Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Lecture with Computer Exercises:

Modelling and Simulating Social Systems with MATLAB

Project Report

Language Formation:

The Nativist and Non-Nativist Dilemma

Afroditi Apostolou & Dimitrios Bolis

Zürich

December 2010

Eigenständigkeitserklärung

Hiermit erkläre ich, dass ich diese Gruppenarbeit selbständig verfasst habe, keine anderen als die angegebenen Quellen-Hilsmittel verwenden habe, und alle Stellen, die wörtlich oder sinngemäß aus veröffentlichen Schriften entnommen wurden, als solche kenntlich gemacht habe. Darüber hinaus erkläre ich, dass diese Gruppenarbeit nicht, auch nicht auszugsweise, bereits für andere Prüfung ausgefertigt wurde.

Name 1	Name 2
Afroditi Apostolou	Dimitrios Bolis

Agreement for free-download

We hereby agree to make our source code of this project freely available for download from the web pages of the SOMS chair. Furthermore, we assure that all source code is written by ourselves and is not violating any copyright restrictions.

Name 1	Name 2
Afroditi Apostolou	Dimitrios Bolis

"...People would like to think that there's somebody up there who knows what he's doing. Since we don't participate, we don't control and we don't even think about questions of vital importance. We hope somebody is paying attention who has some competence. Let's hope the ship has a captain, in other words, since we're not taking part in what's going on...
It is an important feature of the ideological system to impose on people the feeling that they really are incompetent to deal with these complex and important issues: they'd better leave it to the captain. One device is to develop a star system, an array of figures who are media creations or creations of the academic propaganda establishment, whose deep insights we are supposed to admire and to whom we must happily and confidently assign the right to control our lives and to control international affairs...."
- Noam Chomsky

Contents

Abstract

In this project we dealt with the subject of language formation and evolution that has recently drawn the attention of many disciplines such as evolutionary biology, computation, psychology and cognitive science.

More specifically, we tried to shed some light on the so called “Nativist-Non Nativist Dilemma” that deals with the question whether the ability of speaking languages is innate to humans or arises from the application of human’s cognitive skills to the problem of communication.

For that purpose, we implemented the model that Nowak et al [1] proposed. Nowak’s approach assumes that population taken as an input in his model, has no protolanguage and after a number of iterations they start to correspond sounds to objects which actually constitutes language.

In our project, we implemented Nowak’s model and produce results similar to his work. Although we extended that model and used it for population that has some sort of ability to correspond sounds to objects. This was our Nativist extension to the model and the results we produced reveal that in this way language is formed faster and is more effective.

1 Introduction

Language Definition: Language may refer either to the specifically human capacity for acquiring and using complex systems of communication, or to a specific instance of such a system of complex communication. The scientific study of language in any of its senses is called *linguistics.* A language is a system of signs for encoding and decoding information. The English word derives from Latin lingua, "*language, tongue*". This metaphoric relation between language and the tongue exists in many languages and testifies to the historical prominence of spoken languages [3]. When used as a general concept, "language" refers to the cognitive faculty that enables humans to learn and use systems of complex communication.

1.1 Language Formation

Language is thought to have originated when early hominids first started cooperating, adapting earlier systems of communication based on expressive signs to include a *theory of other minds*[1] and shared intentionality. This development is thought to have coincided with an increase in brain volume. Language is processed in many different locations in the human brain, but especially in Broca's and Wernicke's areas. Humans acquire language through social interaction in early childhood, and children generally speak fluently when they are around three years old. The use of language has become deeply entrenched in human culture and, apart from being used to communicate and share information, it also has social and cultural uses, such as signifying group identity, social stratification and for social grooming and entertainment.

One of the main questions that concerns linguists for many years is how language is actually formed. Namely if humans have an innate ability of acquiring languages, or if language acquisition is actually a byproduct of human's cognitive skills, in order to solve the problem of communication. This question seems even more important if

[1] **Theory of Mind:** It is hypothesized that when humans learn the meaning of a novel word, they try to simulate the speaker's thought process to form their own theory of the speaker's intention, thus trying to understand what the speaker meant. To do this, the hearer must know that the speaker has a mind similar to its own and must be able to estimate how the speaker came to its decision to talk about something [5].

we take into account the fact that the use of language is considered to be one of the main differences of human and other animals. A few experiments have been conducted throughout the years based on this concept. Among the most interesting ones, are those that human's closest relatives such as chimpanzees, are taught basic vocabulary. These experiments revealed that even of chimpanzees can really learn a number of approximately 100 words, they cannot use syntax or grammar, thus they cannot combine words in order to make phrases [4]. On the contrary, children are usually fluent speakers at the age of three years old, with no over training at all.

This question can be considered as a philosophical one and this is why it was restricted to philosophers or sometimes linguists for many years. Although the last decades the interest of more disciplines is focused on that. More specifically, recently sciences of evolutionary biology, computation, psychology, and cognitive science deal with the problem of the language formation.

The so called "Nativist vs Non Nativist Dilemma" summarizes all mentioned above and is extensively described in the following section.

1.2 The Nativist vs Non Nativist Dilemma

As already mentioned there are two opposing schools of thought when consider the formation of language. The supporters of the Nativist point of view, such as Noam Chomsky, claim that the linguistic abilities of human beings are innate and base their argument to the procedure that infants acquire their language: with no training but just receiving the inputs of their environment, they evolve in fluent speakers usually at the age of three. On the other hand, the supporters of the Non Nativist point of view, such as Michael Tomasello, reject the idea of an innate universal grammar and instead propose a usage-based theory (sometimes called the *social-pragmatic* approach to language acquisition) in which children learn linguistic structures through intention-reading and pattern-finding in their discourse interactions with others [6].

1.2.1 The Nativist point of view

The corner stone of the Nativists' approach is that the ability of humans to acquire language is determined by biological factors; namely there is an instinct or even specific organs in human's brain that lead a human to learn and use grammar and syntax.

One of the widely known supporters of this school is Noam Chomsky that as mentioned before, bases his arguments to the procedure that children learn how to speak. What Chomsky believes about that is summarized as follows: *Children must innately be equipped with a plan common to the grammars of all language, a Universal Grammar, that tells them how to distil the syntactic patterns out of the speech of their parents. The unordered super-rules (principles) are universal and innate; when children learn a particular language, they do not have to learn a long list of rules, because they are born knowing the super-rules . All they have to learn is whether their particular language has the parameter head-first, as in English, or head-last, as in Japanese. If the verb comes before the object, the child concludes that the language is head-first as if the child were merely flipping a switch to one of two possible positions. The way language works is that each person's brain contains a lexicon of words and the concepts they stand for (a mental dictionary) and a set of rules that combine the words to convey relationships among concepts (a mental grammar)* [7][8] .

Chomsky's approach will be explained in detail in a following section.

There are several reasons to be convinced by the Nativists. The majority of their arguments are based on observations of children under six years old acquire language. A few reasons are mentioned below [9]:

- The strongest reason comes from genetic/biological data and research in child acquisition. Chomsky's original argument was largely based on evidence from acquisition and what he called the ``poverty of the stimulus'' argument. The basic idea is that any language can be used to create an infinite number of productions -- far more productions and forms than a child could correctly learn without relying on pre-wired knowledge. For example, English speakers learn early on that they may form contractions of a pronoun and the verb to be in certain situations (like saying ``he's going to the store''). However, they cannot form them in others; when asked ``who is coming'' one cannot reply ``he's,'' even though semantically such a response is correct. Unlike many other learning tasks, during language acquisition children do not hear incorrect formulations modeled for them as being incorrect. Indeed, even when children might make a mistake, they are rarely corrected or even noticed. This absence of negative evidence is an incredible handicap when attempting to generalize a grammar, to the point that many linguists dispute whether it is possible at all without using innate constraints.

- Another reason is that as Nativists claim, children never make specific mistakes. For instance, consider the sentence A unicorn is in the garden. To make it a question in English, we move the auxiliary is to the front of the sentence, getting Is a unicorn in the garden? Thus a plausible rule for forming questions might be ``always move the first auxiliary to the front of the sentence''. Yet such a rule would not account for the sentence A unicorn that is in the garden is eating flowers, whose interrogative form is Is a unicorn that is in the garden eating flowers?, NOT Is a unicorn that in the garden is eating flowers? The point here is not that the rule we suggested is incorrect -- it is that children never seem to think it might be correct, even for a short time. This is taken by Nativists like Chomsky as strong evidence that children are innately ``wired'' to favor some rules or constructions and avoid others automatically.
- To make this point of view even stronger, it is noteworthy that children that are exposed to language before the age of six, are able to learn fluently and faster a language (regardless of their general intelligence or circumstances of environment) rather than after this age. This reveals that there is a critical period for language and so it should be determined by biological factors rather than the need of communication (as it still exists after the critical period). The power of this argument can be reinforced by two simple examples: firstly, it is observed that it is difficult for children after the age of six or adults to learn a second language; even more they often never adopt the proper accent. Secondly, we can take as example the so called "wild

children". These are the children that have not been exposed to language before adolescence and so they never end up learning a language that even approaches full grammaticality.

- One of the strongest arguments of the Nativists is that language acquisition is not directly dependent on people's intelligence. It is observed that there can be found individuals with normal intelligence but extremely poor grammatical skills, and vice versa, suggesting that the capacity for language may be separable from other cognitive functions. Individuals diagnosed with Specific Language Impairment (SLI) have normal intelligence but nevertheless seem to have difficulty with many of the normal language abilities that the rest of us take for granted. The opposite case of SLI exists as well: individuals who are demonstrably lacking in even fairly basic intellectual abilities who nevertheless use language in a sophisticated, high-level manner. Fluent grammatical language has been found to occur in patients with a whole host of other deficits, including schizophrenia, autism, and Alzheimer's. One of the most provocative instances is that of William's syndrome. Individuals with this disease generally have mean IQs of 50 but speak completely fluently, often at a higher level than children of the same age with normal intelligence.

1.2.2 The Non Nativist point of view

On the contrary of the Nativist point of view stands the Non Nativist one, that claims that human acquires his language because it is needed for communication. It is believed that application of human cognitive skills to the problem of communication will lead to language formation. Non Nativist supporters, such as Tomasselo, use the evidence supporting the Nativist point of view, against it in order to prove it implausible.

Below there are stated the Non Nativists' arguments against the Nativists. For every argument of the subsection 1.2.1 there will be presented the counter-arguments [9]:

- First, and most importantly, there is increasing indication that Chomsky's original ``poverty of the stimulus'' theory does not adequately describe the situation confronted by children learning language. For instance, he pointed to the absence of negative evidence as support for the idea that children had to have some innate grammar telling them what was not allowed. Yet, while overt correction does seem to be scarce, there is a consistent indication of parents implicitly ``correcting'' by correctly using a phrase immediately following an instance when the child misused it. More importantly, children often pick up on this and incorporate it into their grammar right away, indicating that they are extremely sensitive to such correction.

