Simulating Lexical Convergence of Artificial Languages in Populations of Agents Using Evolutionary Algorithms

Gregory Furman
University of Cape Town
Cape Town, South Africa
frmgre001@myuct.ac.za

Dr. Geoff Nitschke
University of Cape Town
Cape Town, South Africa
geoff.nitschke@uct.ac.za

ABSTRACT

The origins and evolution of natural language is notoriously difficult to approach scientifically, namely due to lack of historical linguistic data. This paper contributes to computational methods for increasing our understanding of language evolution. Specifically, we investigate how speakers of different virtual languages eventually converge (agree on) a single common language, given varying degrees of environmental pressure for communication to take place. Experiments were executed using an evolutionary *Naming Game* task environment to test under which conditions a group of *N* language-speaking agents would be able to maximise consumption of food resources by minimising their linguistic differences with other agents. Results indicated that the most spoken language at the beginning of evolution has the largest influence on the converged-upon (common) lexicon at the end of evolution. Furthermore, changes in the rules of the *Naming Game* was found to strongly affect lexicon convergence, while changes to each agent's *talking radius* (field of view) produced negligible differences. Future work will repeat the experiments with different evolutionary linguistic models, language games and population sizes.

1 INTRODUCTION

The ability to conduct experiments and garner data on the origins and evolution of languages is notoriously difficult to approach in a rigorous scientific and empirical manner. This can be largely attributed to the insufficient amount of data required to draw acceptable and statistically sound conclusions[10]. Over the past few years, simulations of artificial language evolution have been adopted in order to investigate simulated agents' ability to self-organise languages with natural-like properties [2].

An artificial language is a language of a (typically) limited size which emerges in computer simulations between artificial agents, robot interactions or controlled psychological experiments with humans. A common or derived lexicon refers to some sort of agreed upon communication system between agents.

While significant work has been done on artificial language evolution and as well as the effect of human population size on language evolution and change, this study is geared towards the intersection of both fields of study [3] [10]. This paper's research applies an *Evolutionary Algorithm* to an *agent-based model* of *artificial language evolution simulation* in which a *common lexicon* must be *converged upon* in order for *agents* to *communicate*, and therefore gain *fitness*.

The Talking Heads experiment [11] placed embodied agents in both physical and simulated environments and observed how a new shared vocabulary was derived by means of playing a type of *Naming Game*. A *Naming Game* is defined as an interaction among *N agents* where the goal is to agree on some common lexical term for a food resource (see Section 4.1). Agents were evaluated at a *Naming Game* task and fitness was a function of their performance at a *Naming Game*. Within the field of Artificial Life (ALife) as well as Artificial Language Evolution, it is commonly agreed upon to view language as a complex adaptive system [5]. This defines language as a system consisting of multiple agents interacting with one another which is adaptive such that an agent's behavior is largely dependent on their past interactions which in turn have an influence on future behaviour [7].

With such context in mind, this project aims to investigate the effects of population size on rule-based agent ability to converge upon a common or agreed upon lexicon. In order to test this, agents were placed in a virtual environment and made to move around and communicate over resources to be consumed. Communication will occur in the form of a modified *Naming Game* [11] [5]. Such communication constitutes agents interacting with one another resulting in the transfer of linguistic units between agents' languages. An agent's language in the context of this study was represented as the lexicon or artificial language of an agent. After each agent moves around the environment K times, the least fit 20% will be culled with agents that have consumed the most resources undergoing *multi-point crossover* [6]. Crossover will occur until the original population size of 150 has been replenished. This is referred to as a steady state population [12]. Fitness, in this case, will be defined by the amount of food resources an agent consumes. More specifically, the resource payouts they receive from consuming said resource which will vary randomly between and including *1* and *10*.

2 BACKGROUND

This work will be borrowing concepts outlined by Nitschke et. al [5] such as the maximising of food consumed by agents using an evolutionary algorithm as well as the usage of a *Similarity Threshold*, which is discussed in greater detail in section 5.6. Maravall et. al [8] showed that using evolutionary algorithms, an autonomous population of agents was able to converge on a common lexicon[1]. This study differs from the research at hand as it attempts to evolve vocabulary with mapped meaning into the convergent lexicon. Alternatively, this paper aims to focus on the effects of population size on an emergent language when agents are incentivised to cooperate. Another study aimed to simulate grammatical agreement on very basic morphemes [2]. The results showed that agreement systems develop to minimize semantic ambiguity in different cultural environments. While similarities exist, this study does not aim to evolve a common language but instead attempts to find

[1] Here, *convergence on a common lexicon* is equivalent to *N* agents agreeing (within a small margin of error) on a single set of *words* as their overall shared language.

under what circumstances agents in a system agree on semantic variation given certain cultural factors (such as the 'gender' of the agent). Bromham et. al [3] demonstrated that a greater population size leads to a higher rate of gaining new words where as smaller populations experience higher rates of word loss. The results found in this study show the effect of demographic factors on rates of language evolution and that rates of word gain and loss can be a function of a particular demographic size.

3 RESEARCH GOALS AND HYPOTHESES

The overall goal of this research is to better understand how speakers of *N* different languages eventually agree on a single shared language when there is environmental pressure for communication to take place over evolutionary time.

We are specifically interested in the role of *population distribution* during evolution. Namely, we hypothesise that if *Language X* has more speakers in the population than every other language, then the final agreed-upon language will be more lexically similar to *Language X* than any other language.

Furthermore, we will be assessing the extent to which factors such as Naming Games between N agents as well as an agent's field of view (referred to as a *Talking Radius*) will affect the convergence of a common lexicon.

An agent type that has a greater population size (i.e. is a speaker of the most commonly spoken language) than all other agent types at the beginning of a simulation will be referred as the majority agent with the those agents that are smaller in number being referred to as the minority agents.

3.1 Language Change

- ***H0:*** *There will be no statistically significant difference present between the average proportionality of a particular language at the beginning of a simulation versus after an evolutionary process has completed.*
- ***HA:*** *There will be a statistically significant difference between the starting proportions and the proportions found after an evolutionary process has completed.*

3.2 Do The Languages Equally Converge

- ***H0:*** *There is no statistically significant difference between the average average proportionality of language types at the end of the evolutionary process.*
- ***HA:*** *There exists a statistically significant difference between the average proportionality of each language type at the end of the evolutionary process.*

4 METHODS

4.1 Naming Game

A Naming Game commences when an agent moves adjacent to a grid cell containing a resource. The agent will then view all agents in a Talking Radius R and engage with either 1 or N agents in the given radius, depending on the game type.

The modified Naming Game would be played between agents as follows:

(1) Speaker: speaks candidate word for resource to hearer comprised of letters from its own lexicon.
(2) Hearer: checks how similar word is. If not similar enough Hearer will mutate the word until it is deemed similar. Each mutation costs 1 resource point.
(3) Speaker and Hearer switch roles as the Hearer speaks the modified word back to the old Speaker.
(4) If old Speaker (now Hearer) accepts: both agents add any letters in agreed upon word to their lexicons that are not currently present and consume the resource with a payout proportional to each agents' lexical influence on the new word.
(5) Else repeat from step 1

4.2 Artificial Life Simulation

Figure 6 shows a graphical representation of the virtual environment.At the beginning of each generation, agents are assigned a default fitness value F. Each agent then moves about a *QxQ* grid M number of times. If an agent appears adjacent to a resource item, every agent in a radius R surrounding it will be added to a queue. An agent will then either engage in a Naming Game with one or N agents within the queue. The conclusion of a Naming Game will lead to both agents being gifted a percentage of fitness or resource points as well as the resource being consumed. If every agent has not attempted to move 100 times, repeat steps 1-5 else progress to evaluating fitness.

4.3 Evolutionary Algorithm (EA)

In the case of this experiment, the an fitness is defined by how many resource points agent is able to garner in a given generation.

(1) ***Fitness Function:*** Agent fitness was calculated as the total value of resources an agent was able to consume during evaluation.

The process by which fitness is evaluated is as follows:

(1) Once all *N* agent movements cease, agents are ranked from fittest to least fittest.
(2) The least fittest 20% of agents in the system are culled. The population size is now *0.8N*
(3) The fittest 20% of agents are then selected and are randomly assigned a breeding partner. These agents are classified as the parent agents.
(4) Breeding pairs then generate two children as offspring. Each child agent's *genotype* is a combination of its two parents created by means of *multi-point crossover* and a *mutation rate* of 0.05.
(5) Breeding will occur until the population reaches its original size of *N*

The EA is therefore structured with an explicit and an implicit goal. Explicitly, the EA aims to breed agents that have consumed more resources and had a greater proportion of their languages in said resource. Implicitly, the algorithm ultimately breeds agents that have a greater ability to communicate and propagate their own languages. Thus it will be assumed that languages bearing linguistic similarity that facilitate this will motivate the minimisation of linguistic differences.

5 EXPERIMENTAL SETUP

Upon the start of simulation, each agent in the system will be assigned a random *Similarity Threshold* between 0 and 1 and conclude once 1000 generations of fitness evaluation and breeding have taken place. An agent's genotype will be its primitive language which will be represented by an array of lexical units as well as a *Similarity Threshold*. 150 agents and 50 food items will be randomly distributed and initialised across a 40 x 40 environment.

A simulation begins with a controlled distribution of different language-speaking agents being placed in a simulated environment with a controlled number of randomly dispersed food items scattered amongst the environment. Agents in the system will play a modified version of The Naming Game which will subsequently lead to some common lexical term for a food item being agreed upon and the resource being consumed. The population used by the EA will be equal to the number of agents in the system - which will be steady state.

5.1 Agents Move

Each agent will move 1 unit about the environment in a Northern, Southern, Eastern or Western direction 100 times per generation. When an agent appears adjacent to a resource it will engage in a Naming Game with N agents within a given radius R (referred to as the Talking Radius). This will occur 100 times. This Naming Game is played between either 2 or N-agents.

5.2 Fitness

The term fitness is interchangeable with resource points in the context of this study. As each resource has a different payout of resource points, fitness refers to the total amount of resource points consumed.

5.3 Candidate Lexical Term

As agents engage in a Naming Game, any term for a resource that an agent has proposed but has not yet been agreed upon will henceforth be referred to as a candidate term. Refer to figure 1. This is otherwise referred to as a *word*.

5.4 Common Lexical Term/Agreed Upon Lexical Term

Once a candidate lexical term has been agreed upon by both agents, it will be referred to as the agreed upon or common lexical term.

5.5 Modification

A modification to a candidate lexical term or word refers to an agent adding a linguistic unit to the candidate term or exchanging a random unit from inside the candidate term that is not present within the agent's genotype. Each single modification an agent makes to a word will cause 1 resource point to be consumed. This modification process will occur until both agents' *Similarity Thresholds* are met or until an agent's fitness is less than or equal to a *Talking Threshold* which will be 10% of the original fitness an agent had before engaging in a Naming Game.

5.6 *Similarity Threshold*

A *Similarity Threshold* [5] dictates what proportion of a candidate lexical term for a resource should be comprised of an agent's own lexicon. Following the scenario illustrated in Figure 1, if the white agent had a *Similarity Threshold* of 0.5 the candidate lexical term offered by the blue agent would be accepted as the proportion of A's lexicon is 0.67 or 2 out of the 3 morphemes.

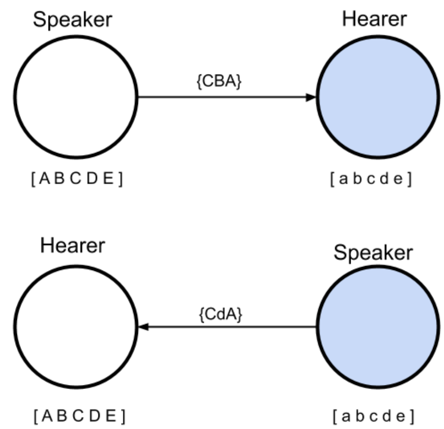

Figure 1: Illustration of Naming Game between 2 Agents

5.7 Talking Threshold

The Talking Threshold refers to the point at which an agent has lost too much fitness during the modification process of a *Naming Game* and will therefore stop modifying the candidate term, if it is in the process of doing so, and accept any candidate term that another agent presents.

5.8 Environment Configuration

The simulated environment is a (40 x 40) two-dimensional discrete grid of cells which contains a set of *agents* and *resources*. A given cell either be empty or be occupied by a single agent or resource at any timestep throughout the simulation. As a generation begins, *resources* and *agents* are dispersed throughout the grid at random coordinates. A graphical representation of the environmental configuration can be seen in Figure 6.

Table 1: Agent and Simulator Parameters

Language Evolution Parameters	
Generations per experiment	1000
Population size	150
Language Mutation Rate	0.05
Similarity Threshold Mutation Rate	Gaussian mutation $(-0.5, 0.5)$
Simulation Parameters	
Total agent movements per generation	15 000
Maximum agent movement per timestep	1
Initial agent / resource positions	Random
Environment width x height	40 x 40 cells
Resource payout	Uniformly distributed between [1,10]
Talking Radius	0.25 / 0.50 / 0.75 / 1 of total grid area
Naming Game	2-Agent (similar) / 2-Agent (random) / N-Agent
Amount of Agents of type A,B and C	A:B:C = 50:50:50 / 100:25:25 / 25:100:25 / 25:25:100

6 RESULTS

6.1 Hypothesis Testing on Equal Language Convergence

For each population ratio of agent types an *F-Test* [1] was conducted in order to test whether or not there exists a statistically significant difference between the average proportions of languages' A, B and C in the average lexicon of an *agent* following the completion of an evolutionary process. Following statistical convention, a confidence level of 0.05 was used. All results can be viewed in Table 3.

- ***Where A:B:C = 50:50:50*** *When the amount of type A, B and C agents are set to be equal in population size at the beginning of a simulation, there is insufficient evidence required to reject H0. It can therefore be concluded that the differences in the proportions of artificial languages A, B and C are statistically insignificant and that there is no statistically significant difference between the average proportionality of language types at the end of the evolutionary process.*

- ***Where A:B:C ≠ 50:50:50*** *Where Population proportion was in a ratio not equivalent to 1:1:1 or where one specific type of agent held the dominant proportion of the total population size at the beginning of a simulation, the results were as follows: There is extremely strong evidence in favour of HA. We therefore reject H0 in favour of HA at the 1% level and conclude that there exists a statistically significant difference between the average proportionality of each language type at the end of the evolutionary process where a Population proportion has one agent type in the majority.*

6.2 Hypothesis Test on Language Change

A one sample *t-test* [4] was used in order to test whether or not there exists a significant difference between between the proportionality of a language type at the beginning of a simulation versus after an evolutionary process had concluded. A one-sample test is used as the proportion of a language type X in the average agent's lexicon at the beginning of a simulation is equal to the proportion of type

	A = B = C	
Population proportions	p-value	conclusion
50:50:50	0.2039	Fail to reject H0
100:25:25	2.32E-05	Reject H0
25:100:25	5.18E-05	Reject H0
25:25:100	4.18E-05	Reject H0

Table 2: The resulting p-values and subsequent conclusions at the 5% level of Experiment 1: Equal Language Convergence.

X agents in a system. This can be seen in the following equation:

$$Proportion_X = \frac{Number\ of\ X\ Agents}{Total\ Agents}$$

All resulting *p-values* can be observed in Table 2.

- ***Where A:B:C = 50:50:50*** *Where Population proportions is equal to 50:50:50 there is insufficient evidence in favour of H0. Therefore we conclude that no statistically significant difference exists between the average average proportionality of language types at the beginning versus after the evolutionary process has completed and that the starting proportions of languages A, B and C are equal to the proportions of A,B and C in the average lexicon of an agent after an evolutionary process has completed. .*

- ***Where A:B:C ≠ 50:50:50*** *In every case of the Population proportions not being equivalent to 1:1:1, there is overwhelming evidence against H0 causing us to reject H0 in favour of HA. We therefore conclude that there exists a statistically significant difference between the proportions of languages A,B and C at the beginning of a simulation versus after an evolutionary process has taken place.*

	A start vs A end		B start vs B end		C start vs C end	
Population proportions	p-value	conclusion	p-value	conclusion	p-value	conclusion
50:50:50	0.3	Fail to reject H0	0.31	Fail to reject H0	0.408	Fail to reject H0
100:25:25	3.39E-06	Reject H0	2.66E-06	Reject H0	5.02E-06	Reject H0
25:100:25	3.71E-05	Reject H0	1.26E-05	Reject H0	6.08E-06	Reject H0
25:25:100	1.18E-05	Reject H0	3.01E-05	Reject H0	1.82E-05	Reject H0

Table 3: Table of resulting p-values and subsequent conclusions at the 5% level of Experiment 1.

6.3 Impact of Naming Games on Results

A modified *Naming Game* can be one of 3 different types with only one type of *Naming Game* to be played in a given simulation. An agent will engage in a *Naming Game* with 1 or N agents within a given radius. A *2-Agent Naming Game* can be one of 2 types - an agent is either selected at random from within a given radius else the agent with the highest lexical similarity to the agent that first encountered the resource is selected. Lexical similarity, in the context of this study, is defined by how many common lexical units one agent shares with another. Therefore, the more units shared - the more similar the agents are determined to be. An N-Agent Naming Game occurs when an agent engages in a Naming Game with all agents within a specified radius.

The total spreads of the 2-Agent Games (similar), 2-Agent Games (random) and N-Agent Games were *0.1918, 0.0329* and *0.0043* respectively. Spread, in the context of this study will be the inter-quartile range. Figure 1 displays all resulting box plots of each *Language Type's Average Proportion.*

The ideal *Average Proportion* of each *Language Type* is 0.33. This value indicates all languages types have an equal proportion in a derived upon or common lexicon. The non-parametric Kruskal-Wallis Rank Sum Test [9] will be used in order to check whether the medians of the Average Proportions of each Language Type are equal.

The results of 2-Agent Games (similar) show the largest spread of Average Proportion for all Language Types. While the median (0.2644) Average Proportion was deemed to be less than the ideal proportion of 0.33, the data is skewed in the positive direction. All medians were deemed equal (*p-value = 0.6312*).

The spread of 2-Agent Games (random) for all Language Types is far less than that of the 2-Agent Games (similar). The Average Proportions of Language Types A, B and C appear to skewed somewhat positively with all medians having no statistically significant differences (*p-value = 0.9585*). The total median of the 2-Agent Games (random) *Average Proportions* was equal to 0.3244.

The smallest spread was found in the Average Proportion of each Language Type when the N-Agent Games were played. Each Language Type's Average Proportion can be seen to be skewed negatively or in the direction of 0.33. The majority of all data can be seen to lie between 0.33 and 0.336 in each Language Type. The difference in all medians between Language Types A, B and C are deemed to be statistically insignificant (*p-value = 0.4545*) with the total median of all all *Average Proportions* across the N-Agent Games being equal to 0.3340.

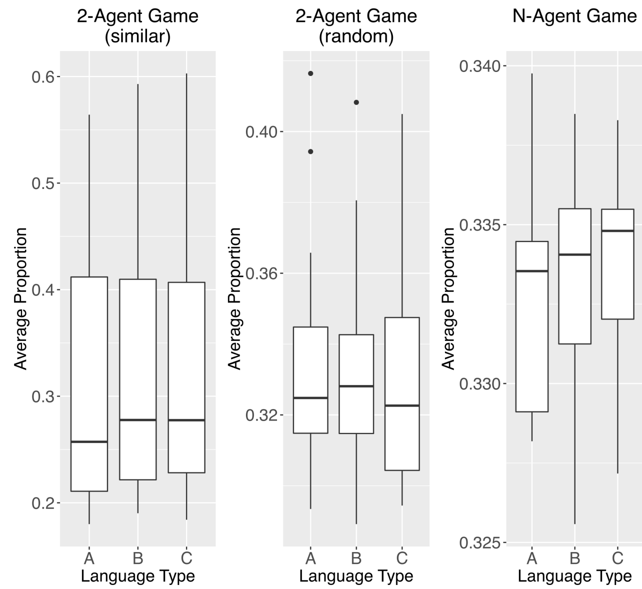

Figure 2: Box plots of the Average Proportion of Language Types A, B and C per Naming Game following the completion of an evolutionary process.

6.4 Impact of Radius on Results

The total area of the environment is *40x40* or 1600 *cells* with a *Talking Radius* ranging from *1/4, 1/2, 3/4 or 1* of the entire grid area. As was discussed in the *Methods* section, the *Talking Radius* is the area of the grid an agent will scan for potential partners in which they can play a *Naming Game* with. Once again, the Kruskal-Wallis Test is used in order to test whether the differences in the median *Average Proportion* of *Language Types* had a statistically significant difference [9]. The total medians of *Average Proportion* where *Talking Radii* are equal to *1/4, 1/2, 3/4 or 1* are *0.3308, 0.3314, 0.3309* and *0.3322* respectively. All subsequent results seen can be viewed in Figure 3.

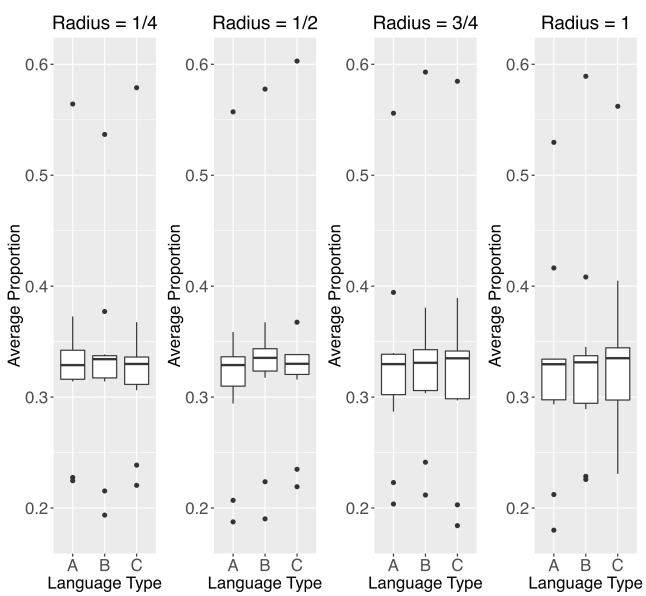

Figure 3: Box plots of the Average Proportion of Language Types A, B and C per Talking Radius following the completion of an evolutionary process

Where a *Talking Radius* (referred to as *radius*) was *1/4* of the whole grid, the differences between the medians where concluded to be statistically insignificant (*p-value = 0.9277*). The total spread of the Average Proportion where *radius* is equal to *1/4* is 0.0224. As the *radius* increases to 1/2 the total spread of all Average Proportions is equal to *0.0213*. It was concluded that all Language Types had no statistically significant difference in their median Average Proportions (*p-value = 0.6489*). We see an increase in the total spread of Average Proportions as the *radius* increased to *3/4* with total spread now being equal to *0.0359*. All differences in medians between Language Types were deemed statistically insignificant (*p-value = 0.9467*).The total spread of the Average Proportion further increased as *radius* increases to 1 with total spread now being equal to *0.041*. All Average Proportions of each Language Group are evidently skewed in a direction less than the median and closer to 0.3. Once again, there is no statistically significant difference between the medians of Language Types A, B and C (*p-value = 0.5468*).

6.5 Effects of Evolutionary Process on Similarity Threshold

The *Similarity Threshold* of a child agent will be bounded by the *Similarity Threshold*s of its parents. If the difference in *Similarity Thresholds'* between 2 parent agents is *0*, then both children will inherit either one of the parents *Similarity Threshold*. If one parent has a larger *Similarity Threshold* than the other, both children will be assigned a *Threshold* equal to a random value between its parents' *Thresholds*. There is a 5% chance of a *Similarity Threshold* having a *Gaussian Constant* between *[-0.5,0.5]* added.

The resulting Figure 5 displays the average *Similarity Threshold* of every agent in a system per generation averaged out across all 48 differing environments and 50 runs per environment. At *Generation 0* the average *Similarity Threshold* is equal to 0.5. This is to be expected as a the *Similarity Thresholds* are uniformly randomly assigned as values between *0.1 and 0.9*. The average would therefore be *0.5*. As the *Generation* increases, there is a rapid decline to a threshold lower than the initial amount. This trend occurs up until *Generation 15* in which the *Similarity Threshold* begins to follow a logarithmic growth pattern plateauing at an average *Similarity Threshold* of *0.62* at around *Generation 650*.

7 DISCUSSIONS OF RESULTS

7.1 Hypothesis 1: Equal Language convergence

The results showed that at the end of every Evolutionary Process the proportion of each original Language Type in the average converged upon lexicon was affected by the initial amount of type A, B and C agents in a simulation.

When the initial amount of agent types were set to be equal in number at the start of a simulation, the proportions of *Language Type* A, B and C had statistically insignificant differences. This is likely due to the probability of engaging in a *Naming Game* with an agent of any type being equal, no matter the size of the *Talking Radius*.

When one type of agent holds the majority proportion upon a simulation beginning, it would only be natural to assume the likelihood of engaging in a *Naming Game* with a majority agent type will be far higher than engaging with an agent in the minority. This is also reflected in the results as the differences in the proportions of Language Types A, B and C following the conclusion of an Evolutionary Process are said to have statistically significant differences.

7.2 Hypothesis 2: Language Change

We can define the Language Change as the extent to which a particular language type will change within the average agent's lexicon throughout the simulation. This change can be measured by means of comparing the proportion of a language at the beginning of a simulation versus the same proportion of said language following the conclusion of an *Evolutionary Process*.

Given an equal ratio of type A, B and C agents upon a simulation beginning, the Language Change of languages A, B and C are concluded to be statistically insignificant. This implies that the proportion of each language type in the average agent's lexicon have not undergone any sort of substantial change. Once again, this is thought to be a direct result of the equal amount of each agent type. The probability of an agent of type A, B or C being present in a given radius and engaging in a language game is not shifted in any particular agent types favour as there is an equal representation of each agent type in the system.

This reasoning is supported by the resulting *Language Change* observed given an initial population of agents not having equal proportions of type A, B and C agents. It was concluded that the differences in the starting proportions of a language type and those proportions observed following the completion of an *Evolutionary Process* were statistically significant. This held true for every single case of agents not being equal in population proportion.

A general observed trend emerged where a language proportion is being measured each generation for a population with a clear majority agent type. The agent type in the majority will always decrease in *Average Proportion* as Generations progress. This is contrasted by the average language proportion of a minority agent which will almost certainly increase in size. This can be seen in Figure 4.

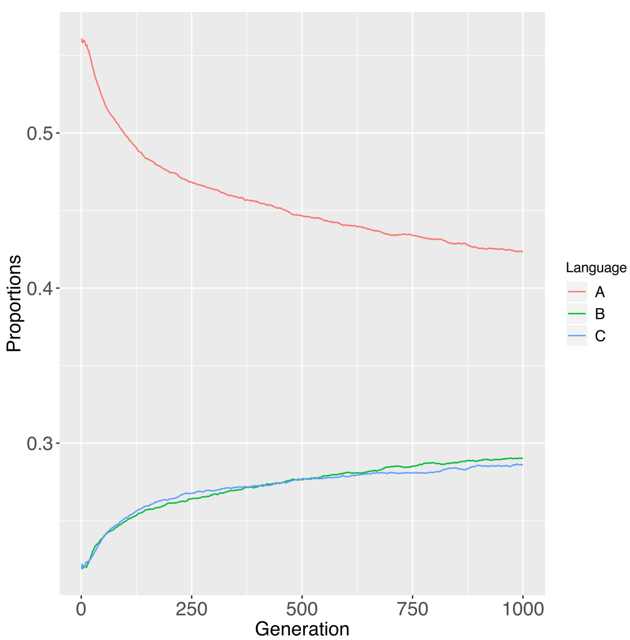

Figure 4: The average proportions of all simulations where A:B:C = 100:25:25

7.3 Naming Games

2-Agent Games (similar) yielded a median smaller than the ideal 0.33 and with the largest spread of all *Average Proportions* across all *Naming Games*. This can be attributed to the lack of diversity such *Naming Games* would fail to promote amongst agents. By only engaging with agents in a given radius that are deemed most similar, the amount of new linguistic information passed amongst agents will be severely limited. The average fitness would be positively affected, as there are far less modifications made to candidate lexical terms for blocks. The average fitness of each agent in the 2-Agent Games (similar) is said to be greater than the average fitness of agents in the 2-Agent Games (random) (*p-value < 2.2e-16*).

The 2-Agent Games (random) outperformed the *2-Agent Games (similar)* with respect to their medians and spread with the random games having a median closer to 0.33 (implying equal convergence) and a spread that was smaller showing less variation in results. The ability for agents to better converge on a common lexicon when a random agent is selected within a given radius as opposed to the most similar is due to the greater passing of linguistic information. Since an agent has no bias regarding the genotype of the agent it is engaging with, there is a greater opportunity for the diversification of the average agent's lexicon.

When comparing the results of both 2-Agent Games to those found in the N-Agent Games, it is clear that the N-Agent Games allowed for the best convergence. The total spread was far lower than those found in both 2-Agent Games and the median of all *Average Proportions* was also much closer to the ideal 0.33. Although more computationally taxing, the ability for an agent to communicate with all agents in a given radius will undoubtedly lead to the greatest transfer of linguistic information. The agent communicating with the N agents in a given radius also acts as a medium for information to transfer between all agents in a given radius. For example: Agents Q, X, Y and Z are in a radius R and agent Q encountered a resource thus starting an N-Agent Game. If agent Q communicates with X, Y and Z (in that order), any information that agent Q gains from agent X has a chance of being transferred to agent Y and subsequently agent Z.

7.4 Radii

The *Talking Radius's* effects on lexicon convergence appear to be negligible. The difference in medians are statistically insignificant (*p-value = 0.4394*). The difference spread between *Radius = 1/4* and *Radius = 1/2* do not seem to differ by a significant amount. But as *radius* increases in size, as does the spread. Such a trend is likely due to a greater *Talking Radius* allowing for one particular language type to spread amongst all agents with far greater ease.

If an agent of type X is a majority agent, the majority of all Naming Games played will be done so by an agent of type X. Having a greater *Talking Radius* would increase the spread of a majority agent type's language and would thus cause such an agent to dominate the average agent's lexicon with greater ease. Such a trend should be investigated further in future research.

7.5 Similarity Threshold

An interesting trend emerges when observing the average *Similarity Threshold* of every agent as generations progress, as seen in Figure 5. There appears to be a stage in which the average *Similarity Threshold* of all agents undergoes a decrease. We theorise this to be due to selective pressure which forces an agent to lower their *Similarity Threshold* until a certain amount of language transfer has occurred. In other words, the lower a *Similarity Threshold*, the less of its own language it requires in order to agree upon a word in a Naming Game. Therefore, in the first few generations, before a large amount of language transfer between agents has occurred, the agents that are more willing to accept a word with a smaller proportion of their lexicon are selected for as they do not yet have a diverse enough language.

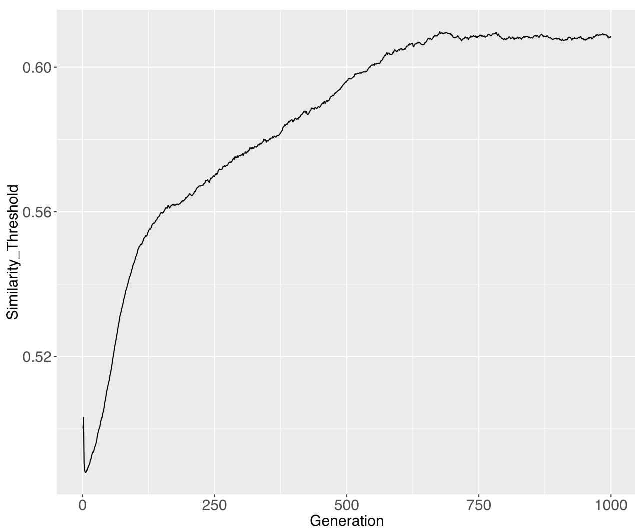

Figure 5: Average *Similarity Threshold* of all agents in a system per Generation averaged over 48 different simulation iterations with 50 runs per iteration.

8 CONCLUSIONS

Results indicated that languages with many speakers have a greater influence in the average agent's lexicon (at the end of evolution) than languages with fewer speakers.

Furthermore, it was found that if every language had the same number of speakers in generation 0, then the final evolved lexicon would be equally constituted of each language in the environment. For example, if languages X and Y were the only two languages, each having the same number N speakers in the beginning, then the final derived lexicon would be 50% constituted of lexicon from language X and 50% from language Y.

Within the first few generations, the average *Similarity Threshold* of each agent was seen to rapidly decrease to a global minimum after a subsequent increased until plateauing at 0.62 was observed. We theorise that such a trend can be attributed to selective pressure favouring agents with a lower *Similarity Threshold* until enough linguistic information was transferred. After which, languages were deemed similar enough to allow for a greater *Similarity Threshold* on average.

Future research will repeat the experiments with different evolutionary linguistic models, language games and population sizes. Additionally we plan to implement a multi-objective evolutionary algorithm as a meta-heuristic to co-optimise maximal agent fitness and minimal linguistic differences.

9 ACKNOWLEDGEMENTS

The author would like to thank Dr. Geoff Nitschke for his supervision and support throughout the project. I would also like to thank Prof. Rajend Mesthrie and Prof. Markus Arnold for their assistance with the philosophy and linguistics of this research.

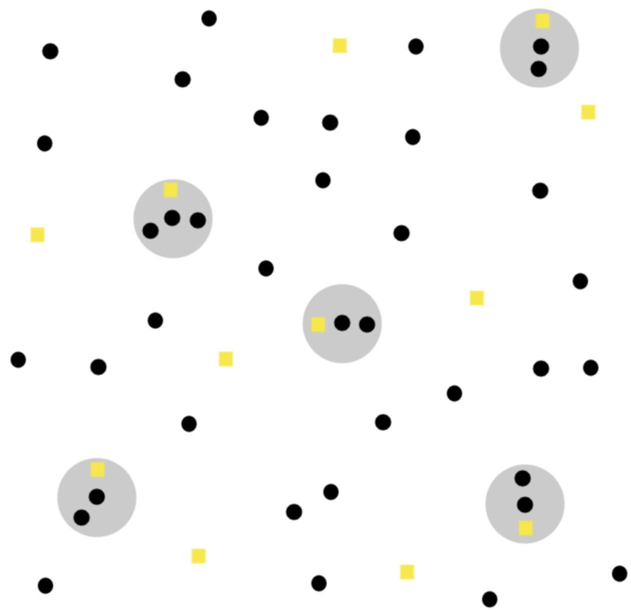

Figure 6: Evolutionary Task Environment with N agents. Food Items are yellow squares, agents are black circles and *Talking Radii* (field of view) are grey circles. Each agent 'speaks' one of 3 different languages and seeks to maximise food consumption. Agents within view of one another engage in *Naming Games*, whereby they attempt to agree (converge) on a common term for a food item by repeatedly mutating and comparing their language's term for the food item. Each mutation costs energy, such that failure to agree on a term results in starvation, while agreement results in food consumption.

REFERENCES

[1] Arne Bathke. 2004. The ANOVA F test can still be used in some balanced designs with unequal variances and nonnormal data. *Journal of Statistical Planning and Inference* 126, 2 (2004), 413–422.
[2] Katrien Beuls and Luc Steels. 2013. Agent-based models of strategies for the emergence and evolution of grammatical agreement. *PloS one* 8, 3 (2013), e58960.
[3] Lindell Bromham, Xia Hua, Thomas G Fitzpatrick, and Simon J Greenhill. 2015. Rate of language evolution is affected by population size. *Proceedings of the National Academy of Sciences* 112, 7 (2015), 2097–2102.
[4] NAC Cressie and HJ Whitford. 1986. How to use the two sample t-test. *Biometrical Journal* 28, 2 (1986), 131–148.

[5] Kobus de Bruyn, Geoff Nitschke, and Willem van Heerden. 2011. Evolutionary algorithms and Particle Swarm Optimization for artificial language evolution. In *2011 IEEE Congress of Evolutionary Computation (CEC)*. IEEE, 2701–2708.
[6] Susana C Esquivel, HA Leiva, and Raul H Gallard. 1999. Multiple crossovers between multiple parents to improve search in evolutionary algorithms. In *Proceedings of the 1999 Congress on Evolutionary Computation-CEC99 (Cat. No. 99TH8406)*, Vol. 2. IEEE, 1589–1594.
[7] "Five Graces Group", Clay Beckner, Richard Blythe, Joan Bybee, Morten H Christiansen, William Croft, Nick C Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, et al. 2009. Language is a complex adaptive system: Position paper. *Language learning* 59 (2009), 1–26.
[8] Darío Maravall, Javier de Lope, and Raúl Domínguez. 2012. Self-emergence of a common lexicon by evolution in teams of autonomous agents. *Neurocomputing* 75, 1 (2012), 106–114.
[9] Patrick E McKight and Julius Najab. 2010. Kruskal-wallis test. *The corsini encyclopedia of psychology* (2010), 1–1.
[10] Luc Steels. 2006. How to do experiments in artificial language evolution and why. In *The Evolution Of Language*. World Scientific, 323–332.
[11] Luc Steels. 2015. *The Talking Heads experiment: Origins of words and meanings*. Vol. 1. Language Science Press.
[12] Gilbert Syswerda. 1991. A study of reproduction in generational and steady-state genetic algorithms. In *Foundations of genetic algorithms*. Vol. 1. Elsevier, 94–101.