Effect of Memory in Emergence of Vocabulary

Galip Ümit Yolcu
Computer Engineering
2016400249
gumityolcu@gmail.com

***Abstract*—Language is thought of as one of the main discriminating traits of humans compared to other animals. In the context of complex systems, language has been studied in two different ways: one corresponding to the evolutionary view, treating language as a way of increasing one's chances of survival and reproduction. The other view focuses on the sociocultural perspective and studies the emergence of language in a population, rather than across generations. Basics of studies in these areas are presented. With inspirations from these models, a new model based on adding memory to agents is explained and results are discussed.**

***Index Terms*—Language dynamics, complex systems, emergence of vocabulary, naming game, evolutionary language game**

I. Introduction

Language is certainly the most important ingredient in human communication and consequently, cooperation. Even though animals too use some degree of language, it is nowhere near humans in terms of complexity and flexibility. This raises the question of determining the mechanisms that enabled the emergence of this complex structure with infinite communication ability, and also the mechanisms to pass it through generations. This latter question seems to be more problematic because of a phenomenon called *poverty of stimulus* or *paradox of language acquisition*. It is the issue that during the language acquisition phase, children are note exposed to enough language related data to acquire the intricate details and nuances of language. Futhermore, they are exposed to very noisy data from the environment but they seem to be able to seperate the language related data from others and learn the language with far less than enough data. This pushes some linguists to think that language has primarily biological roots. This biological root, "the genetic component for language, whatever it is", is called the Universal Grammar.[1]

Moreover, language can be seen as a game. This idea is often attributed to the philosopher Ludwig Wittgenstein. The basic idea is that the relations between linguistic elements and elements of the real world are somewhat arbitrary. That is to say, for example, that words' meanings are determined by their use in different instances of language games.

In order to model the emergence of language, some simplifications have to be made. The most basic language we can think of is a set of signs or signals that are associated with certain meanings. It is reasonable to hypothesize that the roots of language have been established this way. Thus, in the models to be explained, we will focus on the emergence of a certain association of meanings and signals(words). This means that we will be modeling the emergence of vocabulary. The meanings can be seen as objects in the world and signals as different words, corresponding to different sounds or signals some hypothetical primates can generate.

There are two main approaches to this problem: sociocultural(semiotic) and sociobiological(evolutionary) approach.[2] While the sociocultural approach concentrates on the emergence of common vocabulary and self organization of language in the population during the lifetime of an individual, evolutionary approaches focus on how the linguistic knowledge is transmitted and how the language evolves while being transmitted to new generations.

In this paper, we will first explain two fundamental models from these approaches. Then we will briefly mention some models that try to combine features from both views. We will then present a new model that resembles the sociocultural approach, but can be analysed using the mathematical framework of the sociobiological approach and can be used to initate evolutionary dynamics. We will then analyse the results of the model and conclude.

II. Background

A. Naming Game

The sociocultural approach treats language as a self organizing complex system. Each agent organizes her own vocabulary after interactions with other agents and a global vocabulary for successful communication emerges.

Although there are other models for different scenarios(for example, the original Naming Game models the emergence of vocabulary for spatial relations[3]), we will look at a basic version of the Naming Game.[4]

In the model, the population tries to bootstrap a common vocabulary for a number of objects in the environments. Agents have inventories which consists of word-meaning pairs that she knows of. Initially, all agents have empty inventory. Two agents are randomly selected. One of them is assigned the role of speaker and the other becomes the listener. The speaker selects an object to express. If she has no words in her inventory for that object, she randomly creates a word and adds it to her inventory. If she already has some words for the object, she randomly chooses one. Then she conveys that word to the listener. The listener checks her inventory. If that word exists in her inventory, and the meaning associated with it is the right meaning, the interaction is considered successful.