- The second argument of Nativists is rather easily contradicted; Non Nativists claim that children are incredibly well attuned to the statistical properties of their parents' speech [10]. The words and phrases used most commonly by parents will - with relatively high probability - be the first words, phrases, and even grammatical structures learned by children. They also claim that children tend to pay more attention to some words than others and they use them piece-by-piece rather than generalizing rules. These arguments explain why children do not make mistakes that would be expected to be made, without the need for biological factors to play any role. Moreover they claim that the fact that children learn so fast and effectively is the result of their parent's effort to be good teachers for them: when they speak to young children they automatically adjust their own language level to approximately what the child is able to handle.
- Concerning the critical period mentioned by Nativist, Non Nativist base their disbelief to the fact that people can actually learn multiple languages not only until the age of six but even as adults, and they can evolve in fluent speakers. This fact is by itself sufficient to discredit the critical period hypothesis. The biological definition of critical periods (such as the period governing the development of rods and cones in the eyes of kittens) requires that they not be reversible at all [11]. Once the period has passed, there is no way to acquire the skill in question. This is clearly not the case for language. The example of "wild children" is not actually taken into consideration for Non Nativists, as they claim that such children usually suffered extraordinarily

neglectful early conditions in other respects, which might mitigate the results. Moreover these incidents are very few to lead to any conclusions.

- The existence of genetic impairments like Specific Language Impairment seem to be incontrovertible proof that language ability must be domain-specific (and possibly innate as well), but there is controversy over even this point. Recent research into SLI indicates that it arises from an inability to correctly perceive the underlying phonological structure of language, and in fact the earliest research suggested this. This definitely suggests that part of language ability is innate -- namely, phonologicalperception -- but this fact is well accepted by both nativists and non-nativists alike. It is a big leap from the idea that phonological perception is innate to the notion that syntax is.

1.3 Motivations

Since language is considered to be one of the main features that distinguish human from other animals, it raises queries about how it is actually formed. The fact that only human has the ability to use grammar and syntax instead of just vocabulary or even sounds, raises the question if there are biological factors that lead human to form his language. Questions like these were our primer motivation to deal with such a subject.

Moreover, living in the era of multi-modal communication made it very interesting to think about the basis of all that which is the language. And even more, from an engineer's point of view, language is the first protocol that human ever invented. So it was worthwhile to work on that and try to understand the roots of language formation.

2 Individual Contributions

The basis of this project were two papers about the language formation written by M.Nowak [1] and N.Chomsky [2] respectively. Nowak is one of the supporters of the Non Nativist point of view whereas N.Chomsky is one of the most widely known linguists, supporting the Nativist point of view.

We implemented the model that Nowak proposed, but extended it to also include Chomsky's approach.

3 Description of the model

3.1 Nowak's Model Implementation

First we implemented Nowak and attempted to produce results similar to those produced from his work.

At first we have to describe Nowak's scenario: He assumes that language evolved as a means of communicating information between individuals. The population involved in the simulating process is a group of early hominids with no protolanguage and no innate ability to communicate. Namely biological factors play no role at this scenario. A language formed at early stages is based on the correspondence o a sound to a specific object. This process is assumed to lead to the formation of a higher level language. At first he supposes n sounds and m objects.

Communication between two individuals practically means that the speaker will make a sound to describe an object and the listener will understand the same object when hears this sound. Thus individuals are treated both as speakers and listeners. More specifically, p_{ij} refers to the probability that the speaker will make the sound I when sees the object j. These probabilities constitute a matrix P_{ij}, called "**Active matrix"**. On the other hand, q_{ij} refers to the probability that the listener will understand the object j when hears the sound i. These probabilities constitute the Q_{ij} **"Passive matrix".**

To evaluate communication between two individuals, A and B, Nowak calculates a **payoff function F.** When two individuals speaking a slightly different language L (described by P_{ij} and Q_{ij}) and L' (described by P'_{ij} and Q'_{ij}) , then the probabilities of correct speaking and understanding are denoted by p_{ij},q_{ij} for A and p'_{ij},q'_{ij} for B. When A sees object i and signals then B will infer object I with probability

$$q_{ij} = \sum_{j=1}^{m} p_{ij} q'_{ij}$$

Summing this probability over all objects will give a metric to measure A's and B's communication efficiency. So the overall payoff function, if every individual is treated both as a speaker and listener is given by the equation:

$$F(L, L') = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{m} (p_{ij}\, q'_{ij} + p'_{ij}\, q_{ij})$$

The above mentioned are repeated for a number of iterations and every time the payoff function is calculated. At each round this value plays significant role in order to decide which agents will survive to the next round (agents with high values of payoff function have higher possibility to survive) and also how many off springs each agent will have. Off springs are assumed to learn their parents' language.

Running his model for a population of 100 agents, Nowak produced the plot below for 5 iterations:

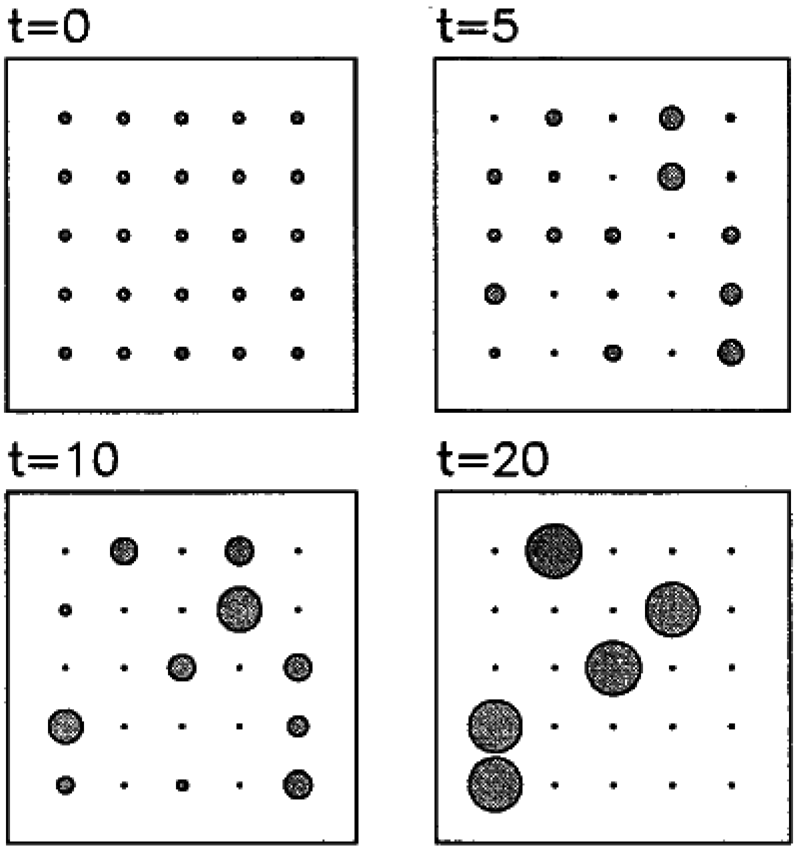


figure 1: *Emergence of a protolanguage in an initially prelinguistic described by this society. The population consists of 100 individuals. Each of them starts with a randomly chosen P and Q matrix. There are five objects and five signals (sounds). In one round of the game, every individual interacts with every other individual and the payoff of all interactions is evaluated. The total payoff of all individuals is calculated. For the next round, individuals produce offspring proportional to their payoff. Children learn the language of their parents by sampling their responses to each object. The figure shows the population average of the P matrix; the radius of the circle is proportional to the corresponding entry. Initially, all entries are about 0.25. After five generations some initial associations begin to form, which become stronger during subsequent rounds. At t = 20, each object is associated with one signal. Signal 1, however, is used for two objects, whereas signal 5 is not used at all. This solution is suboptimum but evolutionarily stable. Interestingly, errors during language acquisition increase the likelihood of reaching the optimum solution.*

After simulating this model producing results that are described at next chapter, we made some extensions to include Chomsky's point of view.

3.2 Extensions

Our first extension concerns the input to this model: we now assume that the population has some kind of innate ability to communicate. This is claimed by

Chomsky. For this purpose, the initial probabilities p_{ij} are not about 0.25. They are still produced pseudorandomly, normally distributed but the standard deviation is now large. This actually means that each agent has an innate ability to make the correspondence between objects and sounds. So at this point, the Nativist point of view is approached.

Moreover, we calculated another metric to evaluate the efficiency of the formatted language. For each object ,this metric represents how "stronger" is the selected sound from the rest four sounds. More precisely, for every object we calculated the mean value of the differences of the selected sound from every other sound. The results produced are given in the next chapter.

The final extension we implemented is based on the idea that effective agents are able to influence the rest of the population.

The final extension we implemented is based, firstly, on the idea that effective agents form larger communities and, secondly, the assumption that the larger the community is the greater the influence to its members will be. People within the same community, communicating mostly with agents within the community itself, will tend to boost the differences between the dominant sounds and the rest ones, in order to maximize the efficacy of the communication. That is, beside the basic general positive feedback mechanism expressed in the cost function, we consider as well, secondary feedback mechanisms inside each community.

4 Implementation

Firstly, the input data are constructed. The program asks from the user to make a selection for the Nativist or Non-Nativist Mode regarding the Language Formation Model. After that, either the "rand" or the "random" function is called for creating pseudo-random data of Uniform or Gaussian Distribution respectively. For the Gaussian distribution the mean was selected as $\mu=0.5$ and the standard deviation as $\sigma=0.01$. The resultant proto-languages are expressed in the form of two 5x5 matrices, P and Q, containing randomly distributed probabilities, as it is described in

Nowak's model. Then, a cost Function F is calculated, using the P's and Q's values, as it was described above, in order to define the communication efficacy, when using each language. This function is also weighted by a factor indicating the strength of each agent after each iteration (generation). It is for the first generation equal to one, and for the next ones, dependant on the strength of each agent, calculating in F. The weights are then normalized in a way that the whole dynamic range is fitted in the range of [0, 100].

Next the Language Adaptation is performed. This procedure is based on the hypothesis that large communities have the power to influence their members intensely and then the members to influence the community itself, forming a positive feedback loop. The basic parameter ("comun") is calculated at the beginning of this section. "Comun" is a binary variable, taking the values 0 or 1 according to a certain probability. This probability is depended on the "W1" variable which is the normalized weight in the range of [0, 1]. The greater is the "W1", meaning stronger agent, the greater the probability of "comun" equals to 1 will be. In the next stage, this variable will define whether or not the dominant sound of the specific language should be increased, resulting from the close interaction of the community members. The increase is not a fixed value but a ratio (1%) of the initial value to be increased.

Then the metric used by Nowak is calculated in the variable "PopAvg", which essentially is a 5x5 matrix, as P, but instead of the P's probabilities, contains the total number of people have maximum probability at this specific element. That means, that the numbers of "PopAvg" express how many people are most likely to use each sound when seeing a specific object.

Then the first step to calculate the metric, proposed by us, is implemented. That is the "Pavg" is calculated, which is the weighted mean value of all the agents' P matrix. In other words, the Pavg(i,j) denotes the mean probability of a random person in the world to make the sound j, when seeing the object i.

Then the elimination of weak agents follows. In this part of the code the weakest agents excluded from the next generations. The selection is based on the minimum values of the weight W.

