

Self-organisation of vowel systems in scale-free networks

Wolf De Wulf

Abstract

Simulations in agent-based systems that model populations of communicating agents have been used to show that due to interactions between these agents and self-organisation, realistic vowel systems emerge. Various independent studies have reported this phenomenon under various parameter configurations. The models used in these studies range from modelling only what is necessary to demonstrate emergence to modelling complex real life factors to investigate the impact on the emerged vowel repertoires. A common factor in a lot of these studies is the unrealistic assumption that every agent regularly communicates with every other agent. This article presents the analysis of a reimplementation of one of the first agent-based models that was created to analyse emergence in vowel systems. The model is extended to run simulations in scale-free networks. While scale-free networks are known to be rare in real life society, the simulations that were ran have allowed for gaining insights into the biases of the original system and into how it could be made to model society more accurately.

1 Introduction

It has been known for some time now that self-organisation through repeated interactions between people explains more about the evolution of the sound systems of humans than innate predispositions do. Liljencrants and Lindblom (1972) showed that realistic vowel systems emerge when a measure of energy, based on thermodynamic statistics, is optimised in an acoustic space of vowels. After that, de Boer (2000) was the first to model a small population of agents that can communicate with each other through a number of steps that are together deemed the ‘imitation game’. The agents in de Boer’s (2000) model had no innate predispositions and were still able to converge to a global and realistic vowel system. Thereby showing that innate predispositions are not needed to explain the evolution of vowel systems. However, de Boer’s (2000) model contained a number of ad hoc methods that could have been left out without influencing the results. Successing models such as those by Oudeyer (2005) and de Boer and Zuidema (2010) did so successfully while also investigating other questions such as whether or not agents should explicitly be able to communicate or not and whether combinatorial structure emerges when instead of single vowels, sequences of vowels are modelled. A common factor in the older models and even in most of the newer models is the assumption that every agent regularly communicates with every other agent. In reality such fully connected networks are only found in very small communities or in families. Therefore, these models might not capture dynamics that are present in large real life populations. To address this, the model that is presented in this article reimplements de Boer’s (2000) model to run its simulations in a scale-free network.

Scale-free networks are networks whose degree distribution follows a power law. Such networks consist of two types of nodes: nodes with a high degree called ‘hubs’ and nodes with a small degree. While many real life networks were thought to be scale-free, recent studies involving new rigorous statistical analysis techniques have showed that this is not the case (Clauset et al., 2009; Broido & Clauset, 2019). Examples of networks that are thought to be scale-free are financial networks (De Masi et al., 2006; Soramäki et al., 2007), computer networks such as the internet (Faloutsos et al., 1999) and some social networks such as the co-authorship of papers (Newman, 2001). In the context of modelling a population, the idea of social networks following a scale-free structure is interesting. Even though a network of, for example, co-authorship in mathematical papers is not one in which phonology plays an important role, investigating what happens when an agent-based model that is designed to model a population of communicating agents is adapted to run its simulations in such a network, should prove to be insightful. This is exactly the case for the de Boer’s (2000) model. The reimplementation of this model that is presented in this article shows that the ad hoc methods that de Boer (2000) uses to model agents introduces a model-specific bias. Insights into how the model can be adjusted to generalise to more realistic networks are gained through a number of experiments.

In what follows, Section 2 gives an in depth discussion on scale-free networks, the methods that can be used to generate them and their . Subsequently, Section 3 repeats the architecture of de Boer’s (2000) model and the adjustments that are made for it to run its simulations in scale-free networks. Afterwards, Section 4 establishes a number of experiments of which Section 5 discusses the results. Section 6 concludes the article with a number of conclusions and possible future work.

2 Scale-free networks

Scale-free networks are networks whose degree distribution asymptotically follows a power law. In a pure scale-free network the probability of a node to have k edges to other nodes is defined as follows:

$$P(k) \sim k^{-\gamma} \quad (1)$$

with γ a scaling parameter typically in the range $2 \leq \gamma \leq 3$. Evidence for scale-free structure in a network can be found using statistical techniques. Broido and Clauset (2019) summarise the different notions of evidence in a set of categories which are not repeated here. For the rest of this article it suffices to know that the used scale-free networks belong to the “strongest” category, which is due to the method that is used to generate them. The scale-free networks in this article are generated using the most known mechanism to do so, called preferential attachment (Albert & Barabási, 2002). The preferential attachment mechanism to generate a scale-free network of size N is given by Algorithm 1.

Algorithm 1 Preferential attachment (Albert & Barabási, 2002)

```

1:  $net \leftarrow \text{fully\_connected}(4)$ ;
2: while  $net.size() < N$  do
3:    $A \leftarrow net.new\_node()$ ;
4:    $count \leftarrow 4$ ;
5:   while  $count \neq 0$  do
6:      $B \leftarrow net.random\_node()$ ;
7:     if  $!net.connected(A, B)$  then
8:        $net.connect(A, B)$  with probability  $\frac{B.degree}{net.total\_degree}$ 
```

The mechanism starts by initialising the network as a fully connected network of four nodes. Subsequently, the network is expanded by adding nodes until the network consists of N nodes. Each time a node is added, the following steps are repeated 4 times: select a random node from the network, if the random node is not connected to the new node, connect them with probability:

$$P(k_i) = \frac{k_i}{\sum_j k_j}$$

where k_i is the degree of the randomly selected node. Hence, preferential attachment is a probabilistic mechanism. New nodes are allowed to connect to any other node in the network. However, if a new node can be connected to a node with degree 2 or to a node with degree 4, the node with degree 4 is twice as likely to be chosen.

Whether or not preferential attachment can be used to accurately describe a population of humans is a topic that has sparked a lot of debate. Intuitively, the ‘hubs’ created by the preferential attachment mechanism seem to be realistic. By varying the scale and the context of the community a scale-free network is supposed to model, its adequacy varies as well. Take, for example, a university community consisting of professors and students. In such a case, a scale-free network seems like an adequate depiction of the underlying network. On the other hand, a much smaller community of hunter-gatherers might be more accurately modelled by a fully connected network. By analysing a variety of real life networks, Broido and Clauset (2019) came to the conclusion that social networks are at best ‘weakly’ scale-free. Nevertheless, making models such as those by de Boer (2000) and Oudeyer (2005) run their simulations in large and complex social networks would result in incorrect representations of reality. When sound systems started emerging for the first time, social networks were surely much smaller and simpler than those analysed by Broido and Clauset (2019). Similarly to how language is said to be the result of the co-evolution of biological and cultural inheritance (de Boer & Thompson, 2018), the evolution of language and the structure of society are two concurrent evolutionary processes that influence each other. Therefore, modelling an evolving sound system in a fixed network structure is not realistic. A more realistic introduction of network structure into models that investigate the emergence of sound systems would require network structure to evolve as well. This does not mean that looking at what happens when a particular fixed network structure is added to a model that investigates the emergence of vowel systems is entirely useless. While the speeds at which language and societal structure evolve are likely to be much more comparable than those of biological and cultural inheritance, fixing network structure in a model can still provide insights into one direction of the co-evolution, namely, the influence of network structure on the evolution of language.

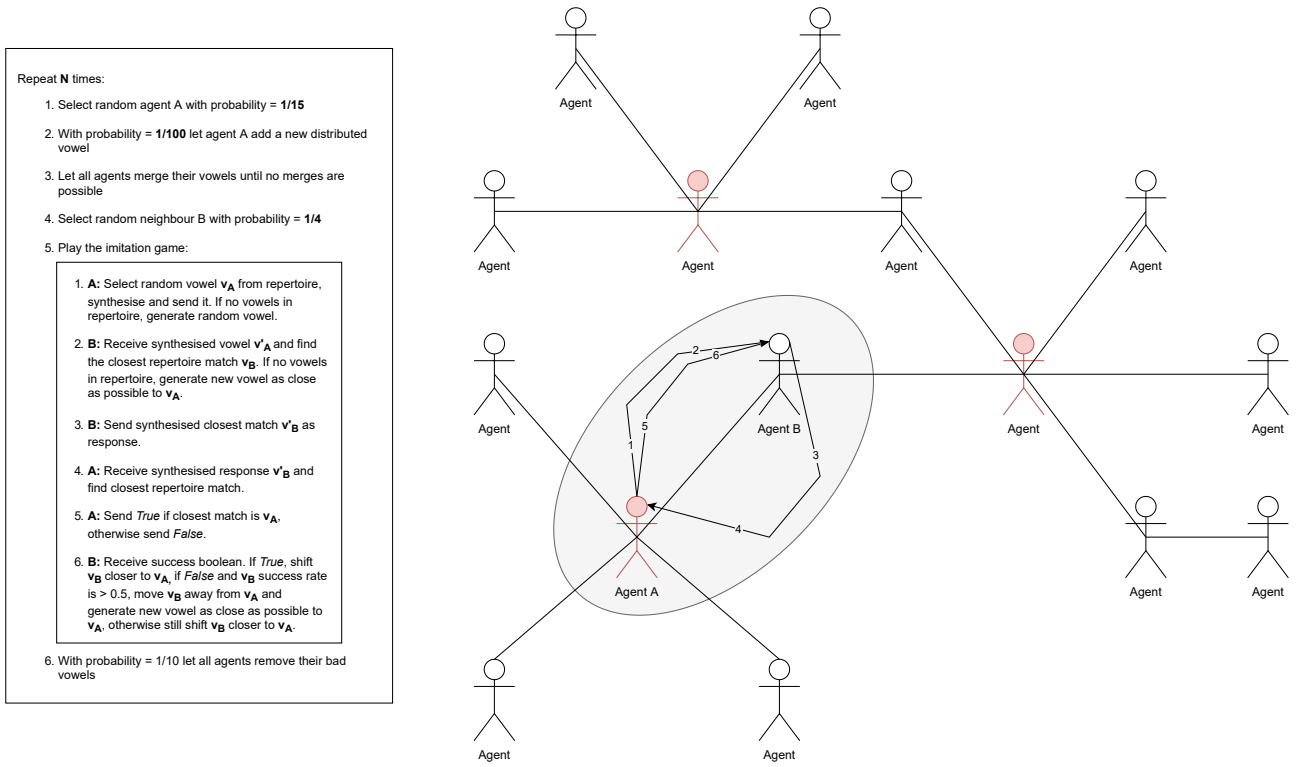


Figure 1: Graphical representation of the steps that are executed during a simulation.

3 Methodology

Aside from two changes, the model¹ that is presented here is exactly the same as that of de Boer (2000). Originally, the model was implemented in C++ and while this did not change, the model was reimplemented for compatibility reasons. The two aspects that were adapted in the first model are the following: the network architecture in which the simulations run and the bark-to-hertz and hertz-to-bark conversions.

Figure 1 gives a graphical representation of the steps that are executed during a simulation. Each iteration starts by selecting a random agent A from the complete population. Agent A is the *sender* of the iteration. With probability $\frac{1}{100}$, agent A is allowed to add a new distributed vowel to their repertoire. The distribution of vowels is implemented just as it was by de Boer (2000): by maximising the distance to all other vowels. Subsequently, every agent merges vowels that are too alike until no merges are possible. The merging threshold is 0.17 in articulatory space and $\frac{\log(1+\text{noise})}{0.1719} - \frac{\log(1-\text{noise})}{0.1719}$, defined by de Boer (2000) to be the maximum distance a vowel can be shifted by ambient noise, in acoustic space. After merging, a random neighbour B of agent A is chosen to be the receiver of the iteration. Originally, de Boer (2000) chose the receiver agent as a random agent from the complete population, thereby modelling a fully connected network. Once a sender and a receiver have been chosen, an imitation game is played. Agent A chooses a random vowel v_A from their repertoire, synthesises it, and sends it to agent B as an utterance v'_A , thereby adding a 10% noise shift. If agent A 's repertoire was empty, it is initialised with a completely random vowel. Agent B hears the synthesised vowel v'_A and finds the most resembling vowel v_B in their repertoire, synthesises it, and sends it to agent A as an utterance v'_B , once more, adding a 10% noise shift. Agent A hears agent B 's response and finds the most resembling vowel in their repertoire. If this vowel is the original v_A then the imitation game was successful and agent A sends *True* to agent B , otherwise the imitation game failed and agent A sends *False* to agent B . Lastly, agent B receives the success boolean and does the following. If it is *True*, agent B moves v_B closer to v_A , since by uttering v_B agent A was able to understand agent B . If it is *False* and the success rate of v_B is > 0.5 then agent B moves v_B away from v_A and adds a new vowel as close as possible to v_A . The reasoning behind this is that v_B is not a bad vowel because it is successful half of the times it is used. It is rather that agent A probably has another vowel that is closer to v_B than v_A is. Therefore, agent B should learn this vowel. If the received boolean is *False* and the success rate of v_B is ≤ 0.5 , then agent B still moves v_B closer to v_A , however, this does not add to the success rate of v_B . Finally, after the imitation game and with a

¹The source code of the model can be found on GitHub: <https://github.com/wulfdewolf/ProjectEoS>

probability of $\frac{1}{10}$, every agent is allowed to remove the bad vowels from their repertoires. Vowels that have been used at least 5 times and have a success rate $< 70\%$ are removed. As mentioned, the network that decides an agent's neighbours is generated using Algorithm 1. Synthesising, shifting vowels and distance measures are all implemented as they were by de Boer (2000).

The bark-to-hertz and hertz-to-bark conversions that were implemented by de Boer (2000) are interpolations based on data from other papers. The second adaptation to de Boer's (2000) model is the implementation of the conversions that are implemented in MATLAB²³, which are based on the work of Traunmüller (1990):

$$bark = \begin{cases} \frac{bark-0.3}{0.85} & \text{if } bark < 2 \\ \frac{bark+4.422}{1.22} & \text{if } bark > 20.1 \\ bark & \text{otherwise} \end{cases} \quad (2)$$

and:

$$bark = \frac{26.81 * hertz}{1960 + hertz} - 0.53 \quad (3)$$

$$bark = \begin{cases} bark + 0.15 * (2 - bark) & \text{if } bark < 2 \\ bark + 0.22 * (bark - 20.1) & \text{if } bark > 20.1 \\ bark & \text{otherwise} \end{cases} \quad (4)$$

While different conversions will most likely not change much in the results, for comparability it is better to use some standardized implementation.

4 Experiments

Two types of experiments are established. Firstly, the experiment and evaluation from Section 4.1 from (de Boer, 2000) is reproduced. This serves to demonstrate that the reimplementation of the model is correct. Both are also done for an instance of the model that runs its simulations in a scale-free network. By comparing the results of the two, insights can be gained regarding the research question that is addressed here. Additionally, for both models, a second experiment is conducted to investigate the influence of the population size on the results. Due to the difference in network structure of the two models it is reasonable to expect different results for this experiment.

The first experiment consists of investigating the emergence of a vowel system in a population of 20 agents under 10% acoustic noise. de Boer (2000) analyses the combined acoustic repertoires of all agents after 20, 500, 2000, and 10000 imitation games. The robustness of the emerged results is evaluated by repeating the simulations 1000 times for 5000 imitation games and plotting the success, size, and energy distributions. Both of these evaluations are repeated here. While the remaining other experiments from (de Boer, 2000) could also be repeated, since the model here is a reimplementation, the mentioned experiments should be enough to show that the reimplementation was done correctly. Running the other experiments for an instance of the model that runs its simulations in a scale-free network could also be interesting but it is not done in this article.

To allow for correct comparisons in the experiments where population sizes are varied, de Boer (2000) implements two changes in the model. Firstly, the number of games is kept per agent and not globally, hence, smaller populations play a smaller number of imitation games. Secondly, the probability of adding new distributed vowels must be made dependent on the number of games per agent instead of on the global number of played games. Therefore, the probability of adding new distributed vowels is set to $\frac{0.2}{N}$, with N the number of agents in the population. de Boer (2000) chose 0.2 as denominator such that for a population of 20 agents the insertion probability is 1%. Both of these changes are implemented here as well.

²<https://nl.mathworks.com/help/audio/ref/bark2hz.html>

³<https://nl.mathworks.com/help/audio/ref/hz2bark.html>

5 Discussion

6 Conclusions

References

- Albert, R., & Barabási, A.-L. (2002, January). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1), 47–97. doi: 10.1103/RevModPhys.74.47
- Broido, A. D., & Clauset, A. (2019, March). Scale-free networks are rare. *Nature Communications*, 10(1), 1017. doi: 10.1038/s41467-019-08746-5
- Clauset, A., Shalizi, C. R., & Newman, M. E. J. (2009, November). Power-Law Distributions in Empirical Data. *SIAM Review*, 51(4), 661–703. doi: 10.1137/070710111
- de Boer, B. (2000, October). Self-organization in vowel systems. *Journal of Phonetics*, 28(4), 441–465. doi: 10.1006/jpho.2000.0125
- de Boer, B., & Thompson, B. (2018, January). Biology-Culture Co-evolution in Finite Populations. *Scientific Reports*, 8(1), 1209. doi: 10.1038/s41598-017-18928-0
- de Boer, B., & Zuidema, W. (2010, April). Multi-Agent Simulations of the Evolution of Combinatorial Phonology. *Adaptive Behavior*, 18(2), 141–154. doi: 10.1177/1059712309345789
- De Masi, G., Iori, G., & Caldarelli, G. (2006, December). Fitness model for the Italian interbank money market. *Physical Review E*, 74(6), 066112. doi: 10.1103/PhysRevE.74.066112
- Faloutsos, M., Faloutsos, P., & Faloutsos, C. (1999, August). On power-law relationships of the Internet topology. *ACM SIGCOMM Computer Communication Review*, 29(4), 251–262. doi: 10.1145/316194.316229
- Liljencrants, J., & Lindblom, B. (1972). Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast. *Language*, 48(4), 839–862. doi: 10.2307/411991
- Newman, M. E. J. (2001, January). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences*, 98(2), 404–409. doi: 10.1073/pnas.98.2.404
- Oudeyer, P.-Y. (2005). The Self-Organization of Speech Sounds. *Journal of Theoretical Biology*, 233, Issue 3, Pages 435–449. doi: 10.1016/j.jtbi.2004.10.025
- Soramäki, K., Bech, M. L., Arnold, J., Glass, R. J., & Beyeler, W. E. (2007, June). The topology of interbank payment flows. *Physica A: Statistical Mechanics and its Applications*, 379(1), 317–333. doi: 10.1016/j.physa.2006.11.093
- Trautmüller, H. (1990, July). Analytical expressions for the tonotopic sensory scale. *The Journal of the Acoustical Society of America*, 88(1), 97–100. doi: 10.1121/1.399849