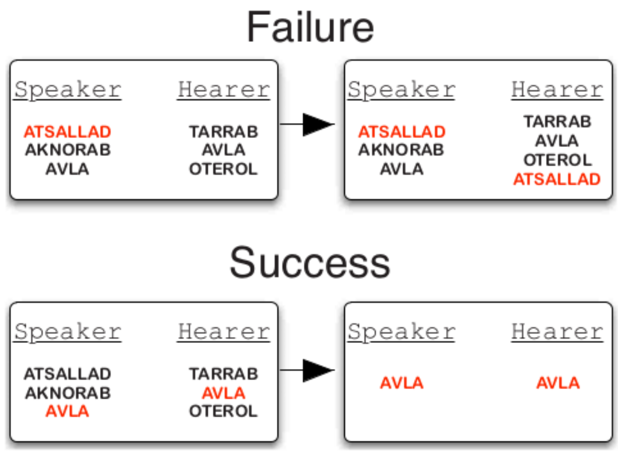


Fig. 1: In case of successful communication, only the word used in the communication is kept in the inventory while all other words are deleted. In the case of failed communication, the listener adds the word-object pair to her inventory. Figure is taken from [4]

If the word does not exist in the inventory, the interaction is considered a failure.

In case of a successful interaction, both parties keep only the chosen word for that object in their inventories, deleting all other associations for that word. If the interaction fails, the listener adds an association between the word and the object to her inventory.

Notice that this model has a number of assumptions:

- Any agent may interact with any agent. A fully connected topology is assumed.
- The number of possible words is assumed to be practically infinite so that the probability that two agents select the same word at different interactions is practically 0. This means that the case where the listener has the word in her inventory but the word is associated with different word than the correct one is impossible. Also, homonyms are impossible.
- From the last assumption: the processes of vocabulary emergence for different meanings are independent. Thus, the model can be analyzed for one word without loss of generality.
- It is assumed that the agents can understand whether the interaction was successful or not. This could be by pointing to objects, for example. This assumption is included to make sure we are discarding any telepathic communication between agents.

For a system with N agents(and 1 object, without loss of generality), there is always a nonzero probability p that the consensus state is reached in $2(N-1)$ interactions. After k interactions, the probability that the absorbing consensus state is not reached is $1-p$ or less. By iterating this process, the probability that the absorbing state is not reached in $2k(N-1)$ iterations is smaller than $(1-p)^k$ which goes to 0 as $k \to \infty$. Thus a common vocabulary will always be reached.

B. Evolutionary Language Game

As mentioned before, this model aims to explain the evolution of language across generations and also compares different strategies in language acquisition.[5]

In the model, agents are characterized by two matrices P and Q. Suppose there are m signals and n objects. The entries p_{ij} in an agent's production matrix $P_{n\times m}$ denote the probability that the object i is associated with signal j for that individual. Thus P is a matrix such that:

$$\forall i \leq n, \sum_{j=1}^{m} p_{ij} = 1 \tag{1}$$

Similarly, an entry q_{ji} in the agent's comprehension matrix $Q_{m\times n}$ denotes the probability that the agent understands object i upon receiving signal j. Thus,

$$\forall j \leq m, \sum_{i=1}^{m} q_{ji} = 1 \tag{2}$$

Therefore, for two agents I_1 and I_2 with matrices $P^{(1)}$ and $Q^{(2)}$ respectively, the probability that I_1 conveys the meaning for object i with success to I_2 is $\sum_{j=1}^{m} p_{ij}^{(1)} q_{ji}^{(2)}$. The ability of I_1 to convey information to I_2 is found by summing this quantity for all objects: $\sum_{i=1}^{n}\sum_{j=1}^{m} p_{ij}^{(1)} q_{ji}^{(2)}$. Finally, overall ability of successful communication between the two agents is found by summing up the abilities of conveying information from one to the other and dividing by two to find the average:

$$F(L_1, L_2) = \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{m}(p_{ij}^{(1)} q_{ji}^{(2)} + p_{ij}^{(2)} q_{ji}^{(1)}) \tag{3}$$

