl, in collaboration with Janet Pierrehumbert and Robert Daland.