Effects of Social Structure on Establishing Lexical Conventions in a Computational Model of Task-Oriented Primeval Dialogue

Martin Bachwerk

April 2013

Declaration

This thesis has not been submitted as an exercise for a degree at this or any other university. It is entirely the candidate's own work. The candidate agrees that the Library may lend or copy the thesis upon request. This permission covers only single copies made for study purposes, subject to normal conditions of acknowledgement.

Martin Bachwerk

Summary

In the field of language evolution, the only way of obtaining empirical data for most of its parts is with the help of computational models and simulations. As a consequence, a large number of different modelling approaches have been introduced in the area over the past two decades, with the level of detail in these models ranging from highly abstract formalisms to experiments with humanoid robots. However, and to some degree precisely because of this discrepancy in the modelling approaches that are spawning in the field, the results obtained within different research projects remain largely incomparable to each other due to the unbridgeable discrepancy in the assumptions made by the respective models. Furthermore, many of these assumptions, such as the availability of telepathic communication, to name the most striking one, are not strongly supported by the findings from the corresponding fields, leaving whole models that have been built on top of these open to questioning.

The current project proposes a new computational framework of language evolution, built on top of an earlier prototype-stage version of the Language Evolution Workbench. The main principle of this framework is that a model should only integrate features that are based on widely accepted findings. Accordingly, only a handful of basic assumptions are made within the model, with almost 50 different parameters allowing the experimenter to configure the remaining properties of the simulated world, agents and the interactions between the two according to his needs. In principle, the proposed framework could be extended to simulate the evolution of a wide variety of linguistic structures, such as phonology, morphology, syntax, etc. However, in the current implementation of the model, the emphasis has been placed on the arguably most fundamental element of language, namely a conventional symbolic lexicon.

Focussing on the task of lexicon acquisition, a range of basic as well as more complex experiments has been performed with the help of the model and evaluated within this research project, in particular with respect to the similarity of the properties of emergent lexicons to those of real human languages. The results of these experiments suggest that agents interacting in pairs or triads have a significantly better chance of developing a reliable lexicon than those who have a significantly larger number of randomly interchanging partners. Given that the common consensus in the field is that language has evolved at the advent of increased hominid group sizes, further experiments have investigated the effects of different social configurations on the evolutionary process. So far, all of these appear to confirm that the early adopters of language need to spend a significant amount of time communicating with one or two select partners in order to have any chance of acquiring a useful lexicon, indicating that language emerged in small subgroups of a population and then spread throughout the whole population, e.g. as a result of group-internal natural selection. However, future experiments involving universally popular or particularly influential agents might still readjust these findings.

Acknowledgements

It is a well-known fact that doing a PhD is not like any other job that one can leave behind when one heads home. No, it constantly follows you around through life, including the evenings, the weekends and sometimes it doesn't even let you go on holiday without packing it with you in your already overweight bag (admittedly, part of the blame here has to fall on Ryanair!).

For this I would like to deeply apologise first and foremost to my loving wife Tahmina and hope that you can forgive me. There are no words that can express how greatly thankful I am for the unwavering support and understanding, with which you have always welcomed me. You are the one who is motivating me to strive for greater things and for this I will be thankful to you always.

My parents also deserve a special mention for never questioning my endeavours, no matter how unexpected, and continuously supporting me, especially in times when doubt and uncertainty was hanging thick over my head. Also, I hope you know how greatly I appreciate your financial support throughout my years at school; without it, none of this would have been possible.

On the academic level, I would like to express a big *kudos* to Carl Vogel, my ever inspiring supervisor who is never short of new ways of looking at things, a feature that has never ceased to amaze and inspire me throughout the years as a PhD student.

Finally, a big hug goes out to Alfredo, Anne, Daniel, Erwan, Francesca, Gerard, Hector, Liliana, Oscar, Roman and Stephan for making me feel at home on the 2*nd* floor of our lovely Georgian dwelling tucked in at the far end of the College. It is never easy to work on any project on your own, but with you around it never felt that way, so thank you for that, you crazy gang!

"HERMOGENES: For my part, Socrates, I have often talked with Cratylus and many others, and cannot come to the conclusion that there is any correctness of names other than convention and agreement. For it seems to me that whatever name you give to a thing is its right name; and if you give up that name and change it for another, tha later name is no less correct than the earlier, just as we changed the names of our servants; for I think no name belongs to any particular thing by nature, but only by the habit and custom of those who employ it and who established the usage."

- *Cratylus*, PLATO

Contents

Chapter 1

Introduction

The question of how human language, one of the key properties of which is an extremely high level of expressivity, stemming from open-ended and recursive construction (cf. Hauser, Chomsky, & Fitch, 2002), could have emerged from an animal-like communication system is at the same time fascinating and unresolved. In fact, while most other traits that make one human have been relatively well studied and documented, the emergence of language,[1] which is arguably the most unique of all human traits, remains largely a mystery. Accordingly, it seems imperative that in order to gain a full understanding of human nature, one should surely be able to reasonably explain the emergence and the evolution of language. In addition to that, the evolution of language has also broad ramifications in the area of natural language and speech modelling. In particular, if one could understand how our extremely distant ancestors learned to associate meanings with seemingly arbitrary symbols, be those symbols gestures or sounds, then one should also have an easier time of engineering artificial systems capable of comparable levels of intelligence.

