Alessandro Lopopolo

6100775

Research MA Linguistics

POPULATION STRUCTURE AND ACQUISITION DYNAMICS INFLUENCE LANGUAGE COMPLEXITY

Simulating the variation in complexity and population structure using the Language Game model

Abstract: this thesis investigates the relation between population structure, type of acquisition and the level of complexity of the language. The hypothesis follows from the observation of Kusters (2003) and Lupyan and Dale (2010) that suggests that the more open a population is, the less complex its language will be. The assumption is that the difference is given by the type of acquisition involved in closed or open populations. In order to confirm these observations, the present work will adopt a computer modeling approach, running a series of experiments based on Steels Language Games (1996, 1997).

1.Introduction

In the recent years, the assumption that human languages are equally complex in structure has begun to be questioned (Sampson et al. 2011, Kusters 2003, Dahl 2004). The availability of a larger amount of data from typological studies provided a clearer picture on the great range of variability in size and structure of the world languages phonological and syntactic repertoires (Pellegrino et. Al 2009, Perkins 1992, McWhorter 1998, 2001, Everett 2005). One innovative study was recently produced by Lupyan and Dale (2010). On the basis of statistical analysis on data taken from the World Atlas of Language Structures (WALS), the study showed that there is a non-trivial correlation between the morpho-syntactical complexity of languages and the structure of the community of their speakers. This not only corroborates the challenge to idea of linguistic uniformity, but gives more prominence on the influence of language use, environment and social dynamics in shaping natural language.

The present work will investigate the idea that language complexity is shaped by the structure of the community of speakers and in the last analysis by the processes of acquisition endowed in the population dynamics. It will move from the observations sketched above and from the idea that language is a complex adaptive system (Ellis et al. 2009) to a computer modeling approach based on the work of Luc Steels and his team (1996, 1997).

This thesis is structured as follow. Section 2 lays down the hypothesis that language is shaped by learning and that the population dynamics around L1 and L2 acquisition are responsible for cross linguistic variation in complexity. The same section contains the research question that guides the experimental work. Sections 3 and 4 contain the theoretical foundations of the thesis: section 3 defines language as a complex adaptive system and argues against innatist theories, whereas section 4 will present a discussion on the concept of linguistic complexity. The same section gives a definition of complexity that can be easily applied to computer models such as the language game. Section 5 is devoted to the methodological choices adopted to investigate the hypothesis: the use of computer simulations and the description of the language game paradigm. Subsequently, section 6 presents two series of experiments consisting of computer simulations of linguistic communities of artificial agents under three experimental conditions: populations whose language is not subject of any process of learning, populations where learning is only by newborn agents (L1 acquisition) and populations open to foreign contacts and where therefore L2 acquisition is allowed by the social structure. Finally, section 7 concludes the work presenting the sum up of the previous section with regard to the question posed in section 2. The final part of this work is also devoted to some theoretical and methodological entailments regarding the experimental design.

2.Research questions

The relation between population structure and complexity of the linguistic system has recently been proposed by a series of works, among which there are the researches of Kusters (2003) and Lupyan and Dale (2010). Both these works formulated a strong claim on the basis of empirical observation: the more open and dynamic a population is, the less complex its language (or a subsystem of it) appears to be. Kusters talks about type 1 and type 2 populations. Type 1 corresponds to those communities where contact with external populations is limited or inexistent and that the only focus is on L1 acquisition. Type 2 populations, on the other hand, are open communities prone to language contact scenario where adult foreign learners are a significant presence. A different terminology and formulation are employed by Lupyan and Dale. They talk about two distinct niches in which languages are learned and used: exoteric and esoteric. Where exoteric corresponds loosely to type 2 and esoteric is equivalent to type 1.

The idea carried on by this line of research is that on a general level complexity is a function of population dynamics and that the actual process behind complexification or simplification is language acquisition, in particular second language acquisition by adult learners. Lupyan and Dale affirm in agreement with Trudgill (2001) that the 'lousy language acquisition abilities of the human adult' work as a filter that simplify the target language when its community is under pressure of immigration from external populations. This is not an isolated claim: in fact it is substantially shared by researchers in the field of language acquisition, who have spent quite some effort in trying to figuring out what are the actual processes that determine a simplification (or complexification) in open and dynamic population on the one hand and in closed and static ones. One convincing answer is contained in the works of Weerman (Weerman 2010) and of Meisel (2011). Weerman, on the basis of the study of the linguistic performance of young foreign learners of Dutch as second language, concluded that the main 'culpable' of the process of change are adults L2 learners that are the typical actors in language contact scenarios. Language contact is naturally a phenomenon that is contemplated in open and dynamic populations, therefore one may conclude drawing from Weerman acquisition is a player in keeping the level of complexity low and that L2 in particular is a major factor of simplification. It appears that both theoretical speculations and typological works introduced above points all to two interlaced factors influencing complexity: the population status, isolated or open to contact, and language acquisition where the age of the learner seems to determine the level of complexity of the learned language.

Therefore, the research questions that the thesis will try to address are: does the structure of the population influence the complexity of the emerging language? And in particular how does a language emerging from a population without L1 and L2 mechanisms differ from one emerging from a population that does implement first language acquisition and language contact? In other words what is the role of language acquisition in directing the trend of complexity of an emerging linguistic system? The hypothesis is simple: language is a product of cultural evolution, and cultural evolution takes the shape of horizontal communication (Language Game) (Steels 1996) and inter-generational transmission (Learning) (Kirby 1999) – in absence of the latter the language will evolve as a highly redundant and complex system. Therefore the effect of language acquisition is to keep the complexity of the system under a certain level and to filter out redundancy, ambiguity and synonymy. As age and population dynamics are concerned, the simplification effect is more evident when adult L2 acquisition takes place, as in cases of population contact, and confirms the empirical evidence that languages spoken by open communities tend to be simpler than languages spoken by isolated ones.

3.Language as a Complex Adaptive System

The idea of language adopted in this work starts by considering a self-evident aspect of human linguistic behavior: language has principally a social function, it cannot be fully understood outside the social dynamics for which and through which it emerged. This idea moves severely apart from generative approaches to language, that tend to construct static and abstract representation of language, disregarding notions such as the distinction between individual and communal language, the role of learning in shaping grammar and the interaction between cultural and biological evolution. That is why it is reasonable to agree with the proposal that language, and especially language evolution, is best understood as a complex adaptive system (Ellis et al. 2009). This is reinforced by the assumption that theoretically unrelated processes, such as acquisition, use and change are aspects of the same system and that cannot be fully understood in isolation. Therefore, if traditional generative linguistics looks for descriptive rules or structure that wishfully are as accurate representation as possible of the competence of an idealized speaker/hearer, the aim of CAS or self-organization approaches to language is to observe and explain how and why such a complex behavior emerges and stabilize. Even if it may sound somehow simplistic, the difference is reducible to an opposition between description and explanation – such difference easily derives from the fact that generative linguistics tends to be a top-down approach to language whereas the other is purely a bottom-up theory.

Language as a CAS involves the following key features: 1) the system consists of multiple agents interacting with each other; 2) the system is adaptive, that is, speakers' behavior is based on their past interactions, this is because the agents modify their internal representation of the language on the basis of acquisition, categorization and other basic cognitive processes and are sensitive to frequency; 3) the agents' behavior is influenced therefore both by their internal design (basic cognitive processes, memory, linguistic framework etc…) and by the social environment in which they are immerged. This rules out the idea of an idealized speaker/hearer as the focal point of linguistic research simply because the process of language emergence from the interaction of independent agents does not guarantee uniformity between the single competences of the single agents. An agent is not necessarily representative of the state of the language spoken in the community.

Agents are defined as simplified speaker/hearer's endowed with basic domain general cognitive processes such as the capacity for social interaction (sociability), ability of applying joint attention, of extracting patterns of regularities from perceptual objects and signals and to imitate the behavior of other agents. It is important to stress the domain generic nature of such processes, since one of the assumptions of language as CAS theories is the rejection of innatist theories of the language faculty and origin of language in general. Given these properties defining the particular behavior or routine of the agent, another defining aspect of the framework is the local level of interaction, that is the fact that the system as a whole is not directed or coordinated by any omniscient factor, but that the dynamic on the global level is the product of the local interaction of the single agents. That is to say that whenever a CAS display a pattern of regularities, such as those observable in language or in biological systems, those are not ascribable to a central intelligent design, but are the byproduct of the work of simple agents that collaborate and that are tendentially unaware of the global pattern. Agents, besides being ignorant of the emergent effect of their behaviors, are also especially as far as language modeling is concerned, unaware of the internal configuration of their counterparts – in other words, when a hearer perceive the signal produced by a speaker and tries to guess the intended meaning and act as a consequence, cannot access the internal linguistic representation of speaker that generated the target signal.

This definition language as an emerging phenomenon from the local interaction of agents as basic cognitive entities entails a series of theoretical and methodological corollaries: the distinction between idiolect and communal language, the differences on the individual level vs. the growth of universal tendencies, the dynamic nature of the system, and the dependencies on the environmental factors.

Language exists and must be treated on two distinct levels: individual and global. The first is the level of the idiolects of the single agents. Each agent picks up its own language from use and interaction with its counterparts and the environment – this means that different agents may have had different experiences and thus partially different languages. The global level corresponds to the communal language of the community and loosely coincides with the similarities between the individual idiolects of the agents composing the community. They are the two faces of the same coin, in fact idiolects are formed through the interaction between the single agents in the communal language, and the communal language emerges as the result of the interaction of the single idiolects – one aspect cannot exists without the other. These similarities are not guaranteed to exists, however, from a naively functional perspective we can say that higher is the level of communality, the more stable the communal language, the more efficient are the communication acts between the agents. The emergence of the communal language is a process of self-organization of a global behavior – similar to the patterns observable in economy, bird flocks and fish schools and so on.

Even if the agents interact in the same environment, it is likely that their idiolects differ from one another. This intrinsic diversity is the natural consequence of the individual's unique exposure and experience of language use (Bybee 2006). This, however, does not rule out the existence of communalities both on the individual (that is between single idiolects) and global (that is between distinct linguistic systems) level. In the first case, it is a basic requirement for communication and in particular mutual intelligibility between different agents. The second represents the case of so called linguistic universals and that can be explained in terms of general cognitive abilities of the agents such as the one postulated above (sociability, joint attention, pattern extraction, imitation)

The assumption that language is shaped by use, and that it is embedded in the social structure of the agents' community justifies the fact that it is in constant change both on global and individual level. Agents constantly update their idiolects as they encounter new linguistic forms that may be introduced by new coming agent in the community, or can create new forms themselves whenever they experience changes in their environment like new objects, situations or relations. This is quite a relevant point, especially for the definition of complexity that will be treated in the next section. Language has the nature of a dynamic system, whose regularities represent temporary states of equilibrium (i.e. a lexicon remains stable as far as the environment or the population do not change), and it is not governed by external rules or forces aimed to preserve an hypothetical state of static equilibrium. In other words, language is an open system that continues to change and adapt dynamically to the variation in the environment that supports it – whereas a closed system will reduce to a stable state or equilibrium (Larsen-Freeman & Cameron, 2008).

As a consequence of what sketched above, the dynamic nature of a linguistic system has to be ascribed to a concomitance of environmental factors. Keeping the basic behavior of the agents invariant, the influencing factors on language change and evolution must be the structure of the community and its internal dynamics and the state of the world described by the language. This defines the linguistic system as sensible or dependent on environmental factors, social structure first of all. Since language emerges through the contact between agents, the frequency of such interactions between different agents carrying their own experiences and language has a crucial effect in the process of language change and variation. Therefore formalism such as social networks (Milroy, 1980) and statistical studies on the composition and degree of openness of a community are valuable variable for the study of the behavior of language systems on a global level.

The present work aims to study the interaction between population dynamics, especially the ones involved with L1 and L2 acquisition in shaping the complexity of the linguistic system. Next section will be devoted in finding a suitable definition of complexity that keeps an eye on the distinction between idiolect and communal language. On the other hand, the second part of this thesis will be devoted to the computational modeling of the problem, with an introduction and justification to this methodology and a description of the Language Game framework. The model will be implemented in a series of experiments where the ideas of language as CAS and self-organizing system will be realized through the creation of populations of artificial agents that interact with each other in a specified environment under the form of Language Games (Discrimination and Naming Games, to be more precise).

4.Linguistic Complexity

The notion of linguistic complexity has become more and more a hot topic in the last decades. Many researchers and research groups have proposed their own definitions of complexity and methodologies to measure it empirically. However, this topic has often stirred up a hornet's nest, so to say, whit implications that not always have to do only with linguistics or cognitive sciences tout court. This chapter will first give a brief review of the controversies around the idea that languages may have different levels of complexity. Then the focus will be on formulating a simple and clear definition of complexity that can be applied to the framework of language games and that can help studying the interaction between language acquisition, population dynamics and, of course, evolution of linguistic complexity. For this last task, however, the work of Kusters will serve as a good starting point, from which the present definition will significantly part.

4.1.Controversies

The notion of linguistic complexity has always been and still is controversial and troublesome. It is controversial because it raises questions about the very nature of natural language, especially the definition that the generative and innatist school gives about it. This controversy generates some interlaced assumptions. First of all, according to those who negate the possibility of differentiate languages according to their intrinsic complexity, all languages are in the last analysis of equivalent complexity because are endowed by the same innate machinery called universal grammar or language faculty. This element is domain specific, of biological nature and shared by all the members of the human species – in some way, claiming that a language is more complex than another amounts to claim that the speaker of those languages have different cognitive capabilities, at least with regard to language. In second stance, complexity can be observed in different subsystems of one language. People may even recognize the fact that languages can have a more or less complex morphology or a more or less complex phonology, but the final stand is against difference n complexity on the global level. Aitchinson (1991) for instance give voice to this believe and affirms that the differences in complexity within language subsystems are likely to be compensated with trade-offs between for instance a simple phonology and a more complex morphology. In a few words, the idea of complexity diversity is still rule out and limited to language subsystems.

Whatever forms this opposition may take, looking at some textbook for undergraduate students of linguistics, it is not uncommon to read positions that a priori rule out the possibility for languages to differ in terms of complexity (McMahon 1994, Aitchinson, 1991). A similar position can be found also in studies on creole languages (DeGraff, 2001). In particular the position of creolists discloses one of the origins of the opposition to the notion of complexity variability. Too often this notion has been cognate of Eurocentric stances on cultural and anthropological comparisons. As Kusters (2003) points out, for a long period of time and especially in the XIX and the first half of the XX century complexity was a correlate of prestige, especially in the human sciences. A highly complex object was considered more valuable. Therefore a European like society, that was deemed more complex, was arguably superior to the organization of some tribal communities of Africa or South America. That was also applied to linguistic descriptivism leading to consider languages with complex morphosyntax such as Latin, Sanskrit and German as superior languages of superior cultures. Although these positions have, officially, been excluded from the linguistic debate, they still work as prejudicial biases against the notion of complexity diversity.

Nonetheless, this debate exceeds the scope of this work. The starting point of this thesis is that language is a self-organizing system, that can be studied only inside the community of speakers that use it and that there is no need for postulating the existence of any innate language specific machinery in order to justify its features. It follows that different languages differ with regard to complexity, and that there is not any a priori reasons to deny this proposition. Keeping in mind what has been said about language as a CAS in the previous section, it is reasonable to say that an eventual variation between languages is not due by individual differences between speakers, but to difference in the social and natural environment that saw their emergence. In other words, what influences complexity are not the cognitive capacities of the speakers, which are assumed to be equal, but the conditions that shapes the way the language is used and transmitted, i.e. the social structure of the community and the environment it is set in.

4.2.Defining Complexity

Complexity is not only controversial as a notion, it is also difficult to define and, in the last analysis, to measure. Can it be measured on the language as a global phenomenon, or is it more proper to consider individual differences in complexity? And also, is it an absolute property of a language or is it necessary to define it in relative terms, for instance in terms of learnability by a particular agent or efficiency with regard to certain communicative scenarios?

Leaving aside these questions, for the moment it suffices to consider how a definition of complexity should look like in order to be applicable to a computational model. A definition of complexity must obligatorily be simple, formal, and of course measurable. These requirements are not mere theoretical preferences, but sine qua non conditions without which it will not be possible to monitor the behavior of the model with regard of complexity. A furthermore characteristic is the applicability to an easy abstract representation of natural language, such as the one adopted by Steels in his language game and talking heads experiments.

In the past decades the number of definitions of complexity for natural language grew significantly. The easiest and possibly more intuitive approach is simply quantitative (McWhorter 2001, Dahl 2004). A language, or one of his sub-systems, is more complex as more elements composing it can be distinguished. For instance, the nominal inflection of Russian is considered more complex than the one of English simply because it encompasses a larger paradigm, more formal features and therefore more morphemes and so on. This quantitative approach can be refined by having recourse to Information Theory. For instance Dahl (2004) adapts Kolmogorov's complexity (Kolmogorov 1965, Mitchell 2009) that states that the measure of the complexity of a linguistic structure is the length of the description of that structure; the longer the description, the more complex is the structure in analysis. Kolmogorov's formulation is also known as algorithmic entropy – it refers to the fact that a linguistic signal, or more generically a string of signs, can be described by the algorithm that generates it. And the measure of complexity is given by the shortest algorithm that generates the target string. Here again, the longest is the shortest defining algorithm, the more complex is the generate string. These informational theoretical formulations of complexity are undoubtly very appealing. However they will not be applied directly in this work because they do not fit directly the formalization of language adopted here. They presuppose a conception of linguistic system more articulated than the one introduced for the study of the evolution of a common lexicon and call the attention on the morphosyntactical component of the language. However, it is not excludable that a further development of this study, whit a more explicit eye on syntax and compositionality, may resort to these formulations.

4.3.Complexity as measure of Economy and Transparency

In this thesis, complexity is defined in quantitative terms as well – the largest is the emerging lexicon, the more complex is the linguistic system. Nonetheless this quantitative notion needs to be integrated with a measure that describes the internal state of the idiolects of the single agents that compose the linguistic community. In order to do so, two independent concepts will be introduced, both of which are inspired by theoretical linguistics: Economy and Transparency.

Economy represents the purely quantitative side of complexity. Given a linguistic system evolved under the pressure of use in communication by a linguistic community, such system will be composed by a certain number of signs, each of which is an arbitrary mapping of form and meaning. Besides this, the linguistic system guarantees a level of communicative success roughly defined as the probability of a speaker of being understood by its hearer while speaking this particular language in a given context. Said so, hypothesizing an arbitrary high level of communicative success, the linguistic system that has a larger lexicon will be less economical.

Transparency is a cumulative measure of the state of the linguistic community. The intuitive definition states that a language with a higher number of ambiguity and synonymy is more opaque and more complex. Therefore a language is transparent if and only if all its forms express one and only one meaning, and, vice versa, all its meanings are expressed by one and only one form. Another definition of Transparency is 1-to-1 principle (Andersen 1984). Therefore transparency is an ideal state where such a 1-to-1 relation between meanings and forms is attested, and complexity is defined against as a deviation from this state. The measure of mismatch, defined in details in the next section, is a way to quantify how much a particular idiolect (language of an individual agent) parts from this ideal. The level of mismatch of the communal language will be the average distance of all the idiolects of the community from the state of perfect Transparency.

The concepts of Economy and Transparency are inspired by the work of Kusters (2003). In his theory, the above mentioned principles, are intended as constraints working on limiting and shaping the structure of morphological systems observed in nature. Such an approach is implemented through the adoption of the machinery of constraints, candidates and evaluation typical of Optimality Theory. The approach employed in this thesis, although starting from Kusters’s terminology and assumptions, tries to look at complexity from a significantly different angle. First of all, the focus is not on morphological paradigms alone, but is way more abstract and embraces a more functional and general definition of language. What is more different, however, is the way these principles are treated. Kusters uses these principles as requirements, that is, tendencies that are desirable for an efficient linguistic system. In fact, Economy, as intended by him, is a requirement on the number of category expressed in morphology and it favors those structures that are simpler than others. Same story for Transparency, which demands that the relation between forms and meanings is as close to the ideal 1-to-1 relation as possible, and that cases of synonymy and ambiguity should be absent from an efficient system of communication. As natural, it follows that Economy and Transparency often work one against the other, and a stable language emerges when equilibrium between the requirements of one and those of the others is attested.

Even though Transparency as defined above cannot be directly applied, some works on the evolution of vowel systems (Liljencrants and Lindbom, 1978, de Boer 2001), therefore oriented to phonology rather than morphosyntax, contain the concept of distinctiveness that can yield some interesting comparisons whit the present case. De Boer observed the emergence of repertoires of shared vowel-like sounds within a population of agents following the Imitation Game paradigm (cognate to the Language Games employed here and introduced in the next section). It is a clear case of self-organization driven by the mechanisms of articulation, perception and recognition between couples of interacting agents. The configurations of these repertoires display patterns comparable to natural vowel systems and are characterized by a tendency to spread the sounds on a vowel space defined in terms of the first three vocalic formants. Dispersion on the vowel space is interpreted as a phenomenon of self-organization that enhances distinctiveness. As there is no meaning in de Boer’s experiments, and therefore no form-meaning mapping, Transparency is not applicable to this model in a strict sense. On the other hand, using a broader interpretation of transparency, one could say that a more dispersed system is more transparent, in the sense that there is less potential for confusion (in the presence of noise) and therefore there is more transparent mapping between what the sender sent and what the hearer heard. Therefore, here Transparency is definable as the state of 1-to-1 mapping between archetype and vowel, both from the perspective of production and perception. In partial opposition to de Boer’s work, Liljecrants and Lindbom older experiments treat exactly the same linguistic subject, the evolution of vowel system, and they too consider the phenomenon of dispersion on the vowel space as an important feature of the evolutionary process. However, they try to reproduce it using a mechanism of optimization. Given a measure of dispersion based on the distance of the vowels on a bi-dimensional space, the system tries to optimize distinctiveness by maximizing dispersion. In this way distinctiveness is not obtained as product of self-organization through the interaction of elementary linguistic agents, but directly implemented in the form of an energy function. Said so, Kusters treats Economy and Transparency as external forces on the linguistic systems and implements them in OT style formalism. The distinction between Kusters’ work and this thesis is similar to the distinction between Liljencrants' and Lindbom's work and de Boer’s work on vowel systems: the first optimizes directly, whereas in the second, it is the result of self-organization.

In a framework that considers language, and language evolution, as a matter of self-organization, Economy and Transparency are mere measures of complexity and efficiency of the system at study. They are not, so to speak, instincts inherent in the linguistic behavior of the speaker and of the hearer, sort of maxims that guides (or should guide) their interactions and shape the language code used for that purpose. They are simply way to keep tracks of the evolution of the system by measuring for instance the size of the lexicon or of the grammar and the level of mismatch between forms and meaning at a particular point in time. It is worth to recognize, that this change of prospective does not pervert the nature of the notion of Economy and Transparency, but has the effect of making them not normative, easier and ready to be applied to a modeling study of the evolution of linguistic complexity.

4.4.What seems to influence Complexity?

In the present framework, the necessity to communicate about objects or states of the word, the process of signal creation by linking semantic categories to linguistic signals, and the process of learning by interacting are the forces that guide the emergence of a linguistic system. On top of that, the assumption is that the structure and dynamics of the population by which the language is created influence the Economy and overall Transparency of the language itself and, as a consequence, its complexity.

In the next section, where the computational model of Language Game will be laid out in details, I will describe the abstraction adopted as representation of language. Here it will suffice to retain a general definition of language as a function mapping forms and meaning. The model is composed by three interlaced elements: the population, the universe and the language as a collective construct. The population is simply a set of agents, each of them characterized by a behavior that determines the way it relates with the environment. Such behavior is shaped around the discrimination and naming games proposed by Steels (1996) and concerns the way agents perceives the objects in their surroundings, create words to communicate about them with other fellow agents and learn from the interaction with the environment, where environment means objects and agents alike. Then there is the universe, that consists of the objects perceived by the agents and about which the agents communicate. As third and more abstract components there is the language. Each agent builds up its own idiolect, in this case a set of lexemes or Sussurean signs, to communicate with the other member of the population. The process of creation and learning ensure a certain level of uniformity between the idiolects of the single agents – and this is what makes the language a collective construct.

Among these three components, the environment will be considered as invariant – the number of perceivable objects and their structure and values will not be changed in the course of the simulation. The population, on the other hand will be the independent variable that will be manipulate in order to study the effect of population structure on the emergent language, which in turn will be the dependent variable, so to speak. The variable of the population will be on the social and not on the individual level. In other words, The number of agents and composition of the population will be subject to manipulation, whereas the basic structure of the individual agents and its behavior will be left untouched and free to adapt under the principles of the language game implemented in the model. Having said so, language is the focus of this thesis and complexity will be measured on it.

It is possible to claim that the notion of linguistic complexity as introduced here is nearer to the notion of absolute complexity (Miestamo et al. 2008, Kusters 2003) rather than cost/difficulty-based relative complexity. This means that the measures employed in the present model do not take into consideration the performance of the individual learning agent directly as an indicator of complexity. To put it differently, Economy and Transparency are properties of language as a collective construct.

5.The Language Game Model

This section will define in details the linguistic formalism adopted in the thesis and already sketched in the introduction. It will also introduce the Language Game paradigm on which the architecture of the model is based and the experiments are implemented. Importance will be given to the architecture of the agents and in particular to their behavior that is, on the abstract level, based on simple cognitive processes such as joint attention, imitation and pattern extraction, and, on the computational level on the procedures described by Steels about discrimination and naming Games (1996, 1997). This description will be preceded by a justification for the use of a modeling methodology. The description of the agents and of their behaviors will be accompanied by concrete examples in pseudo-code.

5.1.Why a computer model can help in answering the research questions

The use of computer models to investigate processes of linguistic evolution is far from being a novelty. Works of Liljecrants and Lindblom (1972) and de Boer (1999, 2001) on vowel systems have already been quoted here, besides them it is worth to remember the works of Simon Kirby and his team (1999, 2002), of Cangelosi and Parisi (2001) and of Wray (2002). Nonetheless, before tackling the issue of language complexity and population dynamics in a modeling way and introducing the Language Game paradigm, it seems useful to remind the reader the utility of this approach (Ke 2004, de Boer 2006).

In section 3, language has been defined as a complex adaptive system and stress has been put on the distinction between individual and global level of linguistic analysis. Agents in a population create their own individual idiolects by interacting with each other and with the given environment – these idiolects are not necessarily equal because the experiences of the single agents are not necessarily the same. However, language as a global phenomenon emerges from the dynamics of idiolect formation and consists in the communalities between the individual linguistic systems. A proof of this is the success in communication between the agents. However the output and actual dynamics of language formation is not easy to predict because it is not easy to keep track of the high number of local interactions between agents and the mutual influence between individual and global level and between population and environment. A researcher may formulate hypothesis on the behavior of the system, but he needs to test it. In order to confirm or reject a hypothesis made on a complex adaptive system such as language, one has to resort to a tool able to handle this level of complexity, and it is here that a computer model comes handy.

The hypothesis and research questions formulated in section 2 revolve on the role of L1 and L2 acquisition on shaping the complexity of the emergent linguistic system. The questions are: does complexity depends on the fact that natural languages are culturally transmitted systems? And more over, does the complexity of such a system drop when the population gets in contact with another linguistic community, triggering L2 acquisition and creating a language contact scenario? This sort of investigation asks the researcher for at least two conditions: the possibility of playing around with experimental conditions and a large amount of time. If the researcher was to investigate these issues only on the filed would have clashed against such requirements – even though he could have observed clear cases of language emergence, such in cases of creolization (McWhorter 1998, 1999, 2005) or new sign languages (Senghas 1995, 2003, 2005), he couldn't have observed the role of learning, for instance, over more than a couple of generations. Besides the time-scale limit, the possibility of exercise some experimental control seems to be even more important. Given the above mentioned hypothesis, the only way to test the role of learning and contact is to confront groups of populations where those conditions are absent with populations where they are present. Obviously this is quite an aleatory situation: it is not easy to find exactly such conditions in real life, and it is not easy to exert such control on experimental subjects groups of a reasonably large size. That's another reason why a simulation is the most handy and economical solution to the present problems, if not the only one. Time is relatively a non-issue for a computer model, stipulated that the computational machinery is not too complex and heavy for the calculator. Therefore a process of language emergence and evolution that supposedly may take hundreds if not thousands of generations can be easily compressed in a few hours or minutes. At the same time a computer model allows the researcher virtually unlimited control on the conditions. As far as this research is concerned, the present experiments will be run in cases where there is no learning and generational turn over in the population at all, cases where learning is implemented and cases where population is structured in two distinct and interacting groups of learning agents, just like in a simplified language contact scenario. Such liberty, as said above, is basically impossible in traditional field work.

Of course this demands abstractions and simplifications, principally because the computing power of modern calculators is limited. Simplification for instance are demanded in language representation, dynamic of the system, and in the case of agent-based models on the behavior of the agent and on the way it interacts with its surrounding and so on. Simplifications and abstractions are not only limitations to the researcher's liberty to model the target phenomenon. They turn out to be potential allies to the investigation itself – in fact the process of finding, for instance, the right and more efficient representation of linguistic behavior in an agent based model like the present, one is forced to make explicit every mechanisms governing it. In other words, a model calls for precision and formalization, a basic assumption for every effective scientific investigation, leaving aside the vagueness that certain theoretical approaches may have (de Boer 2006). Nonetheless, the simplified nature of many computer models is one source of criticism to this field. The impression sometimes is that the process of modeling abstract away from the real nature of the subject matter. In order to avoid such a situation and criticism towards it, the researcher should be honest on the level of simplification he adopts and state how the abstractions maps onto reality, and the reader on the other hand should. Readers on the other hand shouldn't incur into the risk of rejecting modeling results under the erroneous assumptions that computer models returns only what has been put in by the modeler. This is only partially true, in the sense that a model is naturally determined by the design choices of its creator and that "junks in, junks out", that is if a model is built on faulty assumptions, it will only produce erroneous results, is a principle to bear in mind every time one approaches a computer simulation of any kind. However, on the other hand, the modeling of complex adaptive systems can easily give surprising and unpredicted results in force of the very complexity of the system, as it is said above. The point is not what a model can be made to do, which can correctly raise objections of circularity: a model is made to reproduce what the researcher wants it to reproduce (Nettle 1999, Ke 2004). Rather, the point is what assumptions and abstractions are necessary in order to correctly reproduce a given phenomenon. For instance, in the present case, if one assumes that communication proceeds as described in a Language Game, and that linguistic contact works as in this model, and, at the end of the simulations, the results mimic reality, then one is allowed to conclude that such abstractions, namely language game and that particular population dynamics, are actually the real forces behind the natural language phenomena at study.

After these reflections, the continuation of this section will introduce the language game paradigm and describe: the abstraction adopted as representation of language, the abstract speaker/hearer that is the agent and its behavior.

5.2.A model for language as an emergent phenomenon

The computational paradigm adopted here was developed by Luc Steels and his research group from the second half of the 90's. The reflections on the origin and evolution of language brought to the formulation of an emergentist approach which is on the one hand a reaction to innatists and generativists, and, on the other, closely related if not coincident with theories of self-organization and language as a complex adaptive system as sketched above. The hypothesis is therefore that language is an emergent phenomenon implemented by the interactions of simple agents and shaped by the cognitive capacities, or behaviors of these agents. The concept of agent is therefore central for this framework and, in the course of the series of experiments has been implemented both as computer software and as robotic agents.

The idea is that, given an environment composed by a series of perceptual objects and a population of agents able to perceive those objects and communicate by lexicalizing about them with one another, it is possible to observe the emergence of a coherent linguistic system. This linguistic system in a naming game is nothing more than a series of lexemes, or words consisting in a mapping between a form and a certain meaning. The meaning is a category created by the agent itself by playing a discrimination game on the objects in the environment. Therefore the concepts to be defined are: agent, object and population. These entities are dynamically related by three other abstract entities: language representation, discrimination game and naming game.

5.2.1.Static components: objects, agents' channels, and words.

The environment in which the agents have to interact is called in this model Universe, and it is a collection of perceptual objects, that is entities that are perceivable by the agents as an array of dimensions. So, computationally speaking each object is defined by a finite number of dimensions consisting of a feature-value couple. The feature represents a perceptual characteristic of the object it-self, such as height, width, weight or color – the value is simply a number between 0.0 and 1.

Object 1 [[height:0.4811], [width:0.2935], [colour:0.1584]]

For instance, Object 1 is composed by three dimensions (height, width, color) to which are assigned values ranging between 0.0 and 1. Objects such this are contained in a Universe, consisting of a simple list of objects defined by the same set of dimensions.

The agents itself, as a static object, is the counterpart of the object as the preceptor of the perceptual object. It is the central component of the model and has at least two sets (set of channels, and the lexicon) and a series of procedures that enables it to update these sets on the basis of the interaction with other agents and the environment.

The elements contained in the set of channels are the counterpart of the dimensions defining the objects in the environment and therefore consists of the perception of features such as height, color and so on. Naturally the number of channels and their types map with the number and names of the dimensions of the objects in the universe that the agents have to communicate about. Each channel (ex: channel height below) can be thought as a one-dimensional space with limits 0.0 and 1.0. A channel is divided in categories. As a way of example, channels height, color and width of a hypothetical agent are presented below:

height [(0.0;0.5), (0.5; 0.75), (0.75;1.0)]

color [(0.0;0.1)]

width [(0.0;0.5), (0.5; 1.0)]

Channel width is composed by two Categories, each defined by a minimum value (0.0 for the first, and 0.5 for the second one), a maximum value (0.5 and 1.0). In his discrimination games to study the emergence of shared perceptual distinctions, Steels (1996) adopts a slightly different formalization. He represents categories on the perceptual channel through discrimination trees whose final branches divide the continuous domain of the channels into subdomains. The adoption of such formalism, among other things, is justified by the use of the pruning procedure used both for the lexicon in the naming game, and here, on the trees of the discrimination game. Pruning is not adopted in the present version of the language game, the reason for not doing so and the entailments of this choice will be discussed below.

Perceptual channels and subdivisions in categories enable the agent to perceive its environment that is the objects contained in the universe set. When an agent faces an object, it does not perceive its dimensions as described above, i.e. as a couple of feature and value. The agent perceives the object by categorizing it – this consists in the procedure of assigning a category to each dimension of the object from the perceptual channels of the agent itself. Finding the right category for a given value consists in finding the category that contains the value. For instance, assume that object o's width is 0.33 and that the perceiver's width channel consists of categories [0.0;0.5] and [0.5;1.0], the agent will perceive o's width as the first category, and not through its raw value 0.3. As a way of example consider the following scenario where agent a faces object 1. Both agent and object are respectively defined by three dimensions and three perceptual channels (width, height and color):

Agent a: width [(0.0;0.5), (0.5; 0.75), (0.75;1.0)]

height [(0.0;1.0)]

color [(0.0;0.5), (0.5; 1.0)]

Object 1: [[widht:0.3], [height:0.1], [color:1.0]]

Agent a perceives object 1, by categorizing it in the following way:

Object 1 width(0.0;0.5), height(0.0;1.0), color(0.5; 1.0)

The second set of the agent, the lexicon, contains a series of words. A word is defined as a coupling of form and meaning (<f, m>). In this model a form is simply a string of letters and a meaning is a category in a specific channel. Therefore on the basis of agent a's categories above, a possible configuration of a's lexicon could look as follow:

<"kawabo", C(0.0;0.5)>

<"mutifo", A(0.0;0.5)>

<"sarahe", A(0.5;0.75)>

<"kurati", B(0.0;1.0)>

<"wagebu", A(0.75;1.0)>

<"remuwo", C(0.5;1.0)>

In this Lexicon, the first Word maps the random string "kawabo" to Category (0.0;0.5) of Channel C.

5.2.2. Agent behaviors: discrimination and naming.

However, an agent is not only a collection of perceptual channels divided in categories and a lexicon. It is also an active player in the language game: it can update its channels creating new categories to adapt more efficiently to the environment and can create and learn new words as the basic requirements for interacting with the other agents in its population. These two processes are the core components of a language game, at least as the one presented and implemented by Steels in his Talking Heads experiments. The following section will describe the dynamics of the Discrimination and Naming Game.

Discrimination game consists in updating one agent's perceptual channels by subdividing them in new finer categories. The theoretical background from which Steels and his team move for defining this procedure is purely empiricist. Empiricist cognitive sciences claim that concepts and categories employed by humans are not innate, but that are learned by experience and that are the results of general invariance detection processes. An individual facing a series of objects or situations is able to compare them and extract a series of similarities and differences that will converge to the creation of categories or perceptual distinction.

Cognition, at this level of abstraction, is therefore a cultural construct because depends on the experience of the subjects and whose uniformity has to be looked for in processes of learning and cultural transmission. And it is here that language kicks in, since the shared conceptual machinery in a community is spread and maintained through communication. This vision of discrimination game is substantially in line with the assumptions ascribed to a complex adaptive system: the system is not innate but emergentist from experience, it adapts to the environmental conditions and it is the product of local interaction between agent and perceptual objects and between agent and agent.

The basic procedure therefore consists of an observing agent and a section of the reality composed by a series of objects functioning as context for a given object topic. The observer must be able to distinguish the topic from the rest of the objects (context) after having conceptualized the scene (assigning the objects dimensions to categories). If the categories possessed by the observer at a particular time are not sufficient for distinguishing the topic from the context, that is, if the topic is not characterized by not even one category not shared by the contextual objects, the observer has the chance of learning from experience. Learning in this case consist is refining one of its perceptual channels by creating smaller subdivisions of it, i.e. by creating finer categories. These new categories will increase the chances of distinguishing the topic from the context. In this way the repertoire of concepts owned by an agent has the possibility of expanding more and more as new scenarios of topics and contexts are presented to the observer by experience. This possibility entails at least two phenomena. On the one hand, the number of categories expands boundlessly. On the other hand, the more categories an agent has, the higher the probability of being able of distinguish an object from another. In this way the system of perceptual distinctions of an agent adapts itself through experience to the environment it happens to emerge in. However, Steels introduces in his model a selectionist-like process that permits to preserve only certain categories and to eliminate other. This consists in a pruning dynamics in which categories (or tree-nodes in Steel's implementations) that have proven to be successful in previous situations are preserved, whereas useless distinctions are forgotten. It creates an ecology of distinctions which is adapting itself to the environment the agent encounters. A similar procedure is applied also for the agent's lexicon. In a way, pruning is a way to keep an adaptive system (distinctions, lexica or grammars) under control, in the sense that it limits its size and complexity by filtering out elements that for one reason or another are no longer functional to the task assigned. Linking this process with the research questions of this work, it is legitimate to suspect that by implementing a Steels-styled pruning dynamics, one runs into the risk of obtaining a stable reduction of complexity by default, through an agent internal procedure and without leaving this task to cultural transmission and population dynamics. Of course, this is what Steels wanted from pruning, however, for these reasons the present model will leave pruning out, relaying only on the effects of L1 and L2 acquisition as hypothesized in sections 2 and 3 and as it will be investigated in the next section.

Discrimination game consists of learning how to distinguish one object from the other using one's categories. Therefore, given a subset of objects from the universe and choosing an object as a topic out of it, the agent tries to find which categories fit with the topic and not with the context. If the Agent is able to find at least an aspect of the topic that distinguishes it from the context, the discrimination game is a success. However, especially when the agent is still 'young', that is after having played only a few times the game, it is quite likely that it will not be able to find any distinctive category for the topic, in this case the game fails. The failure of the discrimination game pushes the agent to improve itself by modifying its own perceptual channels and creating novel perceptual categories. In order to do so, the agent takes a random perceptual channel (guaranteed that the topic has a dimension corresponding to it), and from that, it takes the category that was assigned to the topic by categorizing it and splits it. This creates a narrower category for the topic categorization. However, since the first step of this updating process is essentially random, it does not always prevent a failure for a future discrimination game with the same topic and context objects. As a way of example, consider the observer agent a facing two objects, 1 and 2, out of which 1 is selected as topic, and 2 functions as context. The aim of the game is to find a category in the observer's perceptual layer that is assigned only to the topic, and therefore distinguishes it from the context.

Agent a:	width [(0.0;0.5), (0.5; 0.75), (0.75;1.0)]
	height [(0.0;0.1)]
Topic	1 - [[width:0.2], [height:0.1]]
Context	2 - [[width:0.5], [height:0.3]]

The observer categorizes 1 and 2 in the following way:

Topic	1 - width(0.0;0.5), height(0.0;0.1)
Context	2 - width(0.0;0.5), height(0.0;0.1)

At this point the observer looks for those categories that belong only to the topic and not to the context. Unfortunately, it turns out that there is not such a distinctive feature, and the game fails. The reason of failure is ascribable to the set of categories of the agent itself that is still too small to handle even a simple task as distinguishing object 1 and object 2. At this point the agent has open only one solution: learning from experience and trying to improve it discriminating capabilities by expanding its set of categories. Therefore, the observer selects a random channel: width [(0.0;0.5), (0.5; 0.75), (0.75;1.0)] and splits category (0.0;0.5). This returns an updated set of perceptual categories:

width [(0.0;0.25), (0.25;0.5), (0.5; 0.75), (0.75;1.0)]

height [(0.0;0.1)]

It is quite straightforward to realize that this procedure, when played by more than one agent in a population does not guarantee that the agents share all the same set of perceptual categories, even when facing the same universe of objects. Similar experiences guarantee that their perceptual capacities are not extremely different, but the randomness of the updating procedure open the way to variability and, in second analysis, to complexification.

Naming game is the procedure employed by a group of agents to build a shared language, or as in this case a set of words defined as form-meaning pairs. The agent, having built up a series of categories, communicates them with the other members of the population engaging in a naming game – it takes category, presented by a discrimination game, and couples it with a string of symbols. This instantiate a word that will populate the agent's lexicon or language, which in turn is nothing more than a simple associative memory. Each agent creates its own meanings and communicates them to the others, this brings to the emergence of a shared linguistic repertoire that is the product of an a priori design not governed on a global level but the result of local distributed activities of the individual agents. Besides the ability of verbalize, that is to create new words for new meanings, the agents must be able to parse this words when presented with them, i.e. to interpret them and extract their meaning. Since agents in a speaker-hearer relation are unaware of the meaning intended by their partner and communicate and perceive only the form or utterance of the word, the parsing procedure is not trivial and takes the shape of a guessing game.

Speaker and hearer perceive the same group of objects, the speaker choses a topic and eventually find a distinctive feature for it, the speaker find a word for this feature and sends to the hearer the form of the word – at this point the hearer first find the meaning of the utterance and on the basis of that tries to individuate the object that was previously intended as topic by the speaker. The game is a success if the hearer finds the right object. However each of these passages is an occasion for failure, making this game more difficult than it may appear at first glance. The agent may fail in finding a distinctive feature for the topic, and here it faces the limitation of a possibly immature conceptual space that still need to be expanded through discrimination games. Moreover the agent may also lack a word for the distinctive meaning, and therefore it has the chance to verbalize it, i.e. create a new word. But the hearer's task is even more difficult. It may not recognize the utterance sent by the speaker. In this case it is invited to learn it by coupling it with an appropriate meaning. At this point the speaker reveals to the hearer what was the intended topic of the conversation by pointing at it. This does not solve directly the learning problem for the hearer, because it may or may not distinguish a feature for the revealed topic – if it happens to find a distinctive feature, it couple it with the utterance and creates a new word for its lexicon, if, as it may happen especially for young agents, it has no category, the game reaches a dead end and fails. But suppose that the hearer recognize the utterance sent by the speaker. Even in this case the pitfalls are present. In fact, the hearer selects a meaning from the recognized word and look in its surroundings for an object that fit the description. Three things may happen at this point. There may be more than one object that corresponds to that specific category chosen by the hearer, in this case the game fails because the topic has not been individuated. It may also happen that the hearer cannot find any object, and this again is possibly due to a lack of perceptual categories. The most obvious results however are that either the object individuated coincides with the topic or not. In the first case the game is a success.

To sum up this section, two are the linguistic abilities of the agent: constructing and acquiring words. In the first case the agent creates a new word by mapping a meaning that it never experienced before with a string of signs. In the second case, an agent hears a word he has never heard before and he may decide to add it to his linguistic knowledge. Both these procedures however are determined by probabilities that in Steel's work are called word creation rate and word absorption rate. In the original model these parameters are set by the experimenter and are equal and constant over time for all the agents (usually these corresponded to probability 1). As it will be explained in the next section , in this experiment both word creation and word absorption rates are connected to the age of the single agent, implementing in this way a simple aging structure.

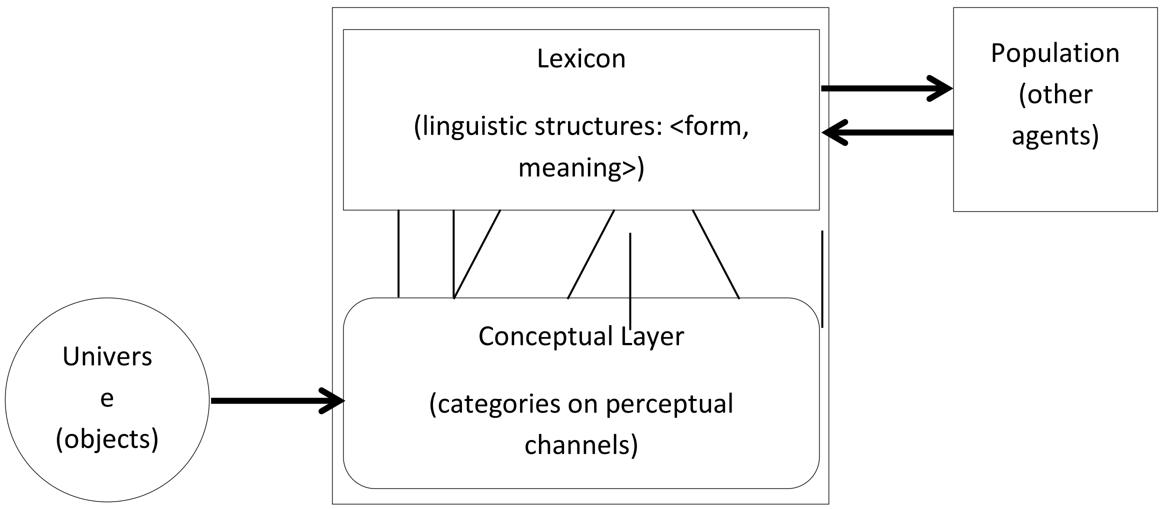


Figure 1 – Schema of the principal components of an agent and its interaction with the universe of objects and other agents.

It is important to stress out the intricacy and fragility of the communication process, because it is the main reason why a communication system is positively redundant and prone to complexity. To be more specific complexity arises because the system allows the emergence of ambiguity (words having different meanings but sharing the same form) and synonymy (words having different forms and sharing the same meaning). These phenomena are the consequences of the way the agents acquire new forms which in turn is grounded on the assumptions that language is shaped by use and that agents cannot access the semantic representation of their counterparts when communicating. Especially when the game is played in large population and in complex environment, it is likely that an agent creates new words for certain meanings ignoring that expressions for those meanings already exists in the communal language. At the same time the process of acquisition, as seen above, is not immune to ambiguity creation because the learner, when memorizing a form by coupling it with a meaning does not know what meaning was conveyed by the teacher. Although the reality that teacher and learner face is identical, the way they look at it may differ because their semantic layers, or perceptual categories sets are shaped independently one from the other.

Ambiguity and synonymy make the language more complex and large and make the life of speakers more difficult because open the way to misunderstandings and inefficiencies. Steels proposed a system of score updating and pruning also for the lexicon of his agents: when a word is used and the result of the game is positive, a score associated to that word is increased and, at the same time, the score of synonyms is decreased of the same rate. Periodically, words that have scores under a certain arbitrary threshold are eliminated from the lexica, keeping in this way a low level of redundancy in the system. However this procedure appears quite artificial and mechanical. It is somehow independent from another important aspect of language: acquisition. In the present model pruning is not implemented. Agents simply update the score of their words. Then, when a word is needed for a given meaning, and the agent has more than one word, it will tend to use the one with the highest score. In case the hearer has to learn the word used by the speaker (because he needs such a word), it will acquire always the most successful and used one from the repertoire of its speaker/teacher companion. In this way, coupling learning with aging and population renovation (the removal of older agents from the population) the effects of keeping a reasonably low level of redundancy is obtained without implementing pruning. This will be the subject of the first set of experiments of section 6.

6.The experiments

This section presents the results of two distinct set of experiments. The first one (6.2) focuses on the role of aging and first language acquisition in shaping the complexity of the emerging system. A population of n agents, immerged in a set of m objects will be let free to develop its own language under the constraints of the Language Game as described in the previous sections. Keeping invariant the type of environment (type and complexity of the objects, number of objects in the context presented to the agents etc.) a second simulation will be run, this time allowing the agents to age and to be periodically replaced in the population in order to implement inter-generational L1 acquisition. The data from the first and second run of the simulation will be compared keeping in mind the following measure: communicative success, lexicon size (Economy) and level of mismatch (Transparency). If in the second run, while keeping a reasonably high level of communicative success, the size of the lexicon and the level of mismatch decrease, it is possible to conclude that learning has as the effect of limiting language complexity. The second set of experiments (6.3) goes further by investigating the effect of population contact and L2 acquisition again on complexity.

Besides the control run, the experimental run consists of a population where, besides aging and L1 conditions, also a contact scenario is implemented. After x repetitions a set of new agents, already equipped with their own language will be merged with the initial population. These newcomers will try to pick up the language of the hosting population and communicate with the native agents, reproducing a case of L2 acquisition. The hypothesis is that the reduction in complexity (Economy, Transparency) will be even more evident here than in the first set of experiments.

6.1.Definitions

6.1.1.Measures of Complexity and Communication Success.

Section 4 has introduced a definition of complexity as a measurement on the average lexicon size of the emerging language and the level of transparency. For these reasons, this measure of complexity has been defined absolute. Section 5 has described the Language Game paradigm on which the present experiments are based. Among other things, that section introduced the formalization of language implemented in the model and the concept of communicative success and the way it is updated. The present section gives a precise formulation of language, of measures of complexity and of communication success.

Language, as defined in section 5, is a function that maps forms to meanings. Forms are simply randomly generated strings of syllable (ex: kawobe). Meanings are categories created on perceptual channels created by the agents by playing Discrimination games on sets of objects selected from the environment. A form-meaning mapping, or word, takes therefore the shape of a couple <kawobe; $_{height}$[0.0-0.25]> where the first term represent the form and the second the meaning, in this case a category between 0.0 and 0.25 on the HEIGHT channel. On a more concrete level, language is a set of these couplings and thus it can be considered quite simply a lexicon of Sussurean-like signs <f, m>.

In order to fit the concepts of complexity to the Language Game paradigm and the formalization of language adopted here, Economy is measured as the average size of the lexica of the agents presents in the population at a particular time. As transparency is concerned, the measure of mismatch can take values between 0 and 1, where 1 indicates a state of perfect 1-to-1 relation between forms and meaning (total transparency) and 0 a state of total opacity. Here again, the measure focus on the communal language and not on the single idiolects. Therefore the measure reported below is computed on the basis of the level of mismatch of each individual idiolect. Each idiolect is assigned a binary value – 1 if it approaches a state of perfect Transparency (one meaning – one form and vice versa), 0 if it parts from that state. Then the global level of Transparency is computed as the average level of mismatch among the population idiolects.

As a way of example, take two languages developed in similar conditions: two populations of the same size grown in similar environment. If the first language is composed by an average of 30 lexemes and its average mismatch is 0.1 and, on the other hand, the second language has a lexicon of 20 lexemes and an average mismatch of 0.6, the latter is less complex than the former.

Given this definitions, the more economical and transparent a language is, the less complex it will be considered. In other words, a language with a small lexicon and 1-to-1 relation between forms and meanings will be considered less complex than a language with a larger lexicon a more opaque relation between lexicon and semantics.

6.1.2.Aging Structure (AS) and Population Renovation (PR) Measures of Complexity and Communication Success.

Aging structure (AS) and Population Renovation (PR) are the main mechanisms applied to the population dynamic that implement inter-generational variation in behavior and, in the last analysis, language acquisition. AS and PR, together with population merging (described in 6.3) are also the methods that implement a language contact scenario and L2 acquisition.

Aging Structure implements a variation in probability of learning and creativity as function of the age of the single agents. In the previous section, learning has been defined as the ability of an agent to update its own language by memorizing a newly encountered linguistic expression. Creativity, on the other hand, is the ability of creating a new expression in order to express a newly encountered meaning or situation. Aging is defined by a numeric value named age assigned to every agent, this value is increased by a constant factor as the simulation proceeds. In other words, the agents grow old during the simulation. The age of the single agents determines the probability with which the agent itself acquires and creates new words during the interactions that constitute the Language Game. As acquisition is concerned, the older the agent, the lower is the probability that it will updates its lexicon by acquiring new words. This is in line with Trudgill (2001) that suggest a relation between learner's age and his ability to acquire a language: the older he is, the less precise or complete his acquisition of the language will be.

Age	*Probability*
0-10	100%
10-30	75%
30-50	50%
50-70	25%
70-∞	1%

Creativity of the agent as well is determined by the age:

Age	*Probability*
0-10	0%
10-50	100%
50-70	75%
70-∞	25%

Population Renovation (PR) concerns the composition of the population itself. An important component of the model is the population, i.e. a set of interacting agents. All experiments are initiated with a population of n agents (25, 10, 4 etc.). Under normal condition the population size and composition remains invariant till the end of the simulation. In this case the agents at run 0 are the same at the final run of the experiment, with the only difference that they will have a much richer set of categories (meaning) and words to express them. If PR is implemented, a fraction of the agents is periodically removed and replaced with new agents, whose age is set to 0 and whose lexica are empty. The primary effects of PR are: keep the average age of the population constant and simulate an intergenerational transmission of the language through learning. This last fact is triggered by the fact that the new agents introduced in the population have a high probability of acquiring the language of the older agents without introducing new, potentially redundant forms.

6.2.Effects of Aging and L1 Acquisition of the Complexity of the System

The experiment has been conducted under two main conditions: absence of aging structure and population renovation, and presence of aging structure and population renovation. The first condition simulates the hypothetical situation of a linguistic system emerging in a closed and static population. The second condition refers to a linguistic system that is transmitted from older to younger agents: that is a system transmitted culturally from generation to generation. The second case therefore implements a closed population where first language acquisition is dominant. In fact there are no other groups of agents equipped with different lexica or conceptual layer involved in learning. The parameters of population size, universe size and number of channels and dimensions are reproduced below.

Population size	20
Universe size	15
Number of Channels	3
Number of Dimensions	3

These parameters are the same for both the first and second simulations. The only variable is given by the implementation of aging and renovation. In the second simulation after a period of 100 repetitions called maturation, the population is partially renovated every 80 repetitions by replacing 2 old agents with 2 new ones. The maturation period is necessary to allow the emerging language to reach a reasonable level of complexity that enables the agents to communicate efficiently, it is however short enough to avoid phenomena of overcomplexification. The hypothesis is that the second set of simulations will display an average lower level of complexity represented by the size of the lexicon and the level of mismatch. The graph below gives an idea of the trend of communication success, lexicon size and mismatch parameters in simulation over 2000 runs of the two experimental conditions (bold line represents the case where first language acquisition is absent, the dotted line represents the case where it is implemented in the population dynamics).

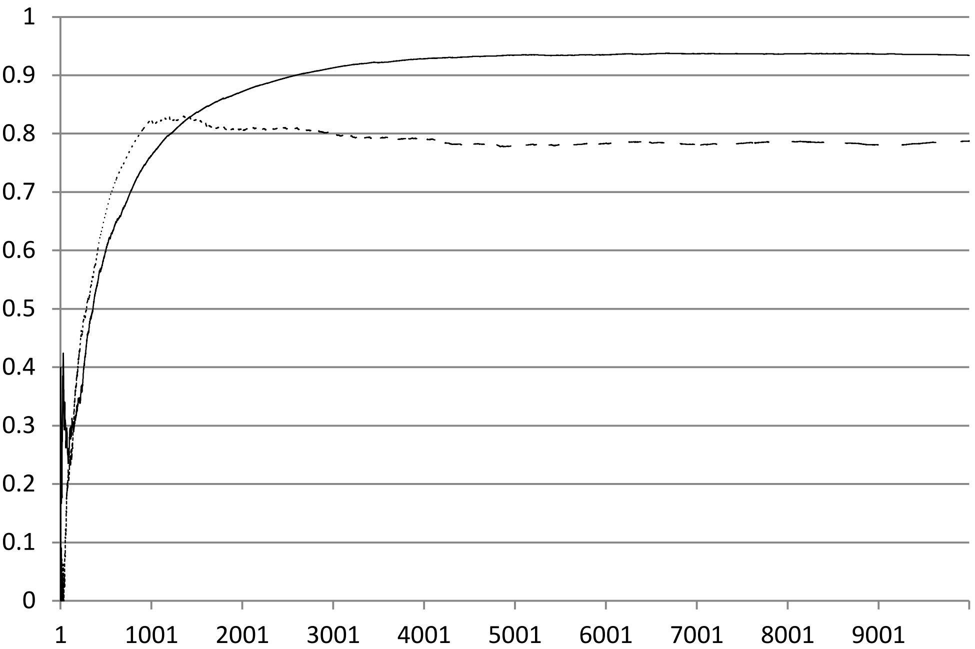


Figure 3 – Success rate (continuous line = without L1, dotted line = with L1)

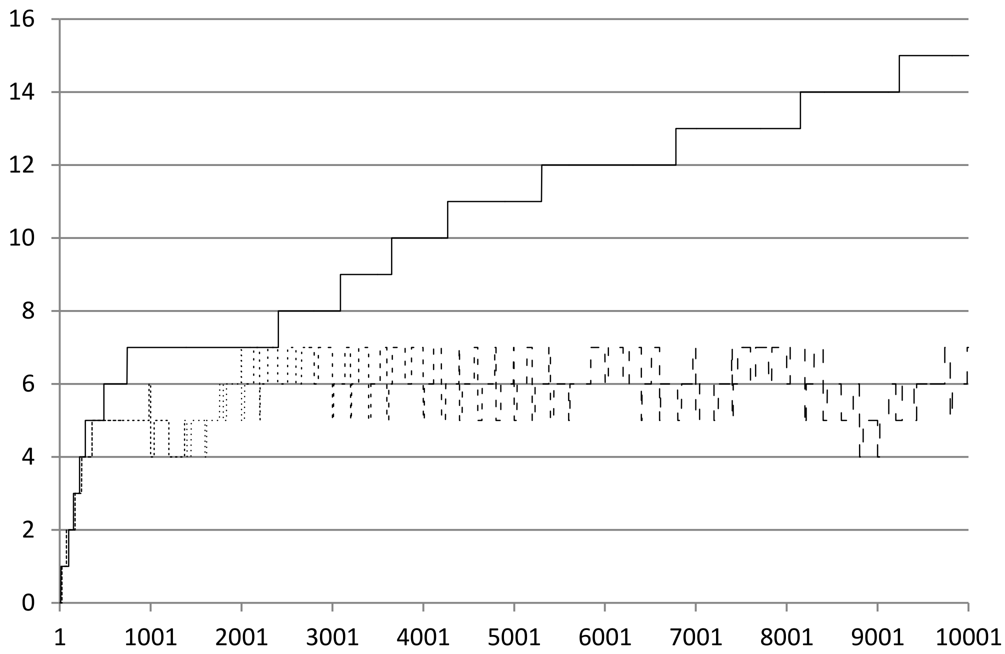


Figure 4 – Lexicon size (continuous line = without L1, dotted line = with L1)

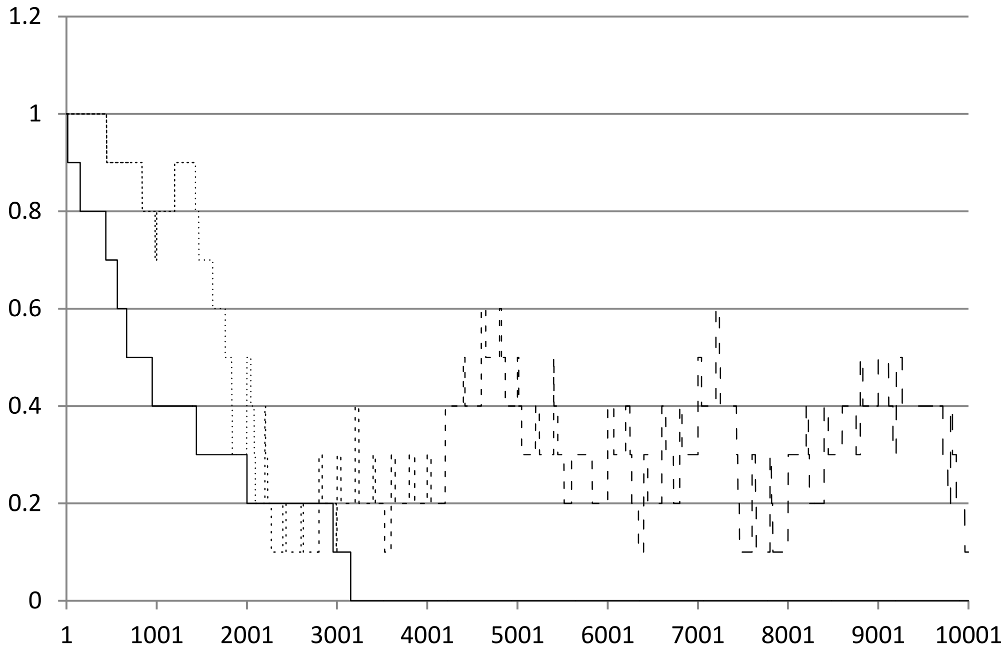
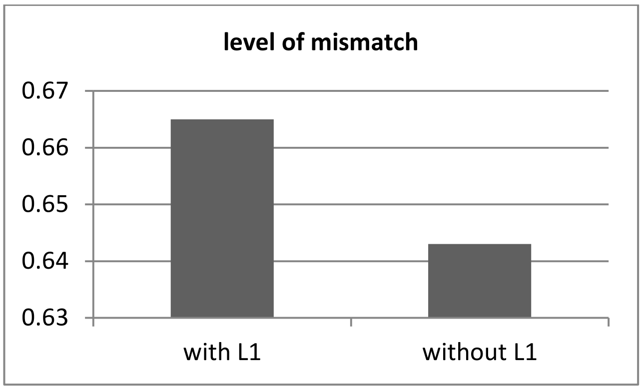


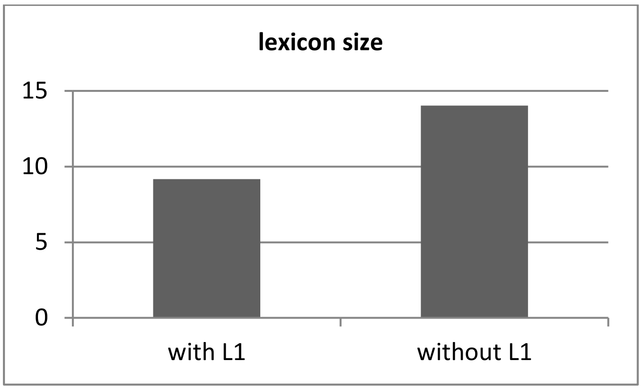
Figure 5 – Mismatch level (continuous line = without L1, dotted line = with L1)

At a first analysis, it is possible to notice that the level of communication success is relativelly high also for the second experimental condition, whereas the lexicon size is evidently smaler and the level of mismatch closer to value 1 (i.e. closer to a state of 1-to-1 relation between form and meaning). The fact that the communication success is still quite high (around 70-80%) indicates that the emerging language, even though simpler, provides an efficient sistem of communication. The simulations under the two conditions have been repeated 100 times, the table below reports the raw average values for lexicon size, success and mismatch.

	With L1	*Without L1*
Lexicon size	9.18	14.03
Mismatch	0.665	0.643
Success	0.657735	0.808425

The histogram below display the frequencies of values of lexicon size and mismatch for the two experimental conditions, here again it is possible to see how these values differ between the two set of simulations.





The t-test confirms the difference in size of the lexicon and in the level of mismatch between the two conditions: size of the lexicon $p<0.001$, $t=-36.0857$; level of mismatch $p<0.05$, $t=2.9598$.

6.3. Effect of population contact and adult L2 acquisition on complexity

The first experiment considers the difference between the effect of having or not having a L1 acquisition-like transmission of the linguistic system. However the type of population is still closed, there is no contact with external population and their systems of communication. This tranche of experiments investigates effect of language contact and L2 acquisition. The procedure will be similar to the one described above and it is based on the comparison of the average success rate, lexicon size and mismatch between population where L1 and/or L2 is implemented and populations where these procedures are not implemented.

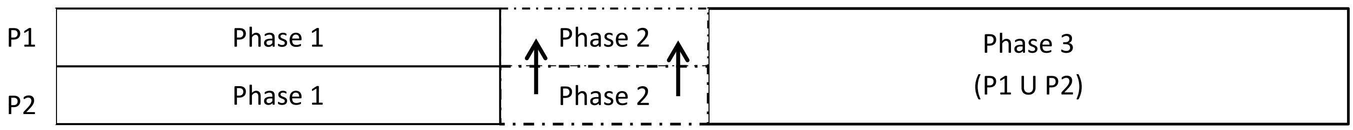


Figure 6 – the 3 principal phases in the second experiment. Phase 1 consists in the independent development of the two populations (P1 and P2), Phase 2 indicates the first period of contact when P2 agents have a high probability of being hearer and therefore acquire P1's language, finally Phase 3 consists in the union of the two populations.

6.3.1.Difference between first and second sets of experiments

In the first experiment, the language community is periodically renovated by the replacement of a portion of its inhabitants with an equal quantity of new agents. These new agents have both an immature semantic layer and an empty lexicon. An immature semantic layer consists of a series of perceptual channels containing only one general category whose values span from 0 to 1 – in other words the channel has not been divided in perceptual categories yet. This new agents will tend to acquire the language of the remaining community without altering it too much. This is possible because the probability of acquiring a new word is very high at the beginning whereas the probability of introducing new forms is nearly null. The second experiment implements a different social structure and consequently a different type of learning. In this case the initial population of agents, the one instantiated at the beginning of the simulation, is open to contact with a group of external agents. In order to simulate a language contact scenario, two distinct populations are created. In a first period this two populations are let to develop their own systems of categories and their languages following the procedures of language game described in the previous section. In this phase the populations are closed and isolated one from the other: that is, there is a 0 probability that an agent of the first one interacts with an agent of the second one and vice-versa. After this period, the populations are merged, allowing the agents to interact freely. This implements an exoteric or type 1 society (Kusters 2003, Lupyan and Dale 2010). Considering that the focus of this thesis is on the effect on the complexity of an artificial linguistic system, only the language of one of the two populations will be considered and monitored with the measures of size and mismatch. Therefore the adult L2 acquisition procedure is implemented by considering an intermediate phase before the actual merging of the two populations (see figure 2). In this phase agents of the second (and therefore incoming) population have a higher probability of learning the language of the host population (the first population) rather than creating new forms while interacting with foreigners. This is quite similar to the situation of new agents in the previous experiments that have a higher chance to learn a language in the first period that are inserted in the population. However the similarities between the two experiments end here. The differences are present both on the individual and the population levels. First of all the new agents of the first experiment have an immature conceptual layer, determined by the fact that they have no precedent experience with the universe of perceptual objects – they start out blank. The incomings agents of the contact population have already their own grown up conceptual layer and their own language developed by their previous experiences antecedent to the contact with the target population. Moreover the differences are reflected also on the age level: every new agent is introduced in the population with an initial age set at 0, whereas the incoming agents have variable ages always bigger than 0 (average incoming agents age 20 years). In the first experiment there is no transitional phase: every agent, independently from being newcomers or not, have the same chance of playing the role of speaker (teacher) or hearer (learner). Then the probability that a hearer (Lerner) actually acquires a newly heard form only depends on his age. In the second experiments however the transitional phase imposes a higher probability on the role played by the hosts and newcomers: the firsts are more likely to act as speakers (teachers) and the seconds to b hearer (learners). Here again, the probability of success is independent from the origin of the agent, and determined by its own individual age. These differences are in line with the generic differences between child and adult learning contexts and are synthesizable in two points: differences in experience (blank vs. mature conceptual layer) and age (greater or smaller ability to acquire a new code). Said so, one is entitled to expect a different evolutionary trend in the emerging.

6.3.2.Simulating contact and adult L2 acquisition: experiment and results

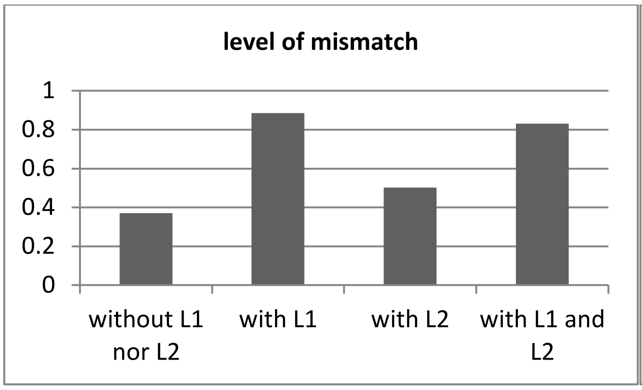
This experiment is divided in four distinct sets of simulations each of which implements one of the two learning procedures (L1, L2), both or none of them. The main focus is to observe what is the effect of having an open population rather than a closed one. The parameters that are used all the four sets of simulations are:

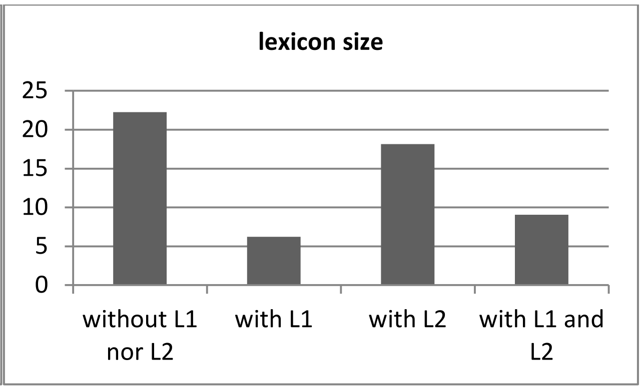
Population size	20
Universe size	15
Number of Channels	3
Number of Dimensions	3

The maturation time is set at 500 runs, and the population renovation consists in replacing 1 agent each 100 runs (this plays a role only when L1 is implemented). The parameters concerning language contact and L2 acquisition are the size of the incoming population (5) and the stage where this contact event takes place (after 1000 runs). The experiments are repeated 100 times each and the averaged results are reported below:

	withoutL1/L2	*withL1*	*withL2*	*withL1/L2*
Lexicon size	22.25815	6.22503	18.12576	9.077988
Mismatch	0.371223	0.885931	0.502482	0.830829
Success	0.742756	0.285898	0.722397	0.418367

The graphs below synthetize the differences in lexicon size, mismatch and success rate between the four experimental conditions.





A two-way analysis of variance on both the lexicon size and level of mismatch confirms what is already evident from the graphical representation above: both L1 and L2 and the interaction between the two factors determines a decrease in size of the lexicon and values closer to 1 of the mismatch measure (higher this measure, the less synonymy and ambiguity a language is supposed to have).

Source	*SS*	*df*	*SM*	*F*	*p*	*F crit*
L2	31.3917025	1	31.3917025	72.18447629	4.01874E-16	3.865048
L1	15926.13961	1	15926.13961	36621.78078	0	3.865048
L1xL2	1153.680995	1	1153.680995	2652.86212	1.3099E-177	3.865048
	172.2131243	396	0.434881627			
Total	17283.42543	399				

The table above summarizes the results of a two-way analysis of variance conducted on the values of lexicon size of 100 repetitions under different population structures and types of acquisition conditions. The analysis shows a significant main effect for the L2 factor, F(1, 396) = 72.18, p < .001; significant main effect for L1 factor (1, 396) = 36621.78, p < .001; and significant effect of the interaction of these two types of acquisition, F(1, 396) = 2652.86, p < .001.

Similar results are met for the level of mismatch, the table below reports the results of another two-way analysis of variance on the experimental results:

Source	*SS*	*df*	*SM*	*F*	*p*	*F crit*
L2	6623.799	1	6623.799	291.6725	2.14E-49	3.865048
L1	6018.821	1	6018.821	265.0329	5.48E-46	3.865048
L1xL2	6537.414	1	6537.414	287.8686	6.43E-49	3.865048
	8993.047	396	22.70972			
Total	28173.08	399				

The table above summarizes the results of a two-way analysis of variance conducted on the values of lexicon size of 100 repetitions under different population structures and types of acquisition conditions. The analysis shows a significant main effect for the L2 factor, F(1, 396) = 291.67, p < .001; significant main effect for L1 factor (1, 396) = 265.03, p < .001; and significant effect of the interaction of these two types of acquisition, F(1, 396) = 287.86, p < .001.

6.4.Success score: is it always high enough?

One of the assumptions that guided these experiments was the idea that even though complexity drops under L1 and L2 acquisition, the success rate remains relatively high. It is undoubtly that the success rate drops when young or foreign agents enters the population an try to learn by using the emerging linguistic system, this is something that Steels already pointed out in his experiments (Steels 1997). However, the point is to prove that the success rate represents a probability of being understood that is still higher than randomness. In other words, a given level of success is not obtained by chance, but it is the results of the structure of a shared and possibly coherent system of signs. In order to demonstrate this, it is necessary to determine which is the chance of performing a successful communicative act, given the language game procedure described in section 5. What is the probability by chance that a speaker/hearer couple agree on which object is the topic of their conversation? If the experimental results show a level of success systematically higher than the one that one would expect by chance, then the communication is guaranteed by a shared language.

The chance of success corresponds to the chance that the hearer recognizes the correct object among the ones presented in the experiment. There are many ways in which this goal may fail. First of all there might be a deficiency on the conceptual layer, that is, the speaker or the hearer or both does not have a category that distinguish the topic from the context of objects in the universe. Second, there might be a deficiency on the lexical level, in other terms, speaker might not have a form to express a given meaning or the hearer might not be able to match an utterance to a specific meaning. In the last analysis, there might be also problems in coupling meaning with an object, this is the case of when the hearer cannot identify one object given the incoming meaning. At this point, it is possible to say that the game has six possible outputs and that the probability of having a positive output is 1/16, i.e. 6%. This is because the event that the hearer recognize the correct object as the topic is conditioned by four binary events: the fact that the speaker has a distinctive feature for the topic or not (1/2); the fact that the hearer recognizes the utterance (1/2); the fact that the hearer individuates 1 and only 1 object from the context (1/2); and the fact that the object individuated by the hearer is the topic meant by the speaker (1/2). Given this .06 threshold for the success rate of the game, if the experimental linguistic systems after maturation are well above this value, it is possible to claim that the system is not random but that its efficiency is due to its emerging structure. All experiments return values clearly above the threshold, even in cases where the success rate is relatively low:

Experiment1		Experiment2			
Without L1	With L1	Without L1 nor L2	With L1	With L2	With L1 and L2
0.808425276	0.657735	0.742756	0.285898	0.722397	0.418367

6.4.Comments

To sum up, the experimental data and their relative analysis show that language acquisition and population structure have an effect in reducing the size of the lexicon and the distance between the emerging system a state of transparency. In other words, the more dynamic and open to contact and change a population is, the less complex its language appears to be. This follows from the definition of complexity adopted in section 2., which is basically quantitative and absolute. That is, it considers complexity on the global level, disregarding the level of the single individual idiolects. However, the same results, especially with regard to mismatch, are corroborated by opening up the system and looking at the states of randomly chosen agents at two different stages of the simulation. The table below reports portions of lexicon of two randomly chosen agents from repetition 100 and repetition 2000 of the same experiment.

Random agent1 at rep 100		*Random agent2 at rep 200*	
[…]		[…]	
karemi	width(0.0-0.50)	turure	width(0.0-0.50)
turure	width(0.0-0.50)	materu	color(0.75-1.0)
jemoso	color(0.25-0.50)	fohuza	height(0.125-0.25)
[…]		[…]	

Random agent1 and agent2 are not necessarily the same agents at two different stages of the population history, this is in fact quite unlikely. The simulation in fact implemented an L1 acquisition process that periodically renovates the population. Observing the two samples, it is possible to note a reduction of synonymy: agent1 had two forms (karemi and turure) for the same meaning width(0.0-0.5), whereas agent2 preserves only one of the two form-meaning pairs. This shows how the process of language acquisition and population renovation filters out possible redundant forms from the system without recurring to ad hoc processes such as pruning.

Keeping in mind the relative abstractions necessary for a computational model such the present one, this is in line with the observations of Kusters (2003) and Lupyan and Dale (2010) and confirms the hypothesis that linguistic complexity is partially determined by population structure.

7.Conclusions

The hypothesis formulated in section 2 was that the complexity of human languages is partially determined by the social environment in which they happens to emerge. Such claim is supported and suggested by a series of studies among which the ones of Kusters and Lupyan and Dale have been considered with particular interest. This thesis wanted to approach this “hot topic” from a opposite prospective: the one computer modeling and simulations. The idea was to start from some theoretical assumpions on: what is language and how it evolves, how linguistic complexity can be characterized and measured and how language as a system can be implemented in a computational simulation. The first step consisted to find a suitable definition of language that encompasses its evolutionary processes and the interaction with its environment. That was found in the theory of language as a complex adaptive system (Ellis and Larsen-Freeman 2009) and in a series of studies on the self-organizing properties of linguistic systems (de Boer 2001). One of the reasons for which CAS theories have been important for the present thesis is the recognition of the dependency of the emerging language to the social and material environment in which it develops. The second step consisted of looking for a suitable definition of complexity that would be easy, formal and measurable, and, in last stance, implementable in a computer model. The definition adopted here considers complexity from a global level and measures it on the communal language rather than the single idiolects (the language of the individual speakers). Then, complexity is split in the concepts of Economy and Transparency, measured as lexicon size and level of mismatch. The less forms a language has, the more economical and less complex is. The closer its mismatch index is to level 1, the more transparent and less complex is. Given these measures, the next step was finding an easy model to simulate the relation between population and complexity. Steels language game paradigm was adopted and modified: a population of agents are asked to build up a conceptual layer in order to distinguish the objects in the environment and to create a language to communicate these distinctions with each other. Such a model presents a series of advantages: it implements a simple but realistic concept of language (as a set of form-meaning pairs), it reproduces the complexity of the communication process (finding the right semantics for a given situation, communicate and interpret it), and it incorporates processes of language creation and acquisition directly in the language use processes (i.e. there is no distinct learning algorithm from the simple communication game paradigm). One main difference between Steels’s original model and the one adopted here is the absence of pruning as a method for keeping the complexity of the language at bay. Here, such a process is not implemented and its effect is demanded to intergenerational and intercommunity language acquisition. Having done the theoretical part, the experimental design and results are reported in section 6. Summarizing this part, it is possible to divide the experiments in three main groups according to the population dynamics and types of acquisition allowed by the populations themselves: static, dynamic but closed, and open. A static population is a population where nor L1 nor L2 acquisition is allowed and where the agents at the beginning of the simulation are the same at the end of it. A dynamic but closed population is a population that allows renovation (periodical change of parts of its agents), aging and therefore intergenerational L1 acquisition. It is closed, however, because the emergent linguistic system and its speakers are not confronted with flux of external agents coming into the population and trying to acquire the language itself (L2 acquisition). Finally, an open population is a population that reproduces the situation of language contact and adult L2 acquisition. The first two typologies loosely corresponds to Kusters’s type 1 and Lupyan and Dale esoteric populations, whereas the last one corresposnds to type2 and exoteric populations.

Experiments	*Kusters (2003)*	*Lupyan and Dale (2010)*
Static	Type1	Esoteric
Dynamic and closed		
Dynamic and open	Type2	Exoteric

The results of the experiments confirms the hypothesis that open populations (Type2, Exoteric) tends to have a relatively simpler language than Static ones, and that Dynamic system in general have simpler languages than static one. This is in line with the observations contained in the literature. The fact that one of the main differences between static and dynamic populations is the fact that the first does not have processes of language acquisition whereas the seconds do (L1 or L2 or both) suggests that language learning is a factor of linguistic simplification and evolution. This final observation is in line with a large portion of recent cognitive and computational studies on language change and evolution such the ones of Christiansen and Chater (2008). Moreover, from a methodological point of view, these observations on the role of language acquisition may open a fruitful discussion on the formal relationship between cultural horizontal transmission developed in Steels language and imitation games and vertical transmission as perpetuated by Simon Kirby's Iterated Learning models (1999).

References:

- Aitchinson, J. 1991 Language change: Progress or decay? Cambridge: Cambridge University Press.
- Andersen, R. W. 1984 The one-to-one principle of interlanguage construction. Language Learning, Vol. 34, issue 4, pp. 77-95.
- Bybee, J. 2006 Frequency of Use and the Organization of Language. Oxford: Oxford University Press.
- Cangelosi, A. and Parisi, D. 2001 How nouns and verbs differentially affect the behavior of artificial organisms. In Johanna D. Moore and Keith Stenning, editors, Proceedings of the Twenty-third Annual Conference of the Cognitive Science Society, pp. 170-175, London: Lawrence Erlbaum Associates.
- Christiansen, M. H., Chater, N. 2008 Language as shaped by the brain. Behavioral and Brain Sciences, 31(5), pp. 489-509.
- Dahl, O. 2004 The growth and maintenance of linguistic complexity. Amsterdam: John Benjamins Publishing Company.
- de Boer, B. 2001 The origins of vowel systems. Oxford: Oxford University Press.
- de Boer, B. 2006 Computer modeling as tool for understanding language evolution. In: Gontier, Nathalie, Van Bendegem, Jean Paul, and Aerts, Diederik. Evolutionary Epistemology, Language and Culture – A non-adaptationist systems theoretical approach. Dordrecht: Springer. pp. 381-406.
- DeGraff, M. 2001 On the origin of Creoles: A Cartesian critique of Neo-Darwinian linguistics. Linguistics Typology, Vol. 5, issue 2/3, pp. 213-310.
- Ellis, N. C., Larsen-Freeman, D. 2009 Language as a complex adaptive system. Wiley-Blackwell.
- Everett, D. 2005 Cultural Constraints on Grammar and Cognition in Pirahã: Another Look at the Design Features of Human Language. Current Anthropology, Vol. 46, pp. 621-646.
- Ke, J. 2004 Self-organization and Language Evolution: System, Population and Individual. PhD Thesis. Department of Electronic Engineering, City University of Hong Kong.
- Kirby, S. 1999 Function, selection and innatenes: the emergence of language universals. Oxford: Oxford University Press.
- Kirby, S. 2002 Natural language from artificial life. Artificial Life, 8(2):185-215. Artificial Life, Vol. 8, issue 2, pp. 185-215.
- Kolmogorov, A. N. 1965 Three Approaches to the Quantitative Definition of Information. Transmission . Problems Inform. Vol. 1, 1, pp. 1-7.
- Kusters, W. 2003 Linguistic complexity. PhD Thesis, Universiteit Leiden. Utrecht: LOT Dissertation Series.
- Larsen-Freeman, D., Cameron, L. 2008 Complex Systems and Applied Linguistics. Oxford: Oxford University Press.

- Liljencrants, L., Lindblom, B. 1972 Numerical simulations of vowel quality systems. Language, Vol. 48, pp. 839-862.
- Lupyan, G., Dale, R. 2010 Language structure is partly determined by social structure. PLoS ONE, Vol. 5, issue 1.
- McMahon, A. M. S. 1994 Understanding language change. Cambridge: Cambridge University Press.
- McWhorter, J. H. 2001 The world simplest grammars are creole grammars. Linguistics Typology, Vol. 5, pp. 125-166.
- McWhorter, J. H. 1999 The Afrogenesis Hypothesis of Plantation Creole Origin. in Huber, M; Parkvall, M, Spreading the Word: The Issue of Diffusion among the Atlantic Creoles. London: University of Westminster Press.
- McWhorter, J. H. 2005 Defining Creole. Oxford: Oxford University Press.
- McWhorter, J. H. 1998 Identifying the creole prototype: Vindicating a typological class. Language Learning, Vol. 74, issue 4, pp. 788-818.
- Meisel, J. 2011 Bilingual language acquisition and theories of diachronic change: Bilingualism as cause and effect of grammatical change. (to appear)
- Miestamo, M., Sinnemäki, K., Karlsson, F. 2008 Language complexity: typology, contact, change. Amsterdam: John Benjamins Publishing Company.
- Milroy, L. 1980 Language and Social Networks. Wiley-Blackwell.
- Mitchell, M. 2009 Complexity: a guided tour. Oxford: Oxford University Press.
- Nettle, D. 1999 Using social impact theory to simulate language change. Lingua, Vol. 108, pp. 95-117.
- Pellegrino, F., Marsico, E., Chitoran, I., Coupé, Ch. 2009 Approaches to Phonological Complexity. Berlin: Mouton de Gruyer.
- Perkins, R. D. 1992 Deixis, grammar and culture. Amsterdam: John Benjamins Publishing Company.
- Sampson, G., Gil, D., Trudgill, P. 2011 Language complexity as an evolving variable. Oxford: Oxford University Press.
- Senghas, A. 2003 Intergenerational influence and ontogenetic development in the emergence of spatial grammar in Nicaraguan Sign Language. Cognitive Development, 18 pp. 512-531
- Senghas, A. 2005 Language emergence: Clues from a new Bedouin sign language. Current Biology, Vol. 15, issue 12, pp. 463-465
- Steels, L. 1996 Emergent adaptive lexicons.
- Steels, L. 1997 Constructing and sharing perceptual distinctions.
- Steels, L. 2000 Language as a complex adaptive system.

- Trudgill, P. 2001 Contact and simplification: Historical baggage and directionality in linguistic change. Linguistics Typology, Vol. 5, pp. 371-374.
- Weerman, F. 2010 Diachronic change: early versus late acquisition. Bilingualism: language and cognition, Vol. 14, issue 2, pp. 149-152.
- Wray, A. 2002 The transition to language. Oxford: Oxford University Press.