As a last step, the “Pcr” (P-Critical), which the metric proposed by us, is calculated. As it is referred above, the “Pavg”, matrix is used. The maximum value for each row is found and the mean distance from the other four values, lying at the same line, is calculated. In other words this metrics quantify the dominance of the most likely sound, regarding to a certain object, in comparison with the rest available in the language.

Finally, in order to make our measurements more general and check the robustness of our models, which are initialized in a random way, we introduce the idea of parallel evolution in a multiple universe. So, we run our models and check the results in each universe distinctly and then calculate the mean metrics to show the general tendency.

5 Simulation Results & Discussion

The results produced from the simulation of the model described above are divided in three sections: first we plotted diagrams similar those of Nowak’s ones (given in 5.1 Nowak’s Diagrams). Secondly, we plotted diagrams that show how the sounds are chosen in order to correspond in one object, for a number of iterations. In this way we can easily compare the performance of the Nativist against the Non Nativist model (given in 5.2 Choice of sound). Finally, we plotted our last metric-described in 3.2, that we called “P-critical”. These plots illustrate the comparison of the effectiveness of the Nativist against the Non Nativist model for multiple testing. (given in 5.3 Effectiveness of Nativist vs Non Nativist model).

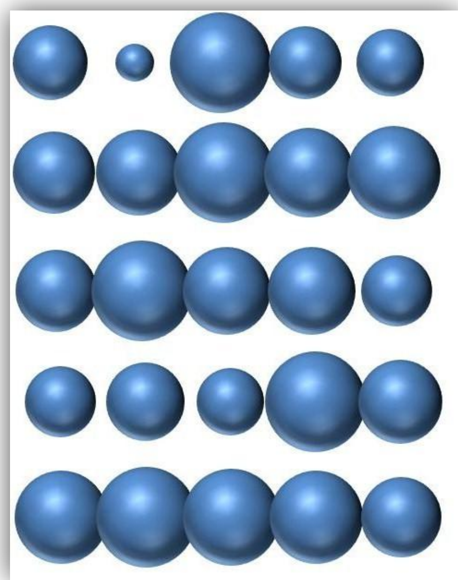
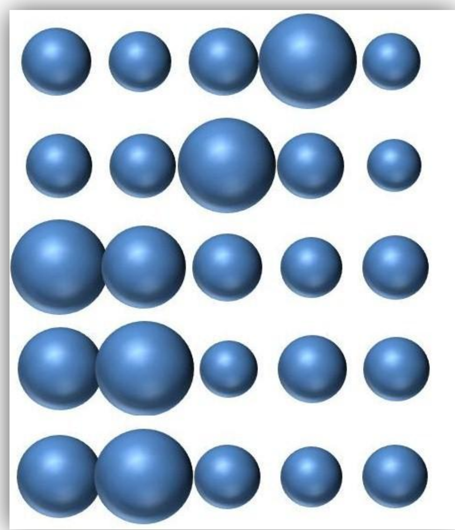
5.1 Nowak's Diagrams

The plots below illustrate better the dominant sounds for each object. We have 5 sounds and 5 objects. The diagrams are slightly different from Nowak's results because our model is much slower than Nowak's one. This deflection can imputed to differences during the implementation of the model such as:

- Nowak assumes random initial population without clarifying the method used to produce it. In our model we used random initial population with normal distribution and very small standard deviation. This means that the initial population has no proto-language or any instinct leading to language formation.
- Nowak claims that uses the values of the payoff function in order to proceed at the reproduction of the population, again without clarifying the specific method used. In our model, we used the values of the payoff function in order to produce a normalized weighting factor that drives the reproduction process. We also implemented the idea of agent's survival to the next generation that Nowak refers to, by eliminating the weakest agent of each iteration.

Differences like the aforementioned are able to lead to the production of different results. We based our model to Nowak's work and to reasonable assumptions for steps that are not clearly defined in Nowak's paper.

Below there are presented the results similar to Nowak's ones for Non Nativist model, for a number of iterations:

ITERATIONS	SOUNDS DISTRIBUTIONS (5 SOUNDS,5 OBJECTS)
t=50	
t=80	

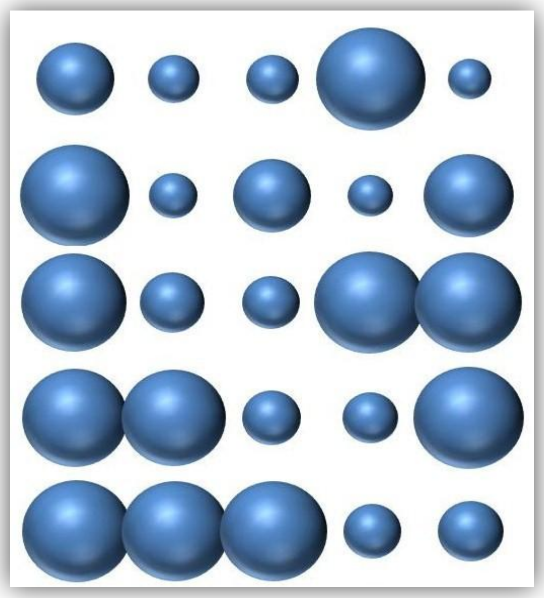
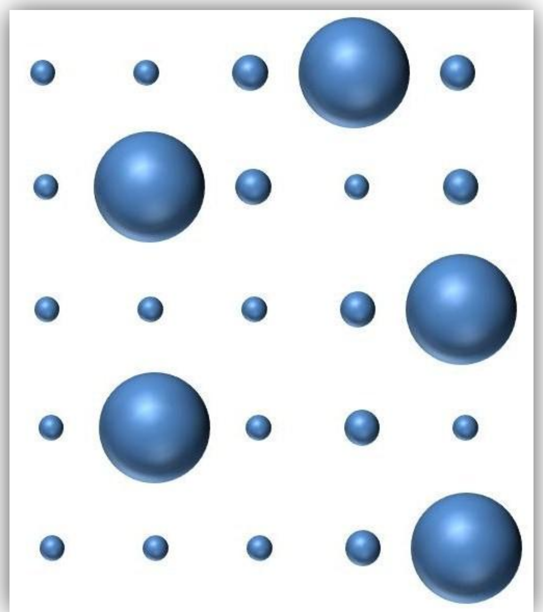
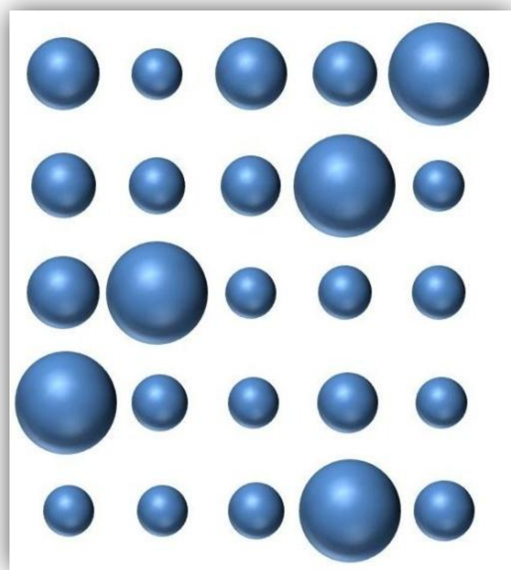
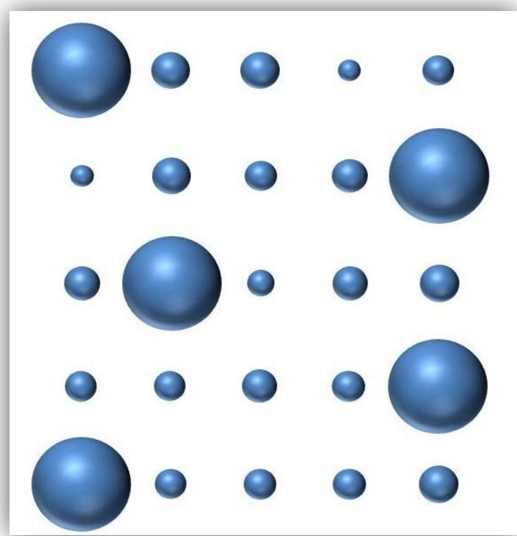
t=100	
t=200	

figure 2: *Language formation process for the Non Nativist model*

Comments: From the figures above one can clearly notice how a language is produced in terms of connecting sounds with objects. We can observe the formation of the language as time goes by (higher number of iterations), which means that sounds start dominating after some iterations. This is clearly described by the last image for 200 iterations.

At this point we thought that it would be noteworthy to produce the same kind of results for the Nativist model that is inspired by the Chomsky's thoughts. The plots describing this are given below:

ITERATIONS	SOUNDS DISTRIBUTIONS (5 SOUNDS,5 OBJECTS)
t=50	
t=80	

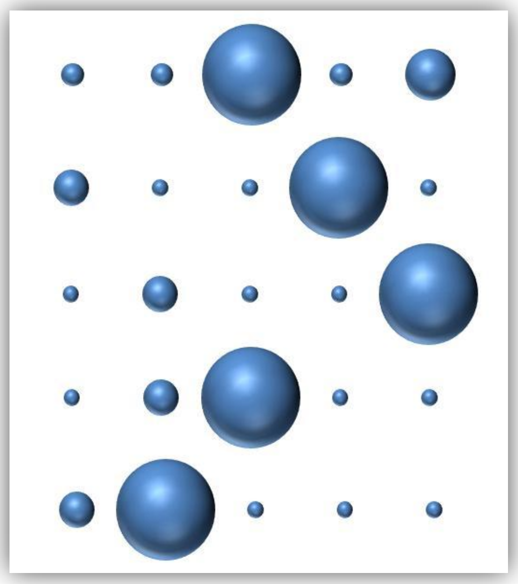
t=100	

figure 3: *Language formation process for the Nativist mode*

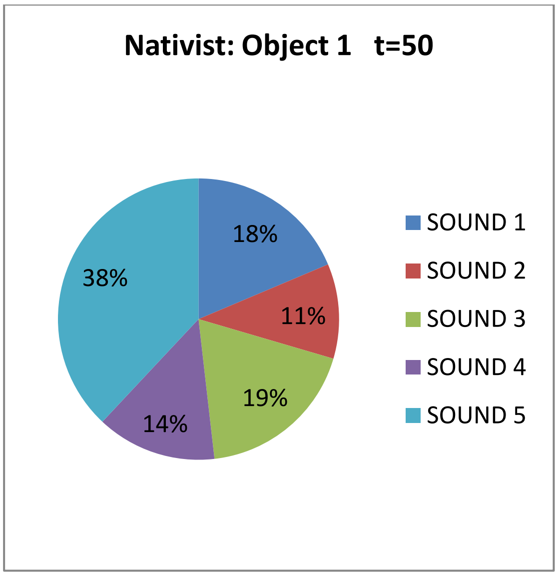
Comments: Our results produced, seem to be more consistent with Chomsky's point of view. Although both models are able to lead to language formation, the Nativist model seems to do so faster and as it will be explained later, more effectively.

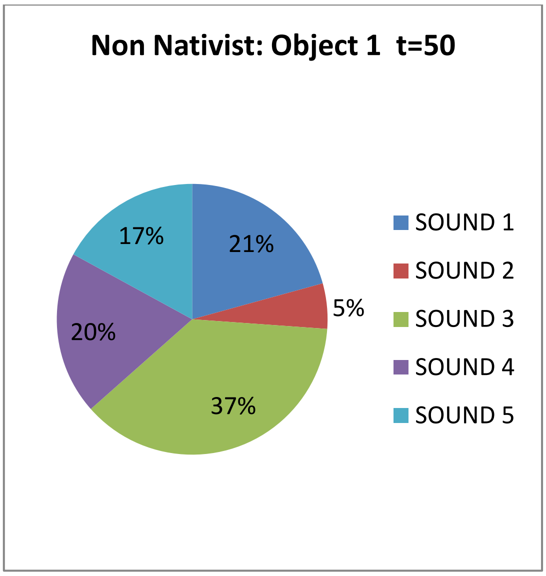
5.2 Choice of Sound

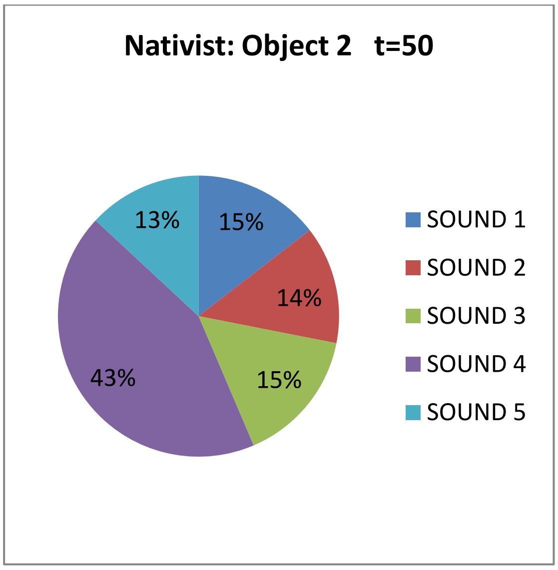
Below there are given diagrams which illustrate the distribution of the sounds chosen in order to describe one object. We can easily observe that for a small number of iterations, all sounds have similar possibilities to be chosen. As iterations increase, plots show that one sound starts dominating for each object. It is though noteworthy that the results produced are suboptimal but evolutionary stable; which means that sometimes the same sound is chosen to describe two different objects. This is also observed at Nowak's results at [1]. Nowak finds a possible explanation in the absence of noise at this stage of the simulation. He claims that "*errors during language acquisition increase the likelihood of reaching the optimum solution*".

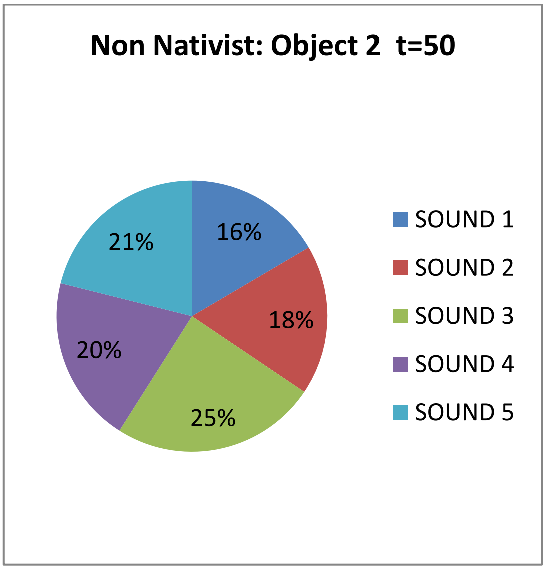
We also note that iterations correspond to the number of generations that are involved in each stage of the simulating process.

- For 50 iterations the results produced for the Nativist and Non Nativist model, for each object are shown below:









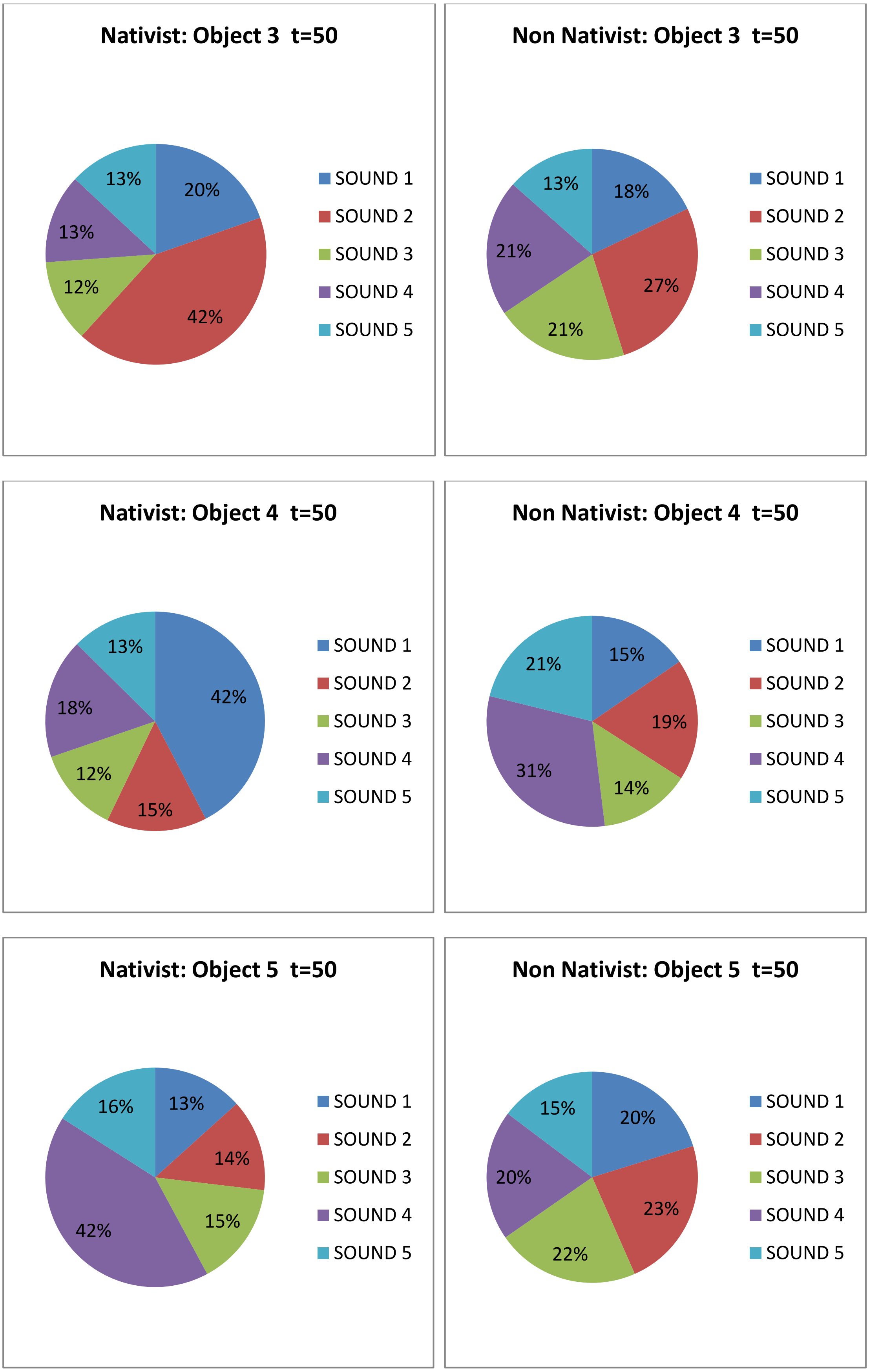
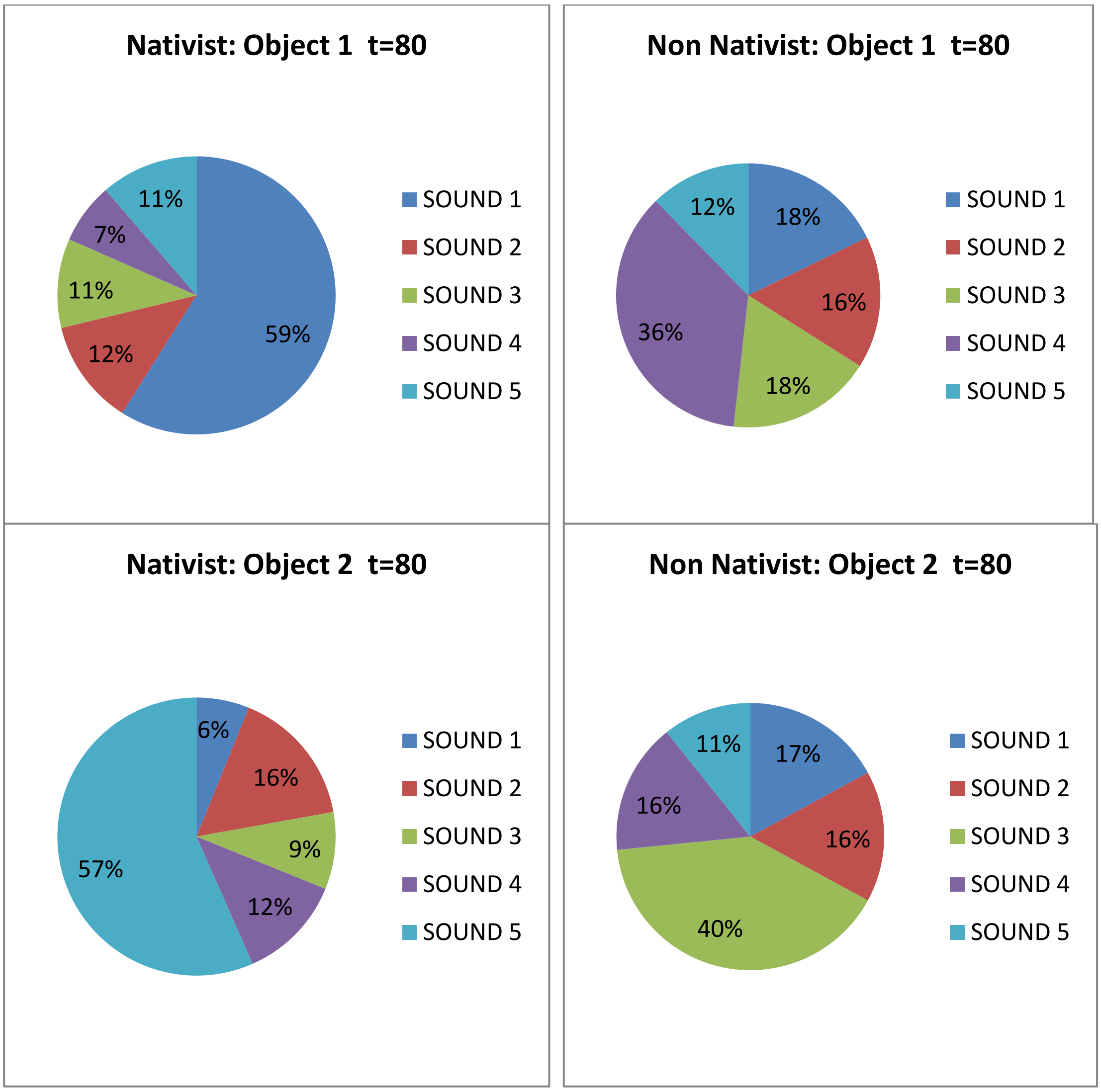


figure 4: *Distribution of sounds for (a) Nativist and (b) Non Nativist model, for 50 iterations*

Comments: from these diagrams we can draw the conclusion that the Nativist model proceeds to the choice of the sound faster than the Non Nativist one. For the Nativist model we can see that even for a small number of iterations, some sounds start dominating; this fact corresponds to the formation of an early stage language. On the contrary at the same time the Non Nativist model does not give any significant results: all sounds have similar possibilities to be chosen for each word.

- For 80 iterations the results produced are:



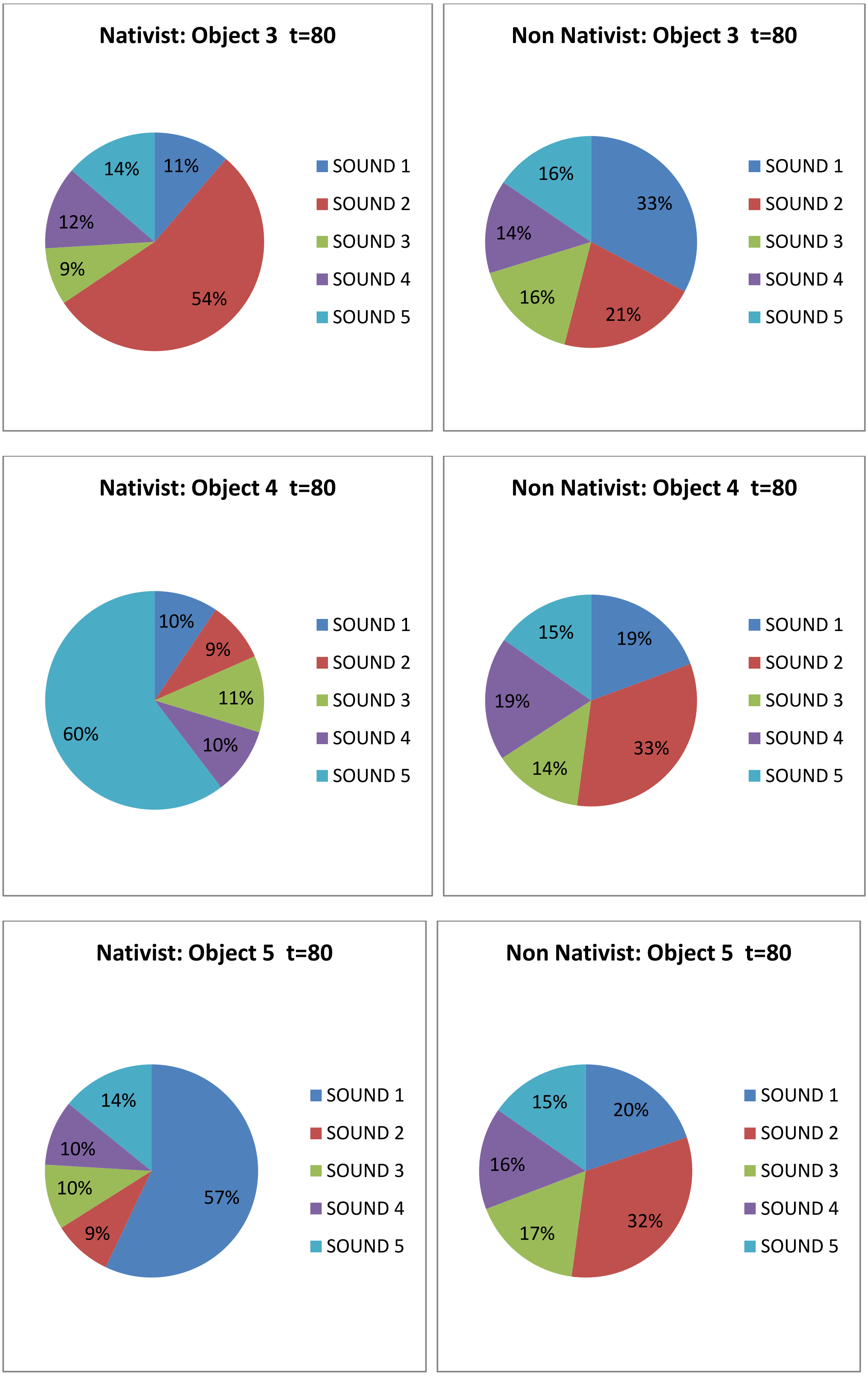
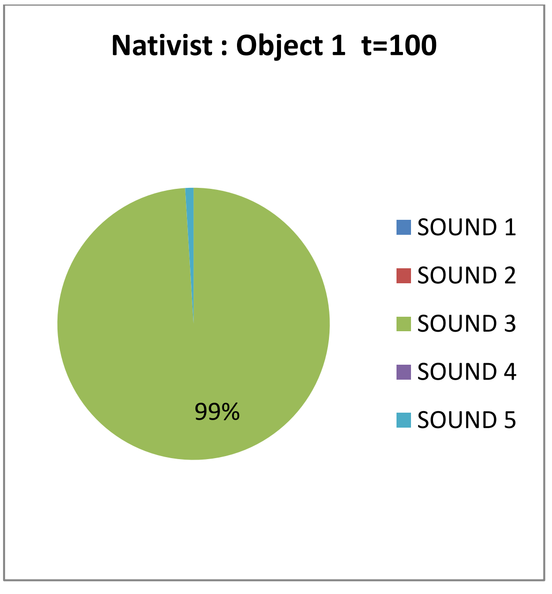
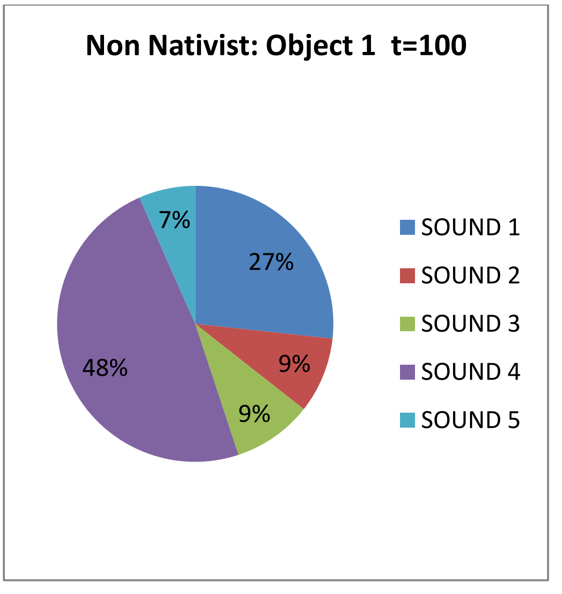


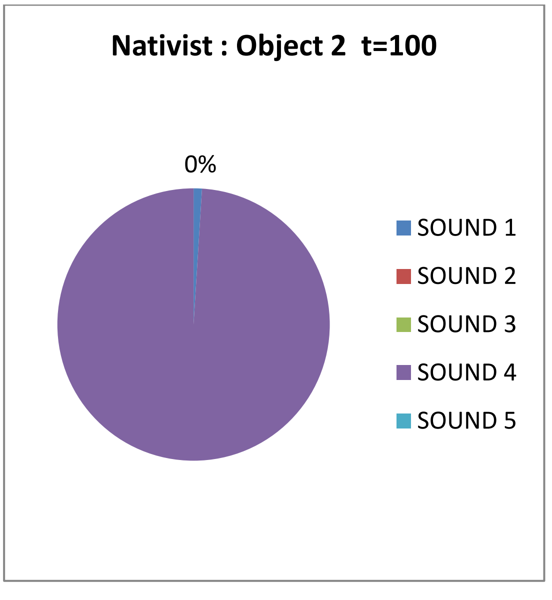
figure 5: *Distribution of sounds for (a) Nativist and (b) Non Nativist model, for 80 iterations*

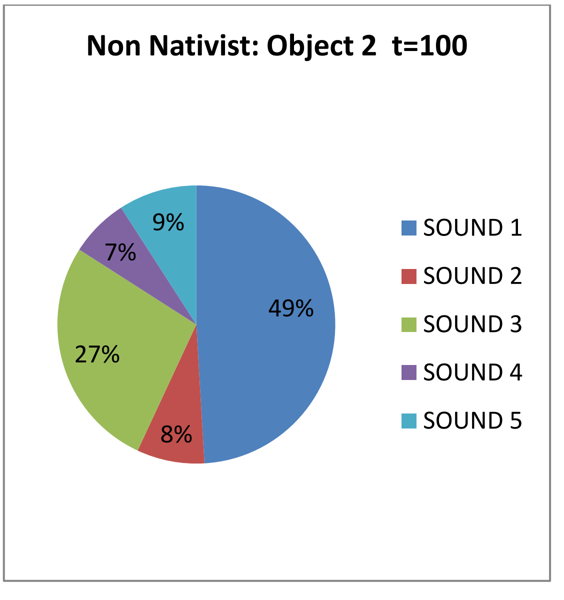
Comments: After 80 iterations the Nativist model has improved its results while the Non Nativist one just started making clearly correspondences between sounds and objects. Both models reach suboptimal solution but Nativist is much faster than the Non Nativist one.

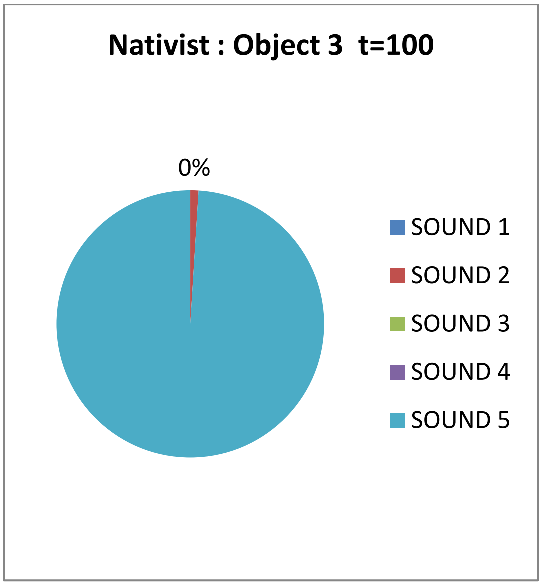
- For 100 iterations the results produced are:

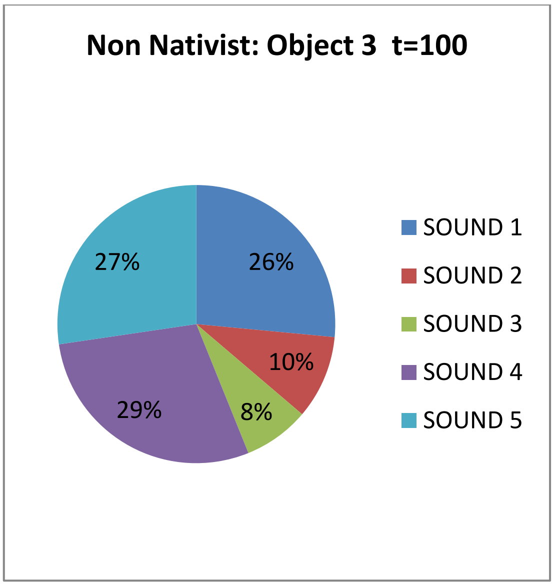


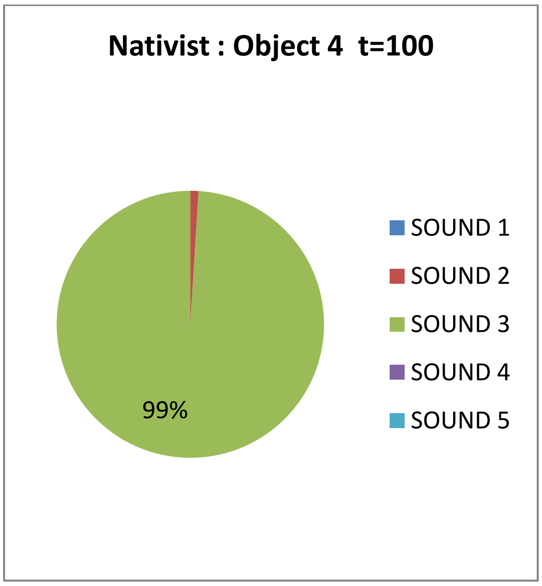


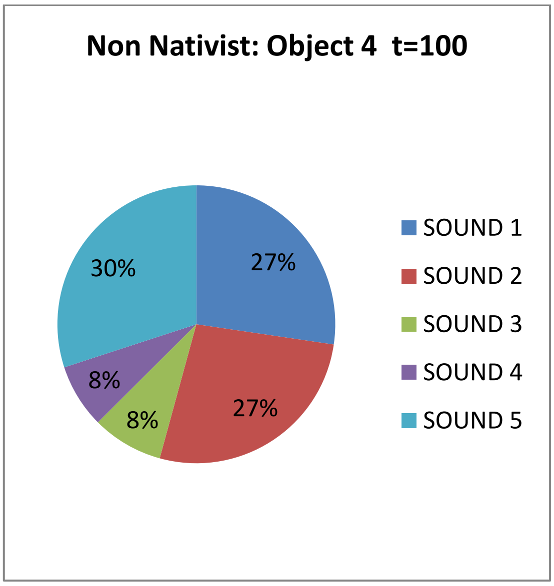


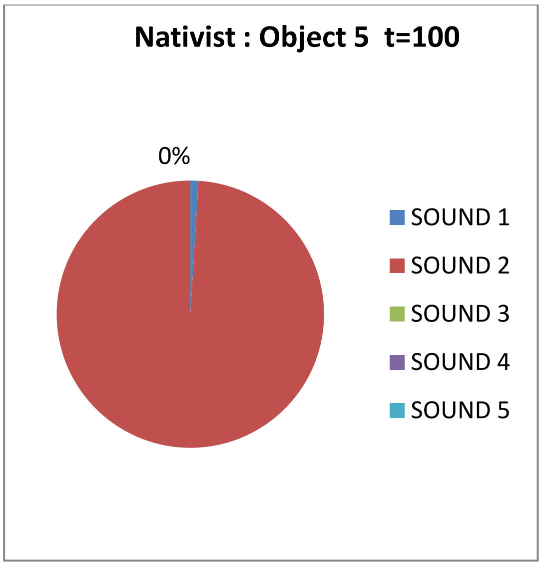



Nativist : Object 3 t=100
0%
SOUND 1
SOUND 2
SOUND 3
SOUND 4
SOUND 5


Non Nativist: Object 3 t=100
27%
26%
10%
8%
29%
SOUND 1
SOUND 2
SOUND 3
SOUND 4
SOUND 5


Nativist : Object 4 t=100
99%
SOUND 1
SOUND 2
SOUND 3
SOUND 4
SOUND 5


Non Nativist: Object 4 t=100
30%
27%
8%
8%
27%
SOUND 1
SOUND 2
SOUND 3
SOUND 4
SOUND 5



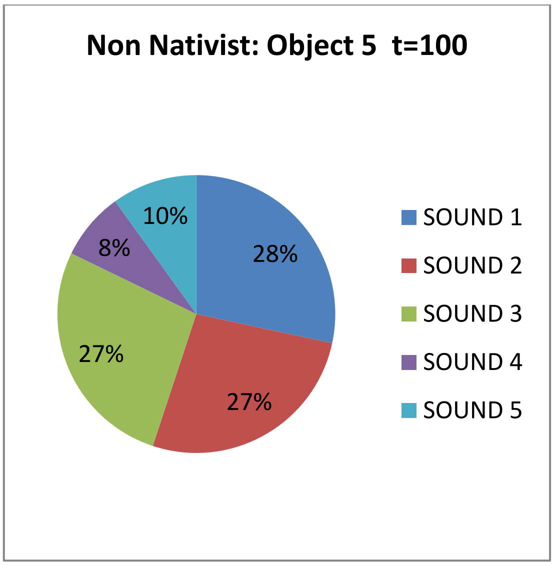
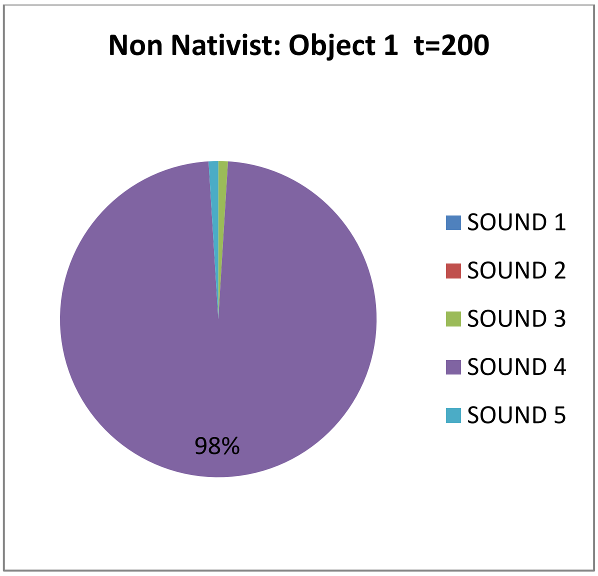


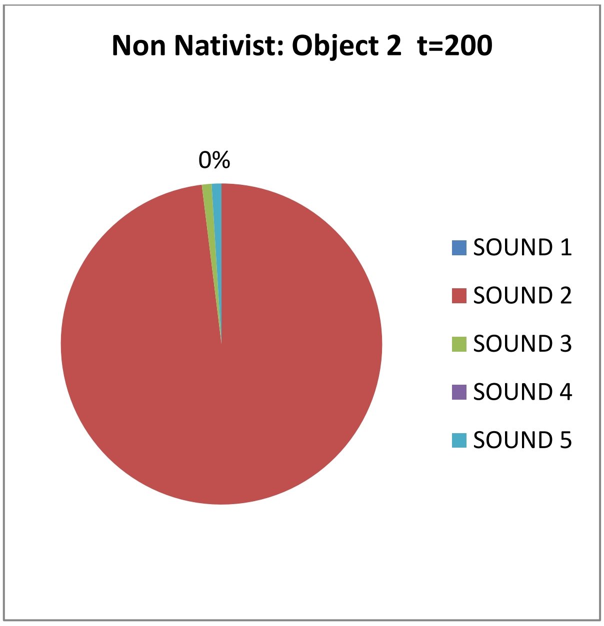
figure 6: *Distribution of sounds for (a) Nativist and (b) Non Nativist model, for 100 iterations*

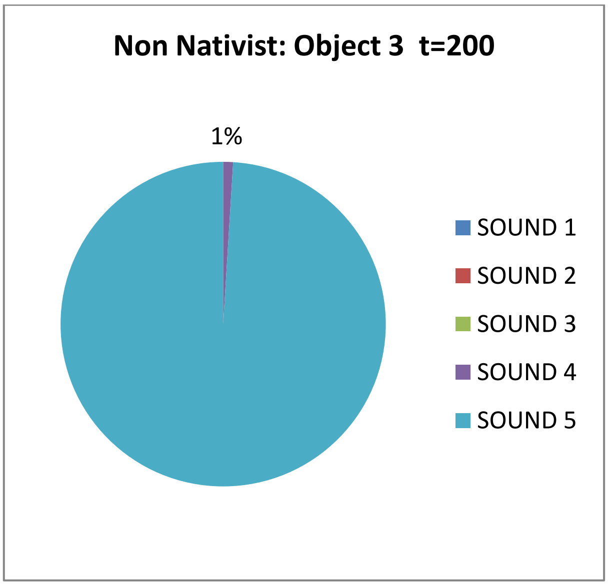
Comments: Figure 4 illustrates that after 100 iterations the Nativist model has dramatically improved its results. We can note that there is only one dominant sound for an object. The Non Nativist model still shows a worse performance. What we can observe is that the Non Nativist model –compared to its performance for 80 iterations- has dramatically decreased the possibilities of the two weakest sounds and seems to choose three dominant sounds.

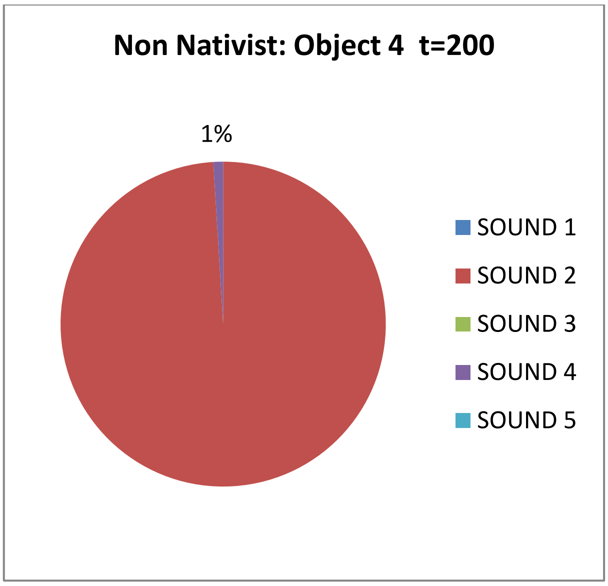
However, at this point, we cannot draw a final conclusion about whether or not the Non-Nativist Model is capable of reaching convergence. For that purpose, we will also present the results of the Non Nativist model for 200 iterations.

- Below there are illustrated the results regarding the choice of the sound, for the Non Nativist model, for 200 iterations. As already mentioned, the Non Nativist model needs more iterations to reach its optimal performance.


Non Nativist: Object 1 t=200
98%
SOUND 1
SOUND 2
SOUND 3
SOUND 4
SOUND 5


Non Nativist: Object 2 t=200
0%
SOUND 1
SOUND 2
SOUND 3
SOUND 4
SOUND 5


Non Nativist: Object 3 t=200
1%
SOUND 1
SOUND 2
SOUND 3
SOUND 4
SOUND 5



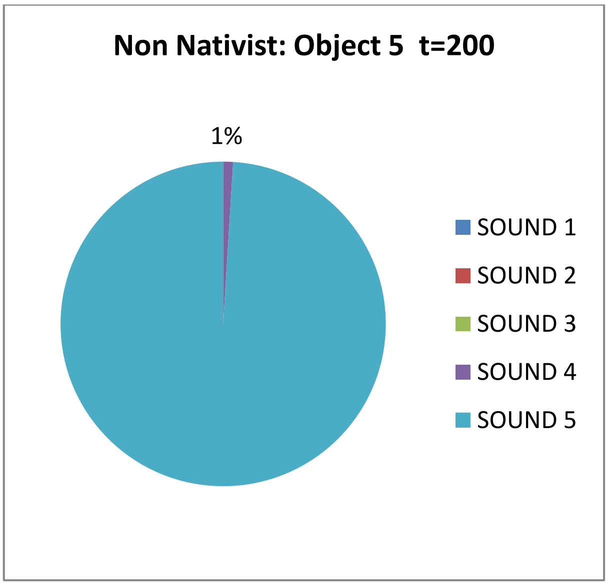


figure 7: *Distribution of sounds for the Non Nativist model, for 200 iterations*

Comments: After 200 iterations the Non Nativist model dramatically improved its results and clearly proceeded to the choice of one dominant sound. Thus, we can now conclude that both models are able to reach the suboptimal solutions, but more iterations are needed for the Non Nativist model to reach its optimal performance.

5.3 Effectiveness of Nativist vs Non Nativist model

At section 5.2 we ended up to conclusions regarding the velocity of each model to reach its optimal performance. At this section we will compare the two models regarding their effectiveness to choose sounds in order to describe objects. More specifically we use our metric that we introduced at our extended model, called P critical. This metric will show how "stronger" is the dominant sound from the other four sounds.

Comparison is made using plots given below, for certain numbers of iterations. Both Nativist and Non Nativist performance are given in the same plot, so that the comparison is done easier.

- For 50 iterations:

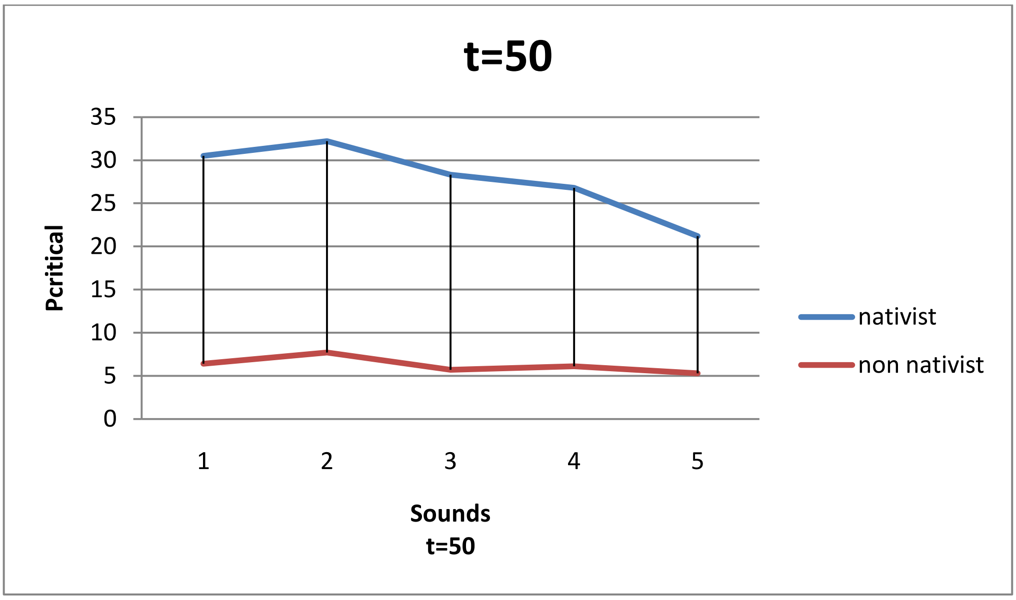


figure 8: *Effectiveness of Nativist and Non Nativist model for 50 iterations*

Comments: for a number of 50 iterations, the Nativist model shows much better performance – in terms of effectiveness- than the Non Nativist one. For instance, for the object 1 the dominant sound that the Nativist model chose is 30% "stronger" than the other sound; while the sound chosen by the Non Nativist model is only approximately 6% "stronger".

- For 80 iterations:

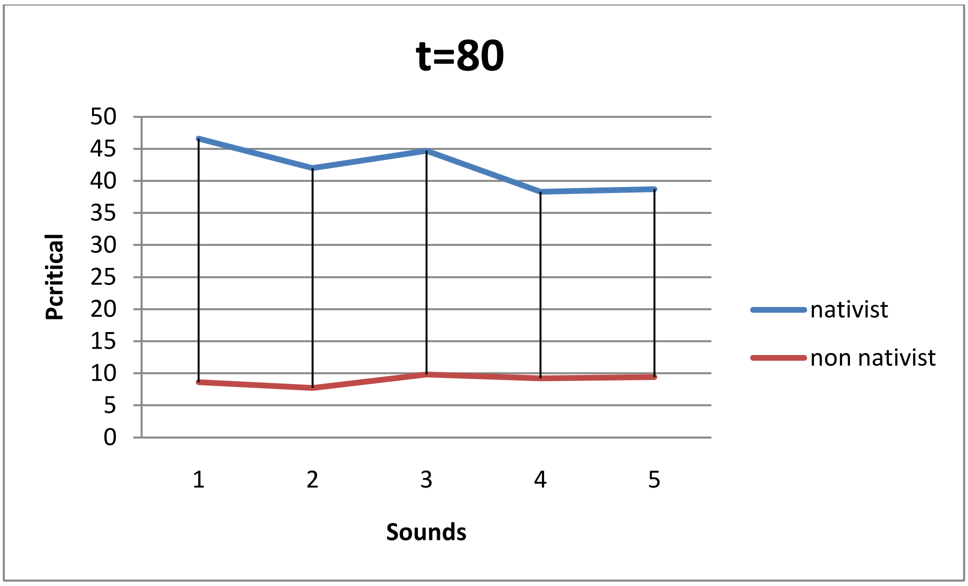


figure 9: *Effectiveness of Nativist and Non Nativist model for 80 iterations*

Comments: for 80 iterations the Nativist model improves its results, so for the object 1 the dominant sound is more than 45% stronger than the other sounds. Non Nativist model has slightly improved its results: the dominant word for the object 1 is approximately 10% stronger than the rest.

- For 100 iterations:

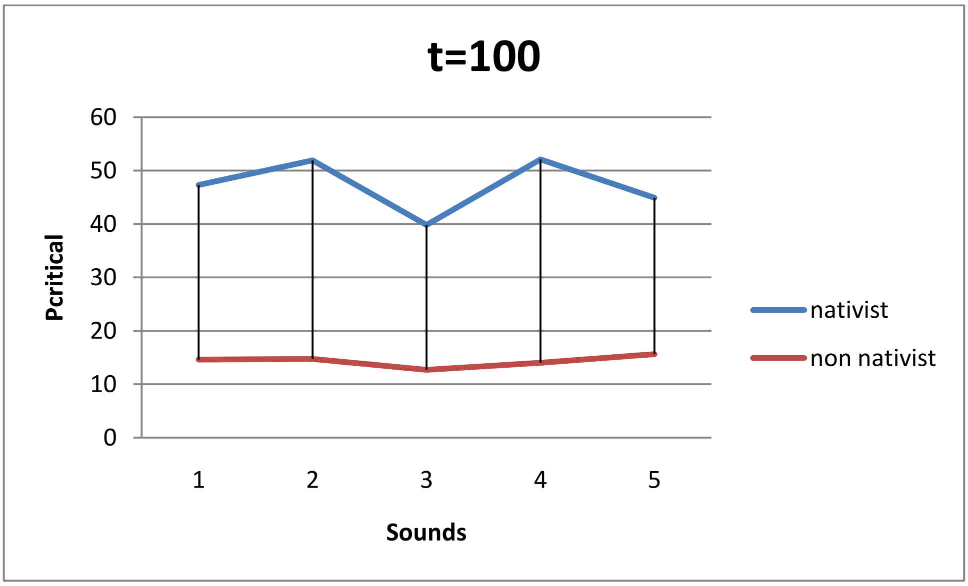


figure 10: *Effectiveness of Nativist and Non Nativist model for 100 iterations*

Comments: after 100 iterations both models keep improving their results. The Nativist model is still better than the Non Nativist one. For object 1, the sound chosen by the Nativist model is almost 50% stronger than the rest sounds. The sound chosen by the Non Nativist model is almost 15% stronger than the rest sounds.

- For 200 iterations:

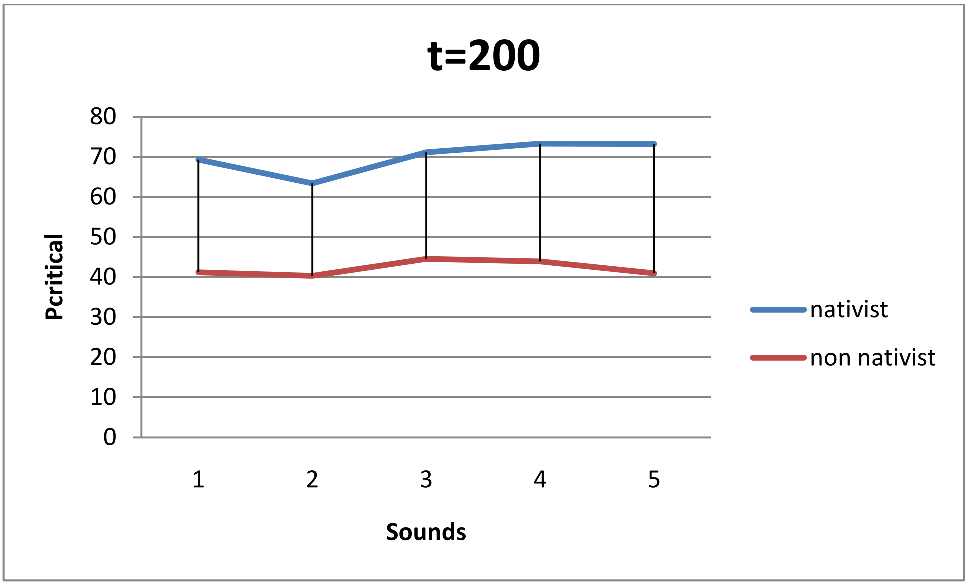
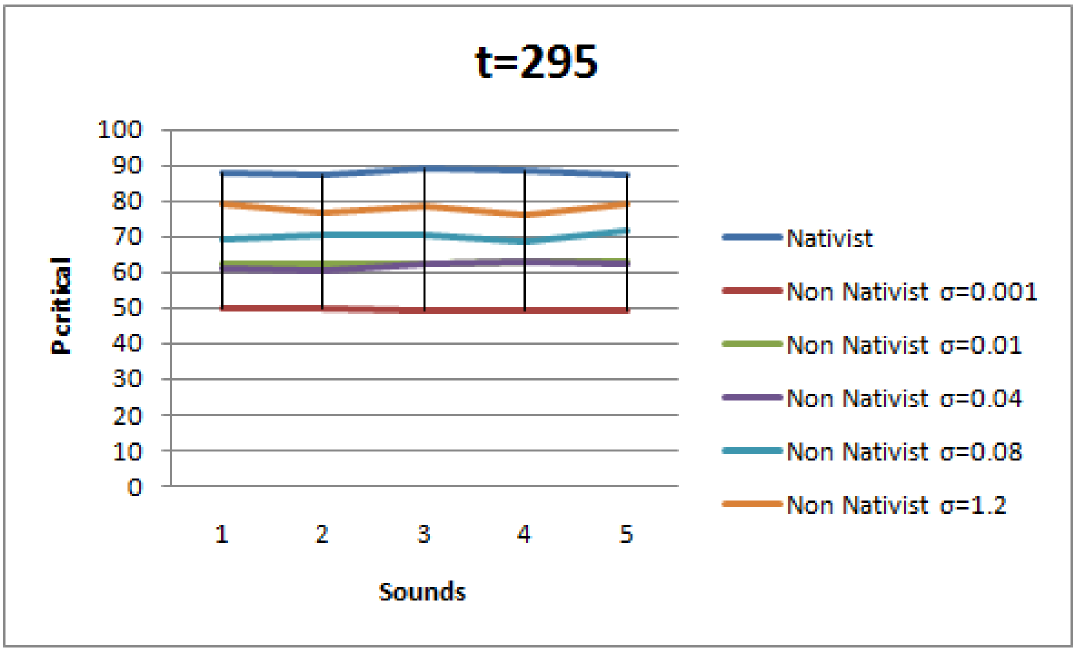


figure 11: *Effectiveness of Nativist and Non Nativist model for 200 iterations*

Comments: As we already concluded from section 5.2, the Non Nativist model needs more iterations to reach its optimal performance. For that purpose we illustrated the effectiveness of both models for a number of 200 iterations. We can now observe that the Non Nativist model is reaching the performance of the Nativist one. The distance between the two lines representing the effectiveness of each model is now much shorter than it was for a smaller number of iterations. For instance, the sound chosen for object 1 by the Nativist model is 70% stronger than the other sounds. The sound chosen by the Non Nativist is 40% stronger than the rest of the sounds.

Finally, the efficiency of the Non-Nativist Model is presented for different values of standard deviation (σ = 0.001, 0.01, 0.04, 0.08, 0.12). It is clearly observed that there is high correlation between the initialization of the input data and the resultant efficiency. Moreover, a tendency of the Non-Nativist Model to approach the efficiency of the Nativist, when it uses greater values of σ, is obviously depicted. This is totally, expectable, as greater values of σ for a Non-Nativist Model means, essentially a "more" Nativist Model.



6 Summary and Outlook

Concluding, the purpose of our work was to quantify the differences between two standard Models for describing the evolution of the Language as the Nativist and Non-Nativist Models. Firstly, we tested whether this two models reach convergence after a certain point of generations and at which rate. The results show that both models reach convergence but the Nativist Model converges with a faster rate. Taking into account that Human Language is a relatively new product, we could assume that our findings are in favor of the Nativist Model as the more likely Model to describe Language Formation. Secondly, we proposed a different metric of calculating the efficiency of the convergence at each stage of evolution. This metric shows that for a little number of iterations (generations), the Nativist Model shows much higher efficiency in comparison to the Non-Nativist Model. However, after a large number of iterations the difference in the efficacy of the abovementioned models becomes shorter. This is explained by the fact that, the Non-Nativist model seems to find optimal solutions as well, but at a slower pace.

After testing our models for different parameters and alternative implementations, we discovered strong correlations between our absolute results and certain parameters, as the initialization parameters, the number of iterations, the reproduction method and the language adaptation procedure. However, the general tendencies described in our report about the two Models, were observed in almost all the cases, less impressively in certain cases; yet still clearly. To be more specific, regarding to the initialization parameters, as the "σ" of the Non-Nativist ,could have a great impact at the results without changing the general conclusions though. Regarding the number of iterations, increased number of iterations lead to stronger convergence and optimized results, but effect in a positive way both models. As far as the reproduction method and the language adaptation section is concerned, they mainly influence the pace of evolution. We implemented the above mentioned sections based on the Nowak's work, making a certain number of intuitive assumptions when it comes to the undefined points.

In the future, further extensions could be made, in order to test the generality of the proposed approach. Additional effort could be put on constructing a rich set of input data, carrying specific isolated characteristics of each Model. Thus, the impact of each single characteristic of the two Models could be defined and safer conclusions about the superiority or not of one model will be drawn. Furthermore, more sophisticated methods of reproduction and adaptation could be utilized in order to simulate more efficiently the evolution procedure in real world.

7 References

[1] M.Nowak, D. Krakauer, "The evolution of language". Proc. Natl. Acad. Sci. USA Vol. 96, pp. 8028–8033, July 1999

[2] N.Chomsky, "Government and Binding". Foris, Dordrecht, 1981

[3] Bloomfield, Leonard (1914). *An introduction to the study of language*. New York: Henry Holt and Company.

[4] FOUTS, R. (1972). *Use of guidance in teaching sign language to a chimpanzee (Pantroglodytes).* Journal of Comparative and Physiological Psychology, 80:515-522.

[5] PREMACK, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences, 1(4), 515--526.

[6] Tomasello, M (2003) Constructing a Language: A Usage-Based Theory of Language Acquisition, Harvard University Press. ISBN 0-674-01764-1 (Winner of the Cognitive Development Society Book Award, 2005)

[7] Chomsky, N. 1957. Syntactic Structures. The Hague: Mouton.

[8] Chomsky, N. 1988. Language and Problems of Knowledge. Cambridge, Mass.: MIT.

[9] Amy Perfors,(2002)." Simulated Evolution of Language: a Review of the Field". Journal of Artificial Societies and Social Simulation vol. 5, no. 2

[10]SAFFRAN, J; ASLIN, D; NEWPORT, E. (1997). Statistical learning by 8 month old infants. Science, 274:1926-28.

[11]GOLDSTEIN, E. (1989) Perceptual development. Sensation and Perception, 3rd ed. Belmont: Wadsworth Publishing Company 327-353.

APPENDIX A

MATLAB CODE

```
%=================================================================
% LANGUAGE FORMATION: THE NATIVIST, NON-NATIVIST DILLEMA (nnnDilemma.m)
%=================================================================

% Afroditi Apostolou, Dimitrios Bolis, ETH Zurich, 2010.

%=================================================================
% Short Description
%=================================================================
% Firstly, the input data are constructed. The program asks from the % user to make a selection for the Nativist or Non-Nativist Mode
% regarding the Language Formation Model. After that, either the
% “rand” or the “random” function is called for creating pseudo- % random % data of Uniform or Gaussian Distribution respectively.
%=================================================================

clc;
clear all;
mode = input('Please select the Mode: Nativist(1), Non-Nativist(2): \n');
clc;

%=================================================================
% UNIVERSES LOOP
%=================================================================
for universe=1:2

display([ 'Iteration ' num2str(universe) ])

%=================================================================
% Parameters Selection
%=================================================================% Define in which degree are the communities restricted
comunFactor = 1;
% Define the number of the proto-languages
proto = 300;
% Define the number of generations
gen=200;

%=================================================================
% Initializer
%=================================================================
W = ones(proto,1);
PopAvg=zeros(5,5,gen);
```

```
%=================================================================
% Proto-Language Type Selection
%=================================================================
if mode == 1
  % Nativist
  P = rand(5,5,proto);
  Q = rand(5,5,proto);
elseif mode == 2
  %Non Nativist
   P = random('Normal', 0.5, 0.01, 5,5,proto);
   Q = random('Normal', 0.5, 0.01, 5,5,proto);
end

%=================================================================
% GENERATIONS LOOP
%=================================================================for t = 1:gen

   %=================================================================% Initializer
%=================================================================
   F = zeros(proto,1);
   parents = ones(proto,gen);

%=================================================================
% Cost Function
%=================================================================
   for m = 1:proto

      for n=1:proto
         for i = 1:5
            for j = 1:5

                if n~=m
                  Fsingle = W(m)*0.5*(P(i,j,m)*Q(j,i,n) + P(i,j,n)*Q(j,i,m));
                  F(m) = F(m) + Fsingle;
                end

            end
         end
      end

   end

%=================================================================
% Weights
%=================================================================
   parents(:,t+1) = F;
   F          = parents(:,t).*F;
   W1          = (F-min(F))/(max(F)-min(F));
   W          = round(100*W1);

%=================================================================
% Language Adaptation
%=================================================================
    PopAvg_temp=zeros(5,5,proto);
    for v = 1:proto
```

```
        [maxV maxI] = max(P(:,:,v),[],2);
        for c=1:5
            PopAvg_temp(c,maxI(c),v) = W(v);

% Perform the intra-community feedback with certain probability
            comun = double(rand < comunFactor*W1(v));
            P(c,maxI(c),v) = (1 + comun*0.01)*P(c,maxI(c),v);
            Q(maxI(c),c,v) = (1 + comun*0.01)*Q(maxI(c),c,v);
        end

% Indicates the amount of the people having maxima in each P's element
        PopAvg(:,:,t) = round(PopAvg(:,:,t)+PopAvg_temp(:,:,v));

    end

%=================================================================
% P-Mean Calculation
%=================================================================
for v = 1:proto
    Ptemp(:,:,v)=W(v)*P(:,:,v);
end
Pavg(:,:,t) = sum(Ptemp,3)/sum(W);
Pavg=round(Pavg*10)/10;

%=================================================================
% Elimination of Weak Subjects
%=================================================================
    for u=1:1
    [Wmin I] = min(W);
    F(I)=[];
    W(I)=[];
    P(:,:,I)=[];
    Q(:,:,I)=[];
    parents(I,:) = [];
    end

    proto = length(F);
end

%=================================================================% Mean
Distance from the Dominant Sound in Each Universe
%=================================================================
Pavg_temp(:,:,1) = Pavg(:,:,t);
Pmax = max(Pavg_temp(:,:),[],2);
PmeanRest = (sum(Pavg_temp,2) - Pmax)/4;
Pcr(:,universe) = (Pmax - PmeanRest)./Pmax;

display('The mean distance from the Dominant Sound, for this Universe is (%): ')
display(num2str(100*Pcr(:,universe)'))

end

%=================================================================
% Total Mean Distance from the Dominant Sound
```

```
%==================================================================

%Pcritical Total
Pcr_total = round(1000*mean(Pcr,2)')/10;
display('-----------------------------------------------------------')
display('The total mean distance from the Dominant Sound, for the selected mode is (%): ')
display([num2str(Pcr_total)])
```