1.1 Research Area

Interest in the origin of human language goes back several centuries, yet due to a lack of any solid physical or empirical evidence, the majority of the papers in the field were historically (and arguably still are) so conjectural that the *Société Linguistique de Paris* had felt the need to ban any kind of publications in this area in 1866. It is only since the reappearance of 'evolutionary linguistics' in the Linguistic Bibliography in 1988 that the question of language evolution has slowly resurfaced again and began to be objectively scrutinized in an increasingly empirically oriented fashion,[2] with a variety of research fields becoming deeply involved in the area. In particular, the number of biological (e.g. Pinker, 1994; Fitch, 2010), social (e.g. Dunbar, 1997; Dessalles, 2007) and linguistic (e.g. Bickerton, 1990; Carstairs-McCarthy, 1999; Heine & Kuteva, 2007) theories on how our distant predecessors could have invented such a complex communicative tool as human language is constantly growing, with hundreds of papers and numerous manuscripts being published every year.

One of the reasons why the number of theories on the same topic (even though, admittedly, a fairly complex one) is constantly increasing, with the newer proposals rarely replacing the older ones, but rather providing more and more alternatives for each aspect of the problem, is that there is basically no empirical evidence available to researchers that could help them make a decision one way or the

[1] Here and henceforth, 'language' will be used to refer to 'human language'.

[2] Also attracting significant public interest, as is exemplified by the broad coverage of recent findings by Atkinson (2011) in *Science* and *The New York Times* among others.

other. There are numerous studies from the fields of experimental semiotics (e.g. Healey, Swoboda, Umata, & Katagiri, 2002; Galantucci, 2005; Garrod, Fay, Lee, Oberlander, & Macleod, 2007; Scott-Phillips & Kirby, 2010), primatology (see King, 1999; Burling, 2005) and language acquisition (e.g. Tomasello, 2003; MacWhinney, 2005) that certainly contribute with some amounts of empirical data that is very closely related to the question of language evolution.

However, such studies will always be set up significantly differently from the situation that early hominids who made the first step towards inventing human language found themselves in. For example, in an experiment by Scott-Phillips, Kirby, and Ritchie (2009) where participants needed to come up with a new communication system, they might have had to invent a new way of communicating, but they were always aware of how conventionalized communication works in principle and how it is normally used by humans, thus putting them at an unfair advantage when compared to the actual inventors of first such system. Similarly, in studies on child language acquisition, the children are being constantly helped by a proficient language user, which could not have been the case for the very first people who started using lexical conventions as a means of communication.

1.2 Computational Framework

Since the advent of cheap and powerful computing power, more and more research fields, in which obtaining real data is extremely hard or literally impossible, are turning to computational models for what could be referred to as simulated evidence. Within this trend, the field of language evolution is no exception, with a multitude of computational models being developed over the last two decades. In effect, just as this project does, these efforts do not provide neither definitive verification nor clear refutation of any theory of language evolution in its entirety, but, modulo underlying assumptions, incrementally build up empirical data for or against components of theories. Within these models, a wide variety of methodologies has been employed over the years, including neural networks (e.g. Cangelosi & Parisi, 1998; Hazlehurst & Hutchins, 1998), self-organizing maps (e.g. Lindh-Knuutila, Honkela, & Lagus, 2006; Worgan & Damper, 2008), Bayesian inference (see Kirby, Dowman, & Griffiths, 2007) and discrimination trees (see Steels & Kaplan, 2002). On top of that, there is a sizeable population of researchers working with embodied agents (see Mirolli & Nolfi, 2010), with some of these including real-size robots (see Steels & Spranger, 2009), underlining the popularity of the field, as well as the amount of funding involved.

Despite the apparent interest in computational models of language evolution, however, there is a clear lack of a widely accepted methodological framework in the field, as indicated by Vogt and de Boer (2010) among others. What this means is that the different models of language evolution are all based on different sets of assumptions that are generally built into the models and can often be heavily disputed, as will be exhibited in section 2.3. This situation makes it particularly challenging to evaluate the relevance of the results obtained with any particular model, as well as compare the results of any one model's configuration with those based on different assumptions.

Considering the above discussion, the field of language evolution would benefit tremendously if a more general framework were available for its branch of computational research that would be compatible with the different evolutionary theories from the many disciplines involved. The framework should be strongly situated in findings from psychology, anthropology, primatology and other related fields, ensuring that no assumption is made without there being strong empirical evidence supporting it. In saying this, it is not implied that every computer simulation should explore every possible combination of the many potential parameters of the process. Rather, it is important that the assumptions of any model are both clearly defined, but also adjustable by other researchers without the need for developing a new system from scratch every time. The goal of this research project has been thus to come up with a framework that would be well situated within the latest findings from the related fields, but also highly generic, allowing one to conduct experiments with an innumerable combination of assumptions and parameters.

For the purpose of developing a generic computational framework suitable for investigating the evolution of language, an early prototype of the Language Evolution Workbench (LEW), as presented by Vogel and Woods (2006), has been adopted within this project. While not being extremely detailed (or efficient) from the beginning, this model clearly showed the most potential for being developed into a reasonable framework as it attempted to avoid integrating any assumption that was not necessarily given as a fact, instead offering the researcher a multitude of parameters to choose his own configuration from. In order to make the LEW into an easily adaptable framework, the model has been significantly modified and extended throughout the years of the research project, with new parameters being introduced throughout, allowing for even more refined experiments (there are almost 50 at the moment of writing), and intermediate simulation results being published by Bachwerk and Vogel (2010), Bachwerk and Vogel (2011), Bachwerk (2011) and Bachwerk and Vogel (2012).

