Age Adaptive Language Competition in Social Networks

Angad Singh Josan
Eastside Preparatory School
ajosan25@eastsideprep.org

May 31, 2024

1 Introduction

Of the 6,511 languages spoken across the world, over 1500 languages will be lost by the end of this century, with the rate nearing one language lost per month. [2] Better models to predict language loss and change are needed.

We propose a network-based model to model competition between languages, which depends on an individual's interactions with other languages, and on the individual's ability to learn a new language, based on their age.

2 Literature Review

Topology for network-based approaches has been a field of study as more complex systems emerge in society. Hallmark approaches include the model of preferential attachment [3], and the small-world model [10].

Earlier approaches to model the dynamics of language death yielded macroscopic differential equations. The Abrams and Strogatz model [1] considers a society with two types of individuals: those who speak language A, and those who speak language B. Their model uses the net probability of an individual switching between states to obtain a differential equation. The Mira and Paredes model [9] builds off the Abrams and Strogatz model, and considers a society where bilingual (AB) individuals can exist.

Previous network based models have modeled the diffusion of a new linguistic feature [5] across varying network types, implemented the Mira and Paredes model in a network-based manner [11], and have utilized sociolinguistic patterns to model the status of a given language [6]. None of these approaches has accounted for the age-related learning curve for new languages.

3 The Model

We generate both a Watts-Strogatz and a Barabasi-Albert network for our experiments. We use $n = 1000$, $m = 3$, $k = 4$, and $p = 0.3$ for all experiments. We initialize each individual's state using a weighted choice between categories: d_a, d_b, and d_{ab}.

We assign each node to a specific age based on data from the United Nations [4]. We assign the age based on the proportion of each age group with respect to the whole population, using a weighted choice mechanism between all ages. Ages range from 0-100.

Each individual switches between states by a probability $p_i j, \rightarrow k$, where j and k are arbitrary states. We multiply all p_i by a function $r(t)$, which is an empirical model [7] that defines an individual's ability to learn a language over time. This will decrease over time, decreasing the probability of a switch

$$R(t) = \left\{ \begin{array}{ll} r_0, & \text{if } t \leq t_c \\ r_0\left(1 - \frac{1}{1+e^{a(t-t_c-\delta)}}\right), & \text{if } t \geq t_c \end{array} \right\} \tag{1}$$

In this experiment r_0 is set to 1. This is the individual's original ability to learn a new language. t is the individual's age, a and δ define the steepness and offset of the function, respectively. We set a to 1 and δ to 0.5 for this experiment. Empirically, t_c is found to be 17.4 [7]. We use $t_c = 17$, as our model deals in integer ages only. These ages do not change as the model iterates.

Based on the work from Castello, Eguiluz, and Miguel [11], we obtain the following probabilities for an individual switching between states

$$p_{i,t,A \rightarrow AB} = s_B \sigma_i^B R(t)$$

$$p_{i,t,B \rightarrow AB} = s_A \sigma_i^A R(t)$$

$$p_{i,t,AB \rightarrow B} = s_B (1 - \sigma_i^A) R(t)$$

$$p_{i,t,AB \rightarrow A} = s_A (1 - \sigma_i^B) R(t)$$

In our model, we use the status $s_A = s_B = 0.5$. The parameter σ_i^j represents the density of individuals of who can speak j (an individual with AB will be double counted) in the neighborhood of node i, and

$$\sigma_i^a + \sigma_i^b = 1$$

4 Experiments and Results

We let the model diffuse over 300 iterations (years) for the Watts-Strogatz and Barabasi-Albert networks. Figures for different scenarios are in the appendices. We tested the following conditions for the distribution d:

Experiment 1:

$$d_a = d_b = d_{ab} = 0.33$$

Experiment 2:

$$d_a = 0.6, d_b = d_{ab} = 0.2$$

Experiment 3:

$$d_a = 0.5, d_b = 0.3, d_{ab} = 0.2$$

Experiment 4:

$$d_a = 0.4, d_b = d_{ab} = 0.3$$

We also obtain a consistent age and degree distribution, due to using a random seed = 300. The degree distribution for the Barabasi-Albert model is generated with $y = 3$, which corresponds to our data for the degree distribution. The presence of high degree nodes are very limited.

For experiments with relatively similar initial d values (1, 4), the spikes in certain states can be attributed to randomness with the distribution of nodes of type A (which are quite sparse). Overall, the number within each state level out after their initial spike, with a jagged pattern during the remainder, signifying consistent change within the population.

The small world model tends to have larger changes within the data, which makes sense as a more dense population is more connected and thus has σ_i^j values that tend to higher or lower numbers, depending on j.

5 Future Work

Potential future ideas with an age-based emerging population model are to add new nodes over time using UN provided birth rates, and to remove nodes using UN provided death rates. A model which changes the age of individuals over time will prove useful, as this will decrease the likelihood of older individuals to keep switching between states, thus providing continuity and reducing the jaggedness of the data.Another population idea to test is the Local World Model [8].A continuity function which decreases probability of state switch as the number of state switches increase is also likely helpful in this regard.

A probability-based model of social linking and delinking could change the distribution of states within communities. Assigning nodes to loose communities within the same linguistic state is another idea to consider.

Status parameters that are not 0.5, could provide the eventual dominance of one language over time. Utilization of link-based language states are another potential improvement, which would allow to model link based phenomenon such as code switching. The Minett and Wang model [6] provides interesting models for the status parameter within certain scenarios.

Varying the effects of more hyperparameters is futile without real-world data to utilize. To find these hyperparameters, a mechanism to fit empirical data would be useful.

6 Acknowledgments

Thanks Ms. Iannucci, my Independent Study Mentor, for all of your support throughout this process.

References

[1] Daniel M. Abrams and Steven H. Strogatz. "Modelling the dynamics of language death". In: *Nature* 424 (2003). DOI: 10.1038/424900a.

[2] Lindell Bromham et all. "Global predictors of language endangerment and the future of linguistic diversity". In: *Nat Ecol Evol* 6 (2022). DOI: 10.1038/s41559-021-01604-y.

[3] Albert-Laszlo Barabasi and Reka Albert. "Emergence of Scaling in Random Networks". In: *Science* 286.5439 (1999). DOI: 10.1126/science.286.5439.509.

[4] United Nations Department of Economic and Social Affairs Population Division. "World Population Prospects". In: (2022). URL: https://population.un.org/wpp/Download/Standard/MostUsed/.

[5] Jinyun Ke Tao Gong and William S-Y Wang. "Language Change and Social Networks". In: *Communications in Computational Physics* 3 (2008). URL: https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=b92e0ec73995acd2b049dbc3f7589f3b3223ebcf.

[6] William S-Y. Wang James W. Minett. "Modelling endangered languages: The effects of bilingualism and social structure". In: *Lingua* 118 (2008). DOI: 10.1016/j.lingua.2007.04.001.

[7] Joshua B. Tenenbaum Joshua K. Hartshorne and Steven Pinker. "A critical period for second language acquisition: Evidence from 2/3 million English speakers". In: *Cognition* 177 (2018). DOI: 10.1016/j.cognition.2018.04.007.

[8] Xiang Li and Guanrong Chen. "A local-world evolving network model". In: *Physica A* 328 (2003). DOI: 10.1016/S0378-4371(03)00604-6.

[9] Jorge Mira and Angel Paredes. "Interlinguistic similarity and language death dynamics". In: (2005). DOI: 10.48550/arXiv.physics/0501097.

[10] Duncan J. Watts and Steven Strogatz. "Collective dynamics of 'small-world' networks". In: *Nature* 393 (1998). DOI: 10.1038/30918.

[11] Victor M. Eguiluz Xavier Castello and Maxi San Miguel. "MODELLING LANGUAGE COMPETITION: BILINGUALISM AND COMPLEX SOCIAL NETWORKS". In: *International Conference on Language evolution* EVOLANG7 (2008). DOI: 10.1142/9789812776129_0008.

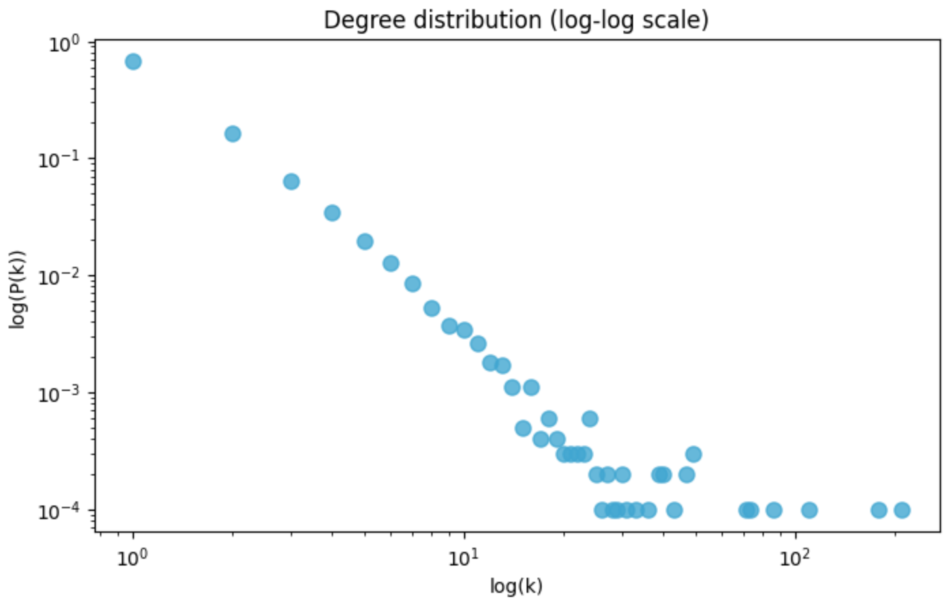

Figure 1: The degree distribution of the Barabasi-Albert Graph, y = 3.

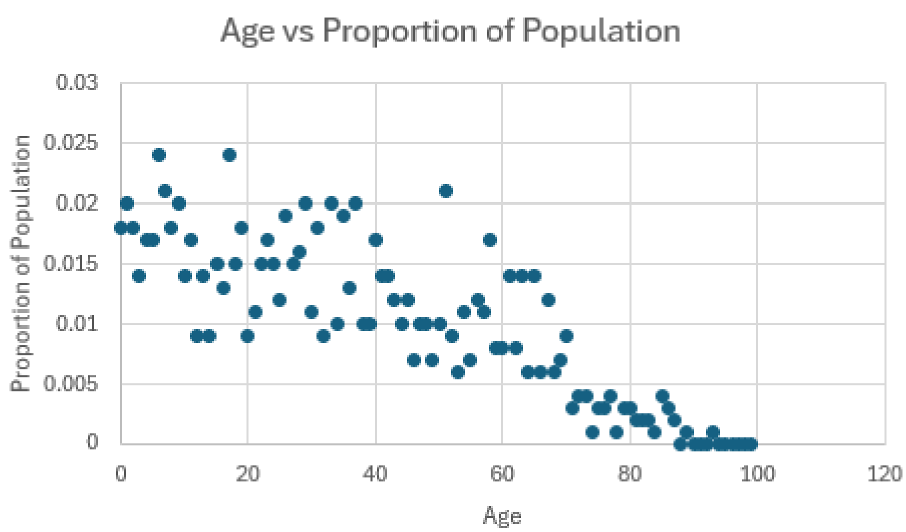

Figure 2: The Proportion of the Population by each age group.

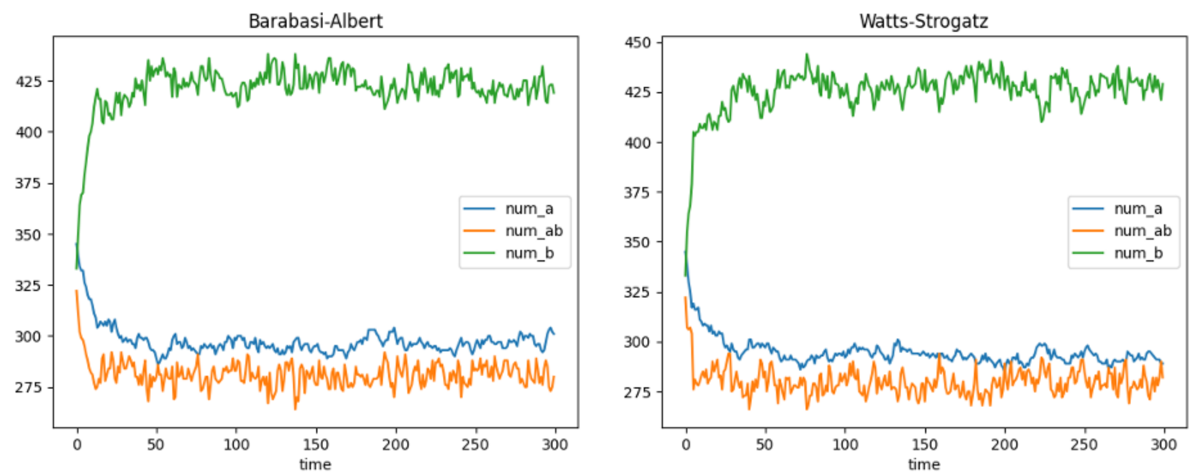

Figure 3: Experiment 1: The number of people in each state over time.

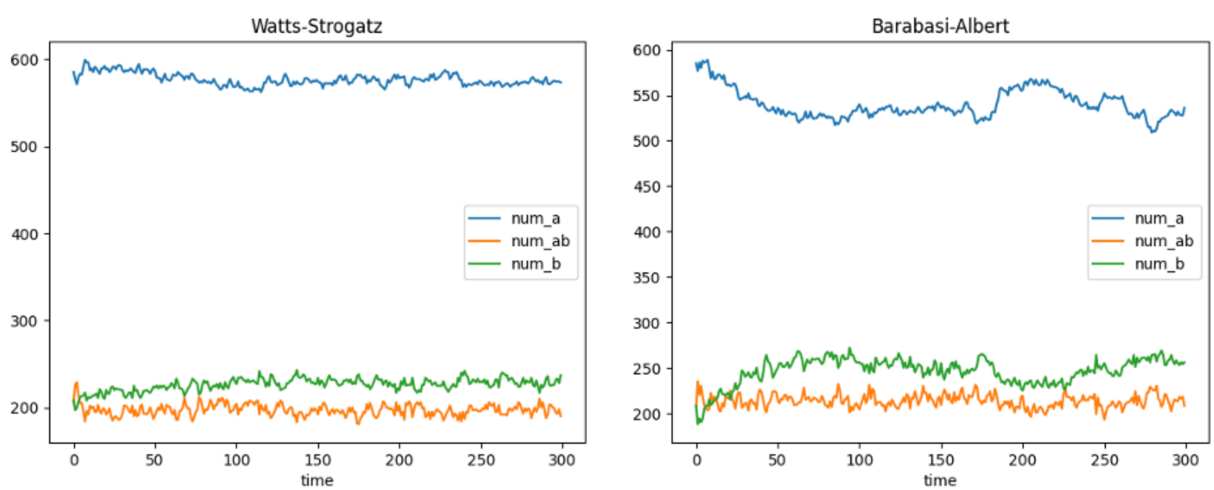

Figure 4: Experiment 2: The number of people in each state over time.

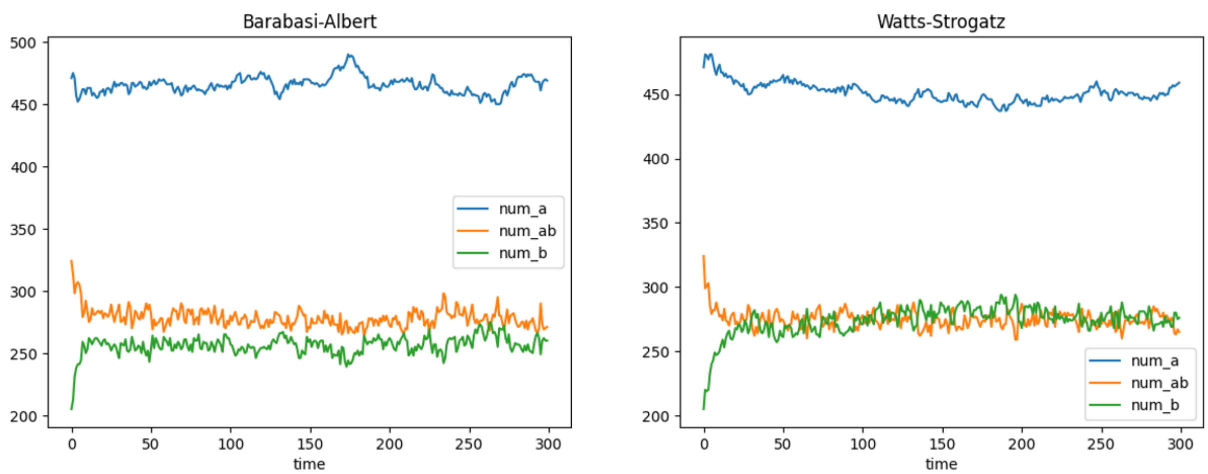

Figure 5: Experiment 3: The number of people in each state over time.

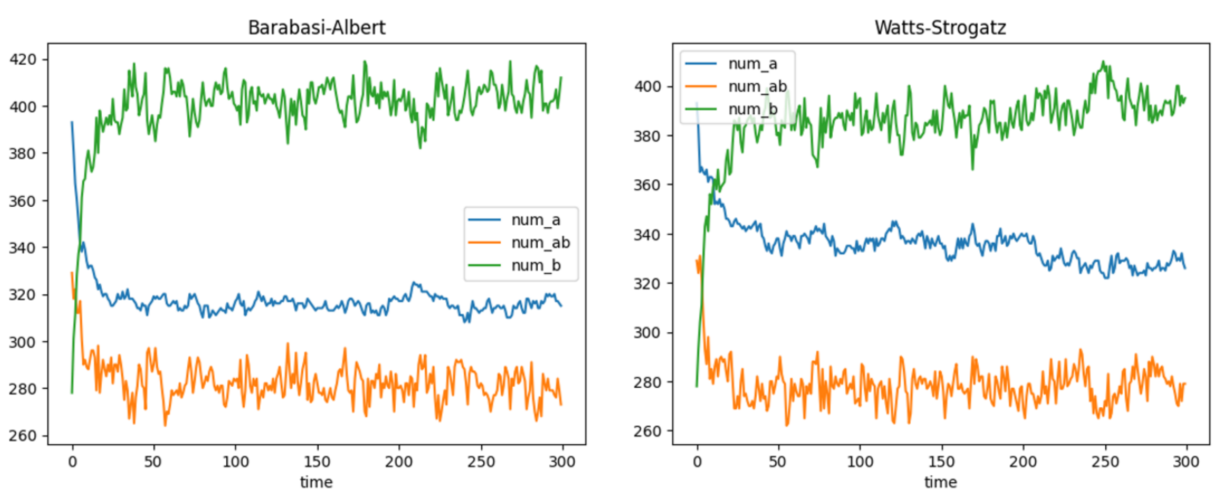

Figure 6: Experiment 4: The number of people in each state over time.