We are following the notation of [5] here. L_k stands for the language of k^{th} agent, since a pair of production and comprehension matrices define a language in the context of the model. The model is based on the premise that being able to communicate successfully increases an agent's chances of reproduction for the next generation in the evolution of the population. Hence $F(L_k, L_l)$ stands for the *fitness* between individual k and l. The individual fitness of an agent is her total fitness with all other agents:

$$F(L_k) = \sum_{l} F(L_k, L_l) \tag{4}$$

In order to measure the success of communication within the whole population, we define the overall success of communication within the population for population C:

$$F(C) = \frac{1}{2\binom{|C|}{2}}\sum_{k}\sum_{l} F(L_k, L_l) \text{ where } k \neq l \tag{5}$$

Finally, when the new generation is being generated, the probability that an agent k has an offspring is proportional to $\frac{F(L_k)}{\sum_l F(L_l)}$.

Notice that in this setup, individuals may have completely uncorrelated production and comprehension matrices. A new kind of matrix A called association matrix is introduced in order to specify how children of the new generation learn the language, also to tie P and Q to a common root.[6]

For n objects and m signals, entries a_{ij} of the matrix $A_{n\times m}$ of each agent holds a value denoting the strength of the association between the object i and signal j. Thus the matrices P and Q can be generated from A as such:

$$p_{ij} = \frac{a_{ij}}{\sum_{j'}^{m} a_{ij'}} \quad q_{ji} = \frac{a_{ij}}{\sum_{i'}^{n} a_{i'j}} \tag{6}$$

When a new generation is generated, the children learn the language by sampling a finite number of meaning-word pairs probabilistically from their teachers. They populate their association matrix A with by setting a_{ij} to the number of times they heard their teacher use word j for object i. We are not saying they learn the language from their parents because difference in these learning strategy is precisely what is being investigated in the early studies of this model.[5]

C. Orthogonal Models

The Naming Game model models the emergence and self organization of vocabulary, thus a horizontal organization, whereas the Evolutionary Language Game models vertical transmission, evolution and organization of language across generations.

It is known that languages change very much from generation to generation and it is probably at least in part due to the use of language among a generation, isolated from other generations. Thus, these two views of horizontal and vertical language organization are interwoven and a complete understanding needs to contain both sides of the picture.

There are a number of models that try to capture both views. They are in essence very similar. They all rely on a mechanism where agents have association matrices like in Evolutionary Language Game, but contrary to it, these matrices are dynamic. They change as the agents interact with other agents as in the Naming Game. During the interaction, the word is selected and understood probabilistically as in both models. After a successful interaction, the association between that word and object is strengthened and after a failure, the association is weakened.

These models have nuances in whether the matrices are understood as probability matrices or if the changes in the association matrices are constant [7, 8] or dependent on the agent.[9]

We will not give the details of the models and refer the reader to [7],[8] and [9] for further reading.

III. The Model

The model to be presented here actually models the self organization of a vocabulary in a population via interaction between agents, i.e. horizantal organization. Nevertheless, as we will see, the resulting structure can be used to generate association matrices as defined in section II.B. for each agent and evolutionary dynamics can be initiated for further studies.

In our model, instead of having inventories or matrices, each individual is defined by her memory for each of the M meanings(objects) in the environment. There are m memory sites for each meaning and each sites holds a word(signal). Number of words, W, may be limited or practically infinite. Initially, all agents have empty memories for all meanings.

Through interactions, agents will update their memories. As we explained in the introductory section, language is a game in its essence. Thus, a population that tries to settle down on a common vocabulary can be thought of as a set of agents that try to understand which words other agents use to convey which meanings. Thus, in our model, agents will hold the **last m words that they heard being used for each meaning** in their memory. Thus, **the same word can be repeated in the memory for a meaning**. In fact, in the optimal case, every agent has the same word in every memory site for a meaning, as will be clear in a moment.

Therefore, for an agent, after hearing m words(either different of identical) for a meaning, the memory for that meaning is full. When she hears a word w being used to convey that meaning one more time, she returns and overwrites the first memory cell with w. The next time, she overwrites the second cell, thus at each point in time, she keeps the last m words that she heard being used for each meaning in her memory for that meaning.

Agents interact in the model: two agents are randomly selected and one is assigned the role of speaker while the other one becomes the listener. Just like in the Naming Game model, the speaker selects a random meaning, let it be named μ. In order to convey μ to the listener, she checks her memory for it. If the memory is empty, she randomly creates a word and adds the word to her memory for μ. If, however, she has one or more words in her memory for μ, she randomly selects one. Then, the selected word(call it w) is spoken/sent to the listener, who tries to understand it. The listener checks her memory for all meanings. She makes a list where each meaning μ_i has k copies where k is the number of times w is found in the agent's memory for μ_i. After this, the listener randomly selects one meaning from the list(interprets w to mean this randomly selected meaning). If the interpretation is correct, the interaction is considered a success; if not, a failure. Notice that if there are meanings that contain w in their associated memory, the constructed list is empty and the interaction is a failure.

Recall that we are trying to model agents that are trying to figure out which words the others around them use to convey which meanings. This justifies our memory update rule:

- In case of a successful interaction, both parties add the word w to their memory for μ, since they both have seen a case where word w successfully conveys meaning μ.
- In case of a failed interaction, **only the listener updates her memory** and adds the word w to her memory for meaning μ while **the speaker does nothing**. This is because even though the interaction failed, the listener has seen that word w is used to convey meaning μ whereas from the speaker's point of view, there is no new information.

At the end of the interactions, the resulting data structure is at most m words for each of the M meanings. This is somewhat similar to inventories in the Naming Game, with the difference that there are repeated entries and there is an upper limit on the number of associations for a meaning. Notice further that this memory structure is equivalent to an association matrix in the spirit of the Evolutionary Language Game. In fact, it defines an association matrix such that each entry is a non-negative integer and the sum of entries on each row is less than or equal to m. If we assume enough number of interaction has been realized, we can safely say that the sum of entries on each row equals m.

We generate the association matrix from the memory structure of an agent as follows:

1) Total number of distinct words in all of the population's memories is counted. This is the number of columns. Let's denote it by D for dictionary.
2) $A_{M\times D}$ is composed of elements a_{ij} which is the number of times the j^{th} word is found in the agent's memory for the i^{th} meaning

Then, from A, production matrix P and comprehension matrix Q can be generated for all agents with equation (6).

IV. Results

Using the association matrices we computed for the agents, we can have a measure of the population fitness as defined by equation (5). Thus, we can study the effects of changing the parameters of our model.

The parameters are:

- N: number of agents in the population
- M: number of meanings(objects)
- m: memory size for each meaning
- W: number of words(signals)
- IT: number of iterations before the simulation is ended. Each agent is the speaker once in an iteration.

We also need a number of outputs:

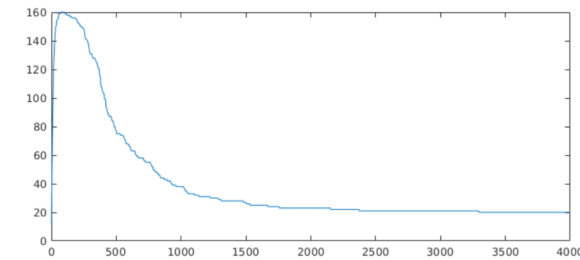
- D: number of words in use by the population
- F: population fitness

Notice that D is also an output of the Naming Game model and F is the population fitness defined in the Evolutionary Language Game.

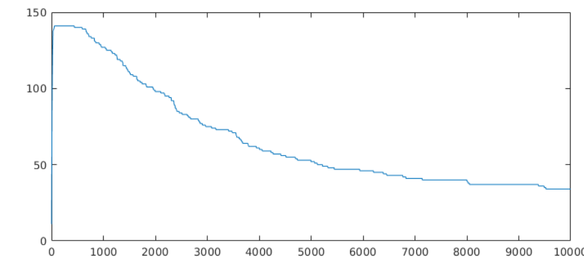
Figure 2 shows the time evolution of D for $N = 20$ agents, $M = 20$ meanings, $W = \infty$ words and different m and IT values.

The findings are strongly parallel with the findings of the original Naming Game model.[4] The system goes through 3 phases:

1) The number of words D increases rapidly as the agents populate their empty memories(inventories) with words.
2) The memories start to become correlated and the rapid increase starts to slow down and eventually stops where D reaches its highest value.
3) The memories start to align which is caused by successful interactions and at the same time increases the chances of more successful interactions. This results in relatively gradual but still rapid decrease in D until it settles down on its final value.



(a) This figure is for $N = 20$ agents, $M = 20$ meanings, $m = 8$ memory cells per meaning and $IT = 4000$ iterations.



(b) This figure is for $N = 20$ agents, $M = 20$ meanings, $m = 30$ memory cells per meaning and $IT = 10000$ iterations.

Fig. 2: Time evolution of the number of words present in the population's memory.

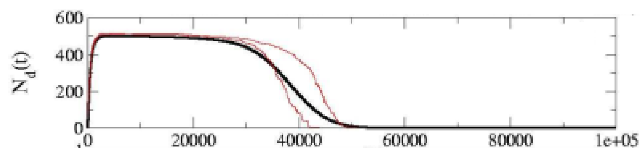


Fig. 3: The evolution of the number of different words used by the population in the original Naming Game model. The figure is taken from [4]

First of all, we used $W = 2^{32}$ for $W \to \infty$. The fact that there is a maximal value which is reasonably small than 2^{32} justifies this choice.

Second, we plotted the results for 2 different memory sizes. It is seen that as the memory size per meaning m increases, the slope of the plot in the 3^{rd} phase increases. Since the slope is negative, this means a slower, more gradual decrease and the vocabulary takes more time to settle down. This is why we ran the simulations on different number of iterations. The speed of increase in the first phase is not dependent on m.

Figures 4 and 5 in the last page show the main result of varying memory size m per meaning. The results are for $N = 20$ agents, $M = 20$ meanings, $IT = 1000$ iterations and different numbers of possible words W. Also we give the plots of the total number of words in use by the population (i.e. the number of columns in the association matrices of the agents) and all the data is averaged over 10 runs of simulation.

We see at Figure 5 that for small W, the number of words used by the population increases as we increase the memory size, until it reaches the maximum value. But after a certain value of W, the maximal value becomes less than W. This means that for a fixed sized population with fixed number of meanings, there is an approximately fixed value for the maximum number of words that the agents will hold in their memory. We can see from the plots that this value is about 150 for the $M = N = 20$ case.

The results suggest that emergence of vocabulary is the result of low memory size. We see that as we increase W, the success rate at big w values increases. This is due to less homonymy caused by limited word space. But in any case, low memory sizes result in greater population fitness.

Looking at Figures 4 and 5 together, we see that for memory sizes that result in the greatest success rate, the number of words used by the population is low. This is because for greater memory sizes per meaning, less words get overwritten and this causes at the same time high D and low F.

Notice that since the resulting structure is nothing more than an association matrix, all results regarding evolutionary dynamics of language transmission and different modes of language acquisition in the context of Evolutionary Language Game [5] applies to this model too.

V. Conclusion

We have presented a new agent based model for the emergence of vocabulary that includes memory. This model is similar to the Naming Game model and the resulting structure is essentially equivalent to the structure in the Evolutionary Language Game model. The difference is that in our model, words can be created and forgotten with time. Thus, it can be seen as an orthogonal model because while it describes the horizontal self organization of vocabulary and can be used to investigate evolutionary language transmission dynamics.

We have shown that low memory size increases the population fitness and decreases the time needed for the vocabulary to settle down.

We now list some further studies that can be carried.

This model can be used to model language change in **evolving environments** since it contains mechanisms for evolving the vocabulary to include new meanings and forget old words. Also, it can be used to model new learning mechanisms apart from those investigated in the context of the Evolutionary Language Game model because it integrates interactions between peers to the framework.

Also, emergence of syntax with interactions between the agents in a population can be investigated in the same spirit as and with inspirations from the studies in this area in the context of Evolutionary Language Game model.[6][10] The same goes for studying the effects of errors in interactions and possible linguistic error limits.[11]

References

[1] Chomsky N., Studies in Chinese Linguistics, Vol. 33, No. 1, pp. 3-16, 2012

[2] Castellano C., Fortunato S. and Loreto V., Reviews of Modern Physics, Vol. 81, pp. 591-646, 2009

[3] Steels L., Artificial Life, Vol. 2, pp. 319-332, 1995

[4] Baronchelli A., Felici M., Caglioti E., Loreto V. and Steels L., Journal of Statistical Mechanics: Theory and Experiment, 2006

[5] Nowak M. A., Plotkin J. B. and Krakauer D. C., Journal of Theoretical Biology, Vol. 200, pp. 147-162, 1999

[6] Nowak M. A. and Komarova N. L., TRENDS in Cognitive Sciences, Vol.5, No.7, pp. 288-295, 2001

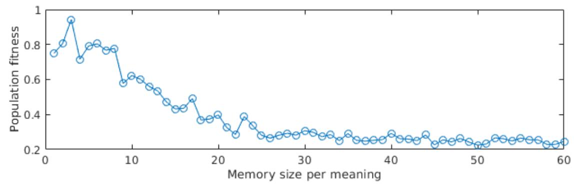
[7] Ke J. Minett J. W., Au C.-P. and Wang W. S.-Y., Complexity, Vol. 7, pp. 41-54, 2002

[8] Lenaerts T., Jansen B., Tuyls K. and De Vylder B., Journal of Theoretical Biology, Vol. 235, pp. 566-582, 2005

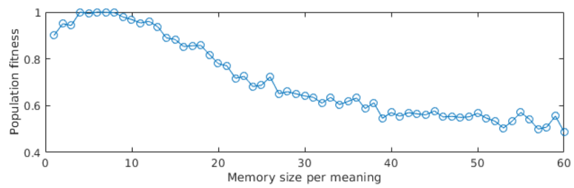
[9] Lipowski A and Lipowski D., submitted for proceedings of Socio-Econo-Physics Summer School in 2007, accessed from https://arxiv.org/abs/0801.1658v2

[10] Nowak M. A., Komarova N. L. and Niyogi P., Nature, Vol. 417, pp. 611-617, 2002

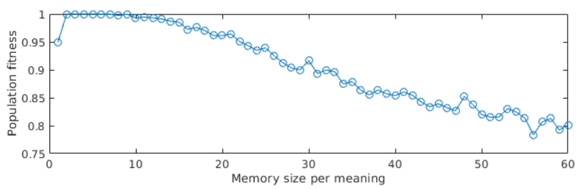
[11] Nowak M. A. and Krakauer D. C., Proceedings of the National Academy of Sciences of the United States of America, Vol. 96, pp. 8028-8033, 1999



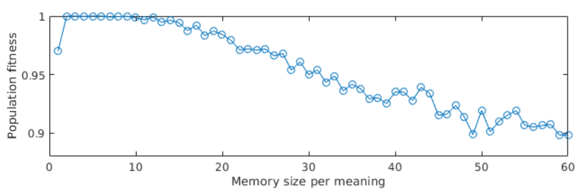
(a) $W = 20$



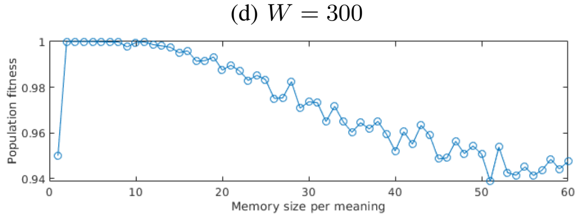
(b) $W = 50$



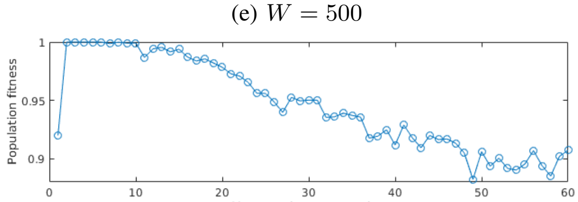
(c) $W = 150$



(d) $W = 300$

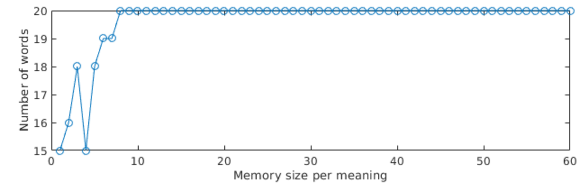


(e) $W = 500$

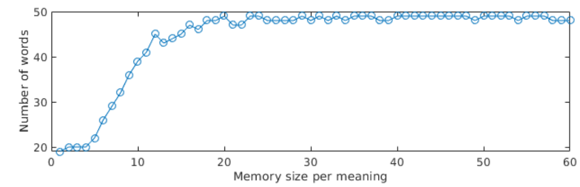


(f) $W = \infty$

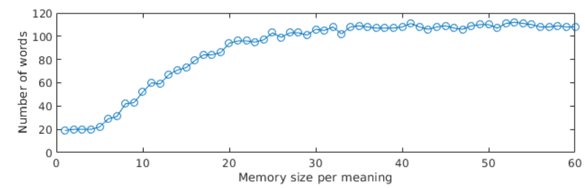
Fig. 4: Population fitness F after $IT = 1000$ iterations with $N = 20$ agents, $M = 20$ meanings where memory size per meaning m ranges from 1 to 60 for different number of possible words W averaged over 10 simulation replications



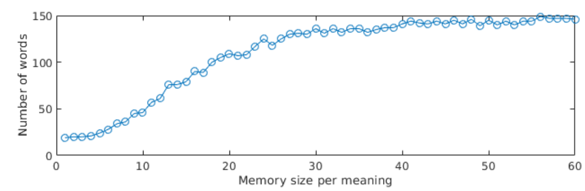
(a) $W = 20$



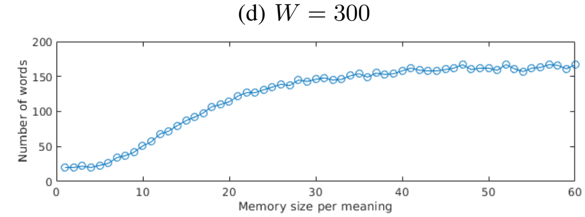
(b) $W = 50$



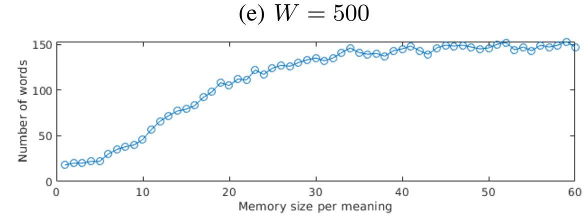
(c) $W = 150$



(d) $W = 300$



(e) $W = 500$



(f) $W = \infty$

Fig. 5: Total number of words D used by the population after $IT = 1000$ iterations with $N = 20$ agents, $M = 20$ meanings where memory size per meaning m ranges from 1 to 60 for different number of possible words W averaged over 10 simulation replications