Finally, it has to be said that the issue of language evolution is an extremely broad one, with a multitude of linguistic, pragmatic, psychological and neurological aspects of it needing to be addressed (for a detailed list of open questions see de Boer & Zuidema, 2009). Accordingly, with time limitations in mind, it was decided to hold off with investigations of the emergence of the more complex linguistic structures, such as syntax, grammar and questions. Instead, the primary effort at this embryonic phase of the framework's development has been put into the seemingly simple, yet as it turns out extremely complex, task of agreeing on the very first lexical conventions or, in other terms, building up a significantly large conventional lexicon.[3] Admittedly, the majority of current linguistic theories would claim that a lexicon of a language contains a significant amount of morphological variation, which is not modelled in this work. However, even if this feature is omnipresent in possibly all modern lexicons, the main definition of a lexicon is still being a set of conventional meaning-form mappings, or words. Accordingly, the only reason for which a lexicon generated within the proposed model could be faulted for is its holistic simplicity, which, while potentially ineffective on the bigger scale, should not disqualify it from being, after all, a lexicon. Further motivation for the task selection, as well as the major challenges and proposed solutions of modelling lexicon acquisition are described in the following section.

1.3 Word Learning Task

Human language is a highly complex form of communication, consisting of a number of sophisticated interacting elements that contribute to the intricate mechanism. However, I would argue that the most important mechanism of language, without which its usefulness as a communication tool would surely have been unimaginable, is the lexicon, which appears to be absolutely essential for the emergence of language in evolutionary terms, as singled out by Gärdenfors (2004) in his account of primeval cooperation. Having said that, I do not intend to claim that the sole task of human language

[3] It should be noted that, despite the choice to focus on lexicon formation at this point, the LEW model can be extended to deal with more advanced features of language in an incremental manner, i.e. without the need for being significantly overhauled for each such extension due to the highly modular structure placed at its core.

is transmitting information and that all other aspects, such as pragmatics in particular, are rather secondary, if not irrelevant. Nevertheless, it is hard to imagine how our distant predecessors could have been motivated to acquire such a complex social skill as language if it did not serve at least some informative purpose.

It comes then as no surprise that the ability to convey information of some sort lies at the heart of almost every theory of language evolution. In particular, it appears that in order for someone to agree on the usage of a certain sound, or sound combination, there needs to be a learning mechanism that would reinforce its correct use. If, however, one assumes that there is no inherent informative function to the use of language, like Dunbar (1997) does in claiming that language has evolved as a more effective replacement to manual grooming in an ever growing population of early hominids, then it remains unclear how any kind of meaning could have ever been associated with the soothing, grooming sounds (cf. Bickerton, 2003). In fact, even if one assumes that, in such a scenario, the invention of a lexical item, i.e. a sound form referring to a meaning, might have occurred by accident or via repetition-induced ritualization (cf. Pika & Mitani, 2009; Ragir & Savage-Rumbaugh, 2009), this still implies that the communicative function of the different sounds was achieved by correlating their use with an informative component.

The conclusion that can be made from the considerations outlined above is that, for a set of lexical conventions to be established in a population of potential interlocutors, at least some degree of agreement about the meaning of the different sound combinations, i.e. words, must have been achieved between them, regardless if this happened intentionally or not. Admittedly, the task of agreeing on a meaning of a certain word might appear if not completely straightforward, but still quite feasible to an experienced user of language (even in a foreign land, the language of which is completely unknown to a person). But if one imagines two interlocutors who have never experienced a conventionalised communication system in action, a number of questions instantly arise regarding the interpretational procedure between these two pioneers:

- How does the hearer know precisely enough what the speaker is talking about?
- What happens if two interlocutors seem to know the words, but in reality are talking about two completely different things (or vice versa)?

Socrates (in Plato's "*Cratylus*"), Corballis (2000) and Hutchins and Johnson (2009) would claim that (at least during the initial stages of language evolution) symbols would be predominantly, if not exclusively iconic and hence pose no difficulty during interpretation. However, the *indeterminacy of translation* thesis by Quine (1960) suggests that even iconicity may be quite unreliable,[4] meaning that the above questions remain largely unanswered by the existing literature on the emergence of language. Presumably (and understandably), this is due to a complete lack of linguistic evidence from the evolutionary process itself. Advances in the areas of computer science and artificial intelligence have significantly contributed to filling this gap by constructing models of language evolution and attempting to generate actual empirical data on the topic. Having said that, it has already been noted in the previous section that few of the computational approaches existing to date have ventured deeper into investigating the arguably quintessential task of lexicon acquisition with a realistic set of assumptions, which was also one of the main motivations for conducting this particular project.

[4]It seems unlikely that all individuals share the same internal representation of things that would allow them to easily interpret the corresponding 'iconic' symbols. As an example, just consider the variety with which animal sounds are represented in different languages, let alone animal names.

The current project aims thus to contribute the research in the field of language evolution with a model that is quite abstract, yet sufficiently realistic, particularly in terms of assumptions made regarding the cognitive capacities of early humans. The model will furthermore implement a variety of cognitive and interpretational constraints that will be applied to a number of different hypothetical population sizes and social structures with the goal of observing the influence of such configurations on the lexicon formation during the very early stages of language evolution. Taking all of the above into consideration, the main questions that will be addressed throughout the experiments performed with the model can be stated as follows:

- What level of agreement is necessary (or sufficient) for a shared lexicon to be established in a population of communicating individuals?
- What are the social constraints on the establishment of linguistic conventions in a population of simulated cooperating agents?

The first question that has been asked above deals specifically with the assumption that language in general, and a conventional lexicon in particular could only have emerged in a cooperative setting, in which the success of a communicative exchange has a direct influence on the livelihood of the communicators, i.e. where correctly interpreting the information transmitted by an utterance is paramount. The intuitive answer to this question would suggest that the higher the agreement requirement, the more uniform would the lexicons of the interlocutors be. However, imposing too high a success requirement on the actors from the very beginning bears with itself a high risk of them getting quickly frustrated with the task at hand and abandoning the attempt to communicate via lexical conventions altogether. A possible insight into this dilemma has been provided by psychological studies that have shown that humans tend to perform at an optimal level when they are able to achieve a relatively high level of success in their task, yet at the same time experience a significant number of near-miss failures that motivate them to keep on repeating the same activity over and over again (cf. Chase & Clark, 2010).

Additionally, if one considers the amount of apparently acceptable, or simply ignored miscommunication in human language, taken together with the relatively high levels of synonymy and homonymy, it raises doubts over whether the formulation of the question should be focussed on what is required for a fully shared lexicon to be established in a population. Instead, the real question is what kind of social constraints result in a communication system, the properties of which are most similar to those of the modern day human languages, which is the second investigation point that has been raised above. When looking for an answer to this question, lexicons of individual agents and whole populations will be evaluated based on two main criteria: the observed communicative success achieved with the help of such lexicons and the comparability of the emergent lexical systems in terms of the lexical properties of existing human languages (a detailed description of the evaluation measures is provided in section 3.4). In the latter case, special emphasis will be put on such arguably universal aspects of modern human languages as the tendency to avoid synonymy, yet tolerate some amount of homonymy (cf. Carstairs-McCarthy, 1999).

1.4 Thesis Outline

With the overall motivation behind this research project, as well as its main aims and objectives being outlined above, this section will present the plan of action that will be executed throughout the remainder of the manuscript. As is appropriate when working on any kind of solution, the first thing that one has to do is properly investigate the problem, along with the solutions that have been previously proposed by others. Accordingly, the following chapter 2 will begin with a review of the current state of research in the field of language evolution. Within this exercise, section 2.1 will begin with an introduction into the field of language evolution, followed by a presentation of some related experiments from the field of communication science in section 2.2. Next, the most prominent modelling approaches in the field are discussed in the section 2.3. Section 2.4 then presents a discussion of the role of cooperation during the emergence of language and introduces the concept of a signaling game that will play a pivotal role within the computational framework presented in this project.

Having provided the reader with (hopefully) sufficient background knowledge, chapter 3 then goes on to present the proposed computational framework in extensive detail. At the beginning of this chapter, the general assumptions and data structures of the implemented version of the model are presented in section 3.2. Following that, the design of the interactions is presented in section 3.3. This section is particularly important for the understanding of the framework as interactions between the model's agents form the core of the project and are also the main component of the conducted simulations. Finally, section 3.4 presents the evaluation measures employed in the experiments conducted with the help of the model for the determination of their relative success or failure.

Clearly, no model can stand on its own without an analysis of its fundamental features. Chapter 4 presents precisely such an analysis, which is based on the results of three experiments performed with a stripped down version of the model. The goal of this chapter is to determine any analytical limits imposed by the model on the different properties of the lexicons emerging within the corresponding simulations in a very basic configuration. For this purpose, experiments with only a very small meaning space and either two, three or ten agents, respectively, are presented and fundamentally evaluated in the three main sections of the chapter. Admittedly, such a basic configuration is a very poor approximation of reality. However, establishing a good understanding of the framework's limits and capabilities is essential for ensuring the plausibility of future, more advanced experiments.

Two such experiments with more complex setups are then presented in chapter 5. The main focus of these experiments is on the different social configurations of the simulated agent populations, motivated by the findings from the preceding chapter, suggesting that agents in smaller groups tend to develop more reliable lexicons, even if the same amount of communicative time is allocated per agent in the different groups. Accordingly, the question that is investigated in this chapter is how the advantages of interacting in e.g. pairs could be transmitted onto the population level, helping all of its members learn the same language.

One solution for this is proposed in section 5.2, in which agents are either allocated fixed partners for the duration of the simulations with a higher-than-average pair-internal interaction rate, or have exclusive interaction partners that are evenly rotated throughout the simulation runs. In the following section 5.3, it is proposed that agents should be able to adjust their social ties based on the previous successes or failures with different partners, contrary to the more common approach of having the social structure of a simulated population fixed throughout any experiment. As usual, the results of both experiments outlined above are evaluated with respect to the different properties of the resulting agent and population lexicons, as well as the communicative success rates of agents. Finally, the goals that were set for the research project as well as all of the obtained results, will be summarized in chapter 6.

Chapter 2

State of Research

This section presents a critical review of the literature from the field of language evolution that is relevant for the proposed research project. While there is an abundance of additional scientific material that is in one way or another relevant to the overall question of language evolution, the presented review will concentrate on works from two areas: communication science and computer science. This selection has been made with regard to the focus of this project, which is placed on providing a computational framework that can explain how the social organisation of our predecessors could have influenced the establishment of first linguistic conventions within the process of language evolution.

Notably, other models presented in this section will mostly provide fairly restricted representations of certain conjectured evolutionary scenarios, with the majority of corresponding assumptions being built into the models. In contrast to that approach, the goal of the current research project is to construct a computational framework that would allow for an exhaustive array of experiments of the language evolution process to be executed, leaving it up to the researcher to evaluate the selected assumptions based on the results of the simulations.

2.1 Language Evolution

One of the most dominant themes in the literature on language evolution is that of human sociability, which is also the main focus of Dunbar's (1997) very prominent 'gossip as grooming' theory of language evolution. Dunbar claims that at some point in its evolution, the hominid species were forced to relocate to territories that were much more open than the woods they previously inhabited, thus making them more susceptible to predation and as a consequence forcing them to live in larger groups. As a matter of fact, a number of studies (e.g. Aiello & Dunbar, 1993; Kudo & Dunbar, 2001) evaluating the limitations and the effort required for a brain to memorize acquaintances suggest that the group size of our predecessors could have gone up to as many as 150 members around 250,000 years ago. Given that the findings are based on solid empirical evidence, there should be no reason for contesting that a strong evolutionary pressure was posited at this stage of human evolution for selecting an adaptation that would facilitate the process of interaction between members of the species, i.e. a proto-language. Dunbar then goes on to argue that, if the above was the case, then 'traditional' grooming alone could not have sufficed for the maintenance of group-internal bonds as the time spent on it would have significantly exceeded the minimum time threshold required for other survivally critical activities such as scavenging for food and sleeping and so the species adapted by making the traditionally time-consuming grooming more efficient.

In his critical reviews of the field of language evolution, Bickerton (2003, 2007) responds to Dunbar's theory by pointing out that baboon groups are actually larger than those of early hominids and thus should have undergone the same selective pressures. However, his main criticism of the 'gossip as grooming' theory is that language must have evolved through a stage where its structure was extremely primitive and the vocabulary very limited, rendering the amount of actual information that could have been transmitted hardly sufficient for anything that could be classified as 'gossip'. Furthermore, it is also unclear how any lexical items could have been agreed on during gossip as it can potentially involve a very broad frame of reference, just as any other form of communication for that matter, but does not involve any non-communicative, or natural punishments or reinforcements that one could learn from (cf. Barrett, 2006, on the need for both types of reinforcements when learning a signaling system).

As an alternative motivation for our ancestors to attempt to extract meanings from the behaviour of their conspecifics, Bickerton (2002) outlines the foraging lifestyle adapted by our predecessors some two to three million year ago. In particular, Bickerton argues that larger groups of hominids used to split up into smaller scavenging parties that went out in search of more prosperous areas. However, the scavenging groups were not large enough to exploit the newly discovered territories on their own and so each party had to describe their findings on return, so that the group could decide if and which of the newly found areas were worth relocating to. In a scenario like this, it is important to note that the success of any primeval utterance would be of extreme significance, considering that it can be extremely costly to relocate a large group just to find out that the new territory turns out to be a scorching desert. As can be imagined, this stands in stark contrast to the grooming theory where success in communicating actual information is much less critical, if at all significant. One of the main goals of this research project is precisely to evaluate the realistic chances of language emerging in such scenarios with the help of a computational model. Importantly, the approach will necessarily abstract over many features of actual human language use and will thus not fully reflect all of the nuances discussed above. Having said that, the model will be able to provide empirical data for an issue that can hardly be empirically examined in other ways.

2.2 Communication Studies

While different evolutionary accounts of language evolution can provide possible scenarios of how and why the adaptation might have occurred, these theories are usually unable to provide a significant amount of solid empirical backing, which is quite understandable considering that there are no fossils of language *per se* and that the phenomenon is hardly reproducible in a laboratory experiment involving humans directly. Nevertheless, a number of studies from different research fields in communication science and linguistics could provide some insights into the mystery of language evolution. Several relevant findings from two of these fields are outlined in the sections below.

2.2.1 Experimental Pragmatics & Semiotics

In the last decade, the field of communication science has seen a major increase in the number of research programmes that have gone beyond the more conventional studies of human dialogue (e.g. Garrod & Anderson, 1987; Garrod & Doherty, 1994) and attempted to reproduce the emergence of conventionalized communication systems in a laboratory (e.g. Galantucci, 2005; Garrod et al., 2007; Healey et al., 2002). The ever growing pool of such studies has prompted Galantucci (2009) to suggest that this line of research should be given a specific term, for which he proposed *experimental semiotics*. In his seminal paper, Galantucci goes further to define experimental semiotics as a more general form of *experimental pragmatics*, whereby he defines that the former *"studies the emergence of new forms of communication"* and the latter *"studies the spontaneous use of pre-existing forms of communication"* (p. 394, (Galantucci, 2009)).

When considering the emergence of the very first conventionalized communication system in the history of human evolution, experimental semiotics certainly provides a novel way of approximately reproducing the process under laboratory conditions. However, experimental semiotics suffers from one major flaw, namely that the results of the conducted experiments cannot be transferred to the question of the primeval emergence of language without the caveat that the subjects of the present-day experiments are quite familiar with the concepts of conventions and communications systems (even if they are not allowed to employ any existing versions of these in the conducted experiments), while our ancestors who somehow managed to invent the very first conventional signaling system, by definition, could not have been aware of these concepts.

In the current research project, the aim is to employ findings from experimental pragmatics and semiotics with the goal of constructing a computational model of language emergence that is maximally representative of human communication. In particular, the interactive alignment model presented by Pickering and Garrod (2004) will be taken as the underlying framework for a wide range of experiments. This model suggests that interlocutors tend to align their representations on all levels, including their perspectives on a situation, lexicon, syntax, style, etc., allowing them to achieve higher communicative success. The explanatory proposal of this approach is that internal world models are made 'public' with projections onto vocal and visual channels and that imitation behaviours eventually give rise to external systematicity between such inner worlds. The current project will concentrate on aspects of world model and lexicon alignment between interlocutors, as these two elements of Pickering's model appear to be absolutely essential for the emergence of language in evolutionary terms, as depicted by Gärdenfors (2004) in his account of primeval cooperation.

2.2.2 Language Acquisition

Similarly to experimental semiotics, language acquisition studies can provide valuable input about how humans process linguistic information, in particular at the very early stages in their cognitive skill development. Admittedly, as is the case with experimental semiotics, findings from language acquisition research need to be put in perspective if one wishes to transfer these onto the field of language evolution. The main issue hereby is the fact that infants are being actively taught the use of a language by fully competent speakers, whereas studies of the emergence of language, by definition, investigate how such linguistic competence could have been acquired in the first place. Nevertheless, one can take advice from the findings in language acquisition studies regarding the minimal cognitive abilities that appear to be necessary for children in order to be able to learn a language.

One account of language acquisition that may be particularly relevant to the question of language evolution is the usage-based theory of language acquisition by Tomasello (2003). With a strong background in primatology and interest in the emergence of language, Tomasello provides a number of insights into the cognitive skills required for learning a language. In particular, Tomasello singles out three social skills as essential to the acquisition of linguistic competence: the establishment of joint attentional frames, the understanding of communicative intentions, and role reversal imitation. In summary, these three competences suggest that in order to learn a language, one needs to be able to share one's attention on the same object or event as one's interlocutor, assume that the speaker is referring to the shared attentional frame with his utterance, and produce own comments in accordance with the communicative scheme employed by others (similarly to the output-input principle observed by Garrod & Anderson, 1987).

Going beyond the social skills, Tomasello claims that children also employ a number of assumptions that facilitate the learning of meaning-form mappings, such as the assumption of mutual exclusivity (see Markman, 1989), according to which an object can only belong to one category and thus only correctly referred to by the word associated with that category, and the principle of contrast (see Clark, 1987), which states that any two distinct forms should differ in their meanings. Tomasello suggests that a mechanism like the principle of contrast is one of the main tools that explains the explosive rate of lexicon acquisition, and inevitable synonymy avoidance, in human language. However, the existence of the principle of contrast and similar learning biases is yet to be conclusively proven to exist in humans, not to mention primates, with experimental results from work by Gathercole (1998) and Deuchar and Quay (2001) among others suggesting that these constraints are perhaps not present in our brains at all.

2.3 Modelling Approaches

More than two decades ago, Hurford (1989) published an influential paper on the evolution of the so called 'Saussurean' sign. The uniqueness of this paper was constituted by the fact that it presented what was probably the first evolutionary linguistic study that was in its entirety based on a computational simulation model. In the model, simulated agents employed three distinct interaction strategies during their lifespan, with their communicative success having a direct effect on the reproductive potential of the agents. With this approach, Hurford was able to show that, given the assumptions of the model, the agents that employ the bi-directional, or 'Saussurean', sign clearly outperform agents who employ alternative strategies (the main characteristic of these alternative strategies was that they utilized *two* sets of rules – one for interpretation and one for production – which were constructed by either directly imitating the productive behaviour of others, or by re-calculating one's own production based on the interpretative behaviour of others). While the model presented by Hurford is not flawless and could be strongly criticized, e.g. due to the meaning space being limited to only six objects, or the impossibility of synonyms emerging during the interactions, the truly ground-breaking part of the approach was that it enabled Hurford to provide empirical support for his hypothesis without relying on predominantly speculative historical linguistics, or very complicated human experiments.[1] In the following sections, several modelling approaches will be presented that build on Hurford's simulations and that have been particularly successful in the field of language evolution.

2.3.1 Iterated Learning Model

The so called Iterated Learning Model (ILM) (see Oliphant, 1999; Kirby, 2001; Kirby & Hurford, 2001) is arguably the most prominent computational model of language evolution of those that have clearly been inspired by Hurford's study described above. At the core of this model lies the assumption that at any given time, a community of agents can be divided into two groups. The first group consists of the teachers, i.e. agents that have an established language and whose sole task is to produce

[1] Another ground-breaking model by Hurford (1991) demonstrates the possibility of genetically driven evolution of a critical period for language acquisition.

linguistic input for the second group – the learners. The learners are agents who are just listening to the teachers' utterances and never speak themselves. The iterative component of the model is represented by an ageing mechanism that defines when learners become teachers, and when 'old' teachers can be replaced by new learners. The first application of this model presented by Oliphant (1999) was to simulate the cultural evolution of a simple lexicon. However, most recent studies performed with the help of the ILM have concentrated on the emergence of syntax from a holistic vocabulary as a consequence of a learning bottleneck and without the need for natural selection or any kind of explicit intervention. In contrast, in the model explored in my work, the distinction between teacher and learner does not exist since all are equal participants, if not equally experienced, in language creation.

While the results of the corresponding simulations have been considered predominantly positive by Kirby et al., the first versions of the ILM have arguably suffered from a broad range of issues that needed to be addressed in order for the simulation results to bear any significance on the question of language evolution in general and the emergence of syntax in particular. For example, the agents in the ILM always attempt to analyse any utterance they hear, even if it can be stored and successfully reused holistically, thus forcing the emergence of compositionality. In a world restricted to 100 meanings, the motivation for doing this is not quite clear. Furthermore, considering that the experiments are designed in a way that prevents the learners from observing an utterance for every meaning (learning bottleneck), making sure that a holistic language simply cannot be learned, the agents will inevitably produce more and more compositionally constructed utterances with every iteration until the language is fully syntactic.

The biggest issue with the ILM simulations, however, lies in the explicit transmission of meaning-form pairs in every interaction between agents. Not only is such telepathic communication unrealistic, but it also explains the relative ease with which the agents are able to learn a language that is completely stable, fully expressive and perfectly unambiguous (i.e. synonym and homonym free), none of which is a property of real human language.[2] Surprisingly, this aspect of the ILM has not been widely criticized, with only a small selection of scholars attempting to bypass explicit meaning-form transmission in their studies. In one such study, Smith (2001) has constructed a version of the ILM in which agents needed to discriminate the subject of an utterance by categorizing the properties of several objects without any additional help, e.g. in the form of meaning transmission or explicit feedback. The results of the corresponding simulations show that agents are still able to achieve high levels of communicative success even without telepathy. However, further experiments have shown that an increase in the number of features deteriorates the learning process significantly, with the additional effect of very high lexicon synonymy levels appearing in the population.

In an attempt to tackle this problem, Smith (2005) has constructed another model without explicit meaning-form transmission, but with a number of representational and interpretational constraints such as the assumption of mutual exclusivity (see Markman, 1989) and the principle of contrast (see Clark, 1987) added to the system. The experimental results of this model suggest that integrating such constraints does indeed improve the communicative success of the emergent language, while keeping it comparable to existing human languages. However, as has been mentioned previously in section 2.2.2 above, the presence of these additional constraints in either humans or primates remains largely unproven, which suggests that this approach may not necessarily be a valid reflection of reality. In summary, there might be an argument for a weaker notion of telepathy being present at the early

[2] See section 3.4 for estimated values of synonymy and homonymy in the English language.

stages of language evolution that could help the agents establish an initial common ground when acquiring a lexicon. This could be achieved for example via a joint attention frame being established through gaze following, pointing, and similar indicators (cf. Tomasello, 2003), as attempted in the model by Gong, Minett, and Wang (2009), which will be presented in section 2.3.3.

One final remark in relation to the original ILM is that the separation of agents into teachers and learners is also highly contentious, considering that children begin producing words and utterances of their own long before they finish learning language. This weak point of the ILM has been addressed in the recent years by a number of proposed extensions, most notably one by Vogt (2005). In his experiments, Vogt has introduced so called *horizontal transmission* into the ILM by also allowing the 'children' of a population to produce utterances as well. In fact, children turned out to be the driving force behind the emergence of compositionality as they were quite frequently prompted to express meanings they never encountered before, thus introducing a sort of *implicit bottleneck*. Unfortunately, these experiments were largely based on the setup of the guessing game and were thus subject to such weaknesses of that approach like the presence of explicit feedback, as described further in the next section.

2.3.2 Language Games

Language games that are referred to in this particular context are evolutionary computer simulations that are based on a computational model introduced by Steels (1996). This model itself builds on the so called signaling games, as introduced by Lewis (1969), which were designed with the goal of conducting game-theoretic simulations of the emergence of communicative conventions, i.e. so called signaling systems. Like the ILM, the language games are significantly more complex than the original 2-state/2-term signaling games, but rather fall into the class of n-state/n-term systems described by Huttegger and Zollman (2011). Accordingly, the task of the agents shifts away from straightforward learning to inducing a meaning of an utterance from a given context, usually presented in the form of a game. Additionally, although the approach *per se* does not require embodied agents for the execution of valid simulations, most experiments performed on the basis of different language games by Steels et al. involve physical robots from small LEGO vehicles (Steels & Vogt, 1997) to AIBO dog-robots (Steels & Loetzsch, 2008) and from static Talking Heads (Steels & Kaplan, 2002) to strikingly human-like QRIO robots (Wellens, Loetzsch, & Steels, 2008).

Two variants of the language games employed by Steels and his colleagues that are particularly relevant for this research project are the *naming game* and the *guessing game*. The naming game (Steels & Vogt, 1997) is effectively the simplest form of language games used so far. It involves only one object that two agents interact about, both of whom know what the object is. The goal of this game is to observe the dynamics of agents converging on identical lexicons under the condition that if both agents understand the form used to describe an object, they instantly delete all of the competing forms from their lexicons (see also de Lara & Alfonseca, 2002, for alternative negative feedback mechanisms that can be employed in the naming game). The guessing game (Steels & Kaplan, 2002) extends the naming game by introducing a *context* of objects to the model. The context is usually a set of objects, from which the speaker selects one as the *topic* of his utterance, whereas the hearer needs to correctly discriminate the topic with the help of two modules. The first of these is the conceptualisation module, which stores discrimination trees that are used by the agent to characterize an object based on the categories that no other object in the context belongs to. The second one is the verbalisation module, which stores the linguistic forms that correspond to meanings (i.e. categories).

Finally, each time an agent experiences a particular meaning-form combination, lateral inhibition is applied to all competing mappings in order to eliminate synonymy from the agent lexicons.[3]

In criticism of the language games, it should be pointed out that the setup of even the most recent experiments involves explicit feedback at the end of each interaction, which is quite questionable as it is not based on any clear empirical evidence from studies in either primate or human language acquisition. Accordingly, the following sections will argue that communicative feedback experienced by interlocutors is usually implicit, if at all present. In fact, the implementation of the explicit feedback in the guessing game could be actually described as post-interactional telepathy. What happens in these experiments is that while the agents do not have access to the correct meaning at the time of an interaction, they *always* experience the exact meaning afterwards. This setup is equivalent to agents knowing the meaning of an utterance, but, for reasons similar to sporting interest, choosing to avoid using that knowledge at first to try and determine the meaning on their own. However, the robot's 'honest' attempt at decoding the meaning of an utterance on his own does not make him less of a 'cheat' if it resorts to taking the correct solution from his pocket in the end. In the following section, a computational model will be presented that fully avoids any kind of direct meaning-form transmission and also introduces several innovative ideas to the computational approach.

2.3.3 Coevolution Model of Language and Social Structure

As mentioned above, one of the main issues that Gong, Ke, Minett, and Wang (2004) see with the majority of computational models of language evolution is that they rely on explicit meaning-form transmission to be present in the system in order for the agents to learn a language. Furthermore, Gong et al. (2004) also identify three further shortcomings of the ILM and Language Game models presented in the preceding sections: failure to model syntax, use of random interactions, which ignore the potential influence of social organization of agents, and homogeneous populations. Since the focus of the current research project is equally on the pre-syntax phase of language evolution, the lack of syntactic structures in current models will unfortunately have to remain unanswered by this work, although all the other issues are of particular interest here, along with the solutions proposed by the so called Coevolution Model of Language and Social Structure.

According to Gong, Minett, and Wang (2008), the main purpose of their model is to reproduce the emergence of systematic lexical compositionality and syntax-like regularity in the lexicons of the simulated agents. For this purpose, two types of predicates are available in the model as potential meanings, with either one or two arguments respectively. The agents have the ability to express these predicates with the help of holistic, compositional and syntactic rules that are stored in their lexicons along with a corresponding confidence value, ranging between 0 and 1. During utterance interpretation, the hearer selects a set of applicable rules from his lexicon, and, depending if the combined weight of all utilized rules, together with the strength of the cue, exceeds the confidence threshold parameter, either transmits positive or negative feedback to the speaker, which in turn results in the agents either rewarding or penalizing the rules that they utilized during the interaction by a fixed value.

When it comes to agreeing on a conventionalized signal-meaning mapping scheme in a reasonable amount of time, the reasoning made by Gong et al. (2009) is that interlocutors need at least some

[3] Further variations of language games include predominantly embodiment-oriented approaches, such as experiments on perspective reversal (Steels & Loetzsch, 2008) and, most recently, the so called action game (Steels & Spranger, 2009). Since the model proposed in the current research project does not include any (abstract) embodiment features, these variations will not be analysed any further in the current work.

degree of help from either each other or the environment. What serves this purpose, they suggest, is the skill of intentionality sharing, e.g. via gaze following, pointing or other informative gestures, which has been identified by Tomasello (2003) as one of the crucial cognitive skills that allowed humans to develop language. In their model, Gong et al. (2009) have implemented this skill in the from of an *environmental cue*, i.e. a certain meaning that is shared between the two participants of an interaction. Since being able to always know what the speaker is talking about would take the model back to full telepathy, Gong et al. (2008) have introduced an additional parameter that defines the reliability of the environmental cue (set to 0.6 in the majority of reported experiments), which essentially defines the ability (either intentionally or otherwise) of two agents to share a very specific joint attention frame in an interaction.

A further interesting aspect of the model described above is the introduction of a number of various social structures and constraints. Firstly, Gong et al. (2008) introduce a single so called *popular agent* into their model, whose involvement in an interaction is defined by an additional *popularity rate* parameter, with the popularity of all other agents being equal. Building on the idea of an agent having a specific popularity among others, the modellers have conducted an additional experiment in which agent popularities followed a power-law distribution, as opposed to the usual equal distribution. In a third experiment, they then distributed 20 agents in two 10-agent groups and defined different proportions for inter- and intra-group interactions. The results of both these experiments indicated a boundary in terms of the popular agent's popularity rate, the popularity distribution's parameters and the intra-group interaction rate respectively, when it comes to establishing a community that is sufficiently integrated to learn a communal language.[4]

A further experiment conducted by Gong, Puglisi, Loreto, and Wang (2008) approached the aspect of social organization of a population of agents from an emergent perspective. In this experiment, agents were distributed across a virtual space, with a distance-based communication constraint being imposed on the possibility of two agents interacting with one another. Furthermore, agents moved closer together or further apart, depending on the success of their interactions. The analysis of the corresponding simulations suggests that different levels of clusterisation emerge in the population, depending on the size of the virtual space and the distance constraint (see Gong & Wang, 2005, for a more detailed network analysis of the results).

2.4 Cooperating with Language

In the preceding section 2.3, three different computational models for simulating the evolution of language were presented: the Iterated Learning Model in section 2.3.1; the Language Games approach in section 2.3.2; and the Coevolution Model of Language and Social Structure in section 2.3.3. The first two of these models focus predominantly on lexicon formation, i.e. the process of assigning meanings to arbitrary symbols and spreading such mappings between all agents of a population, with the goal of reliably communicating about future events. Notably, the configurations of these models include explicit meaning-form transmission, making it hard to evaluate how the results obtained from these relate to the real world, where there can be almost no precise feedback available to the agents. The third model by Gong et al. (2004) does well to address the issues of explicit feedback in the former two approaches by introducing an alternative feedback mechanism based on a confidence value of the

[4]See also works by Nettle (1999) and Tamariz, Gong, and Jäger (2011) for further investigations of the effect that reputation might play on the establishment of linguistic conventions in a group of simulated agents.

utilized lexical items. However, this model is focussed on the emergence of syntax, rather than the establishment of the first lexical items (these are given to the agents from the beginning in this model), which is the focal point of this particular project.

Given the above three approaches to modelling the evolution of language, the question that begs itself is if it is possible to simulate the formation of the first lexicon without any form of telepathy, but with a rational feedback mechanism that would at least approximately inform the inventors of the first linguistic items about their success or failure. In this case, an 'approximate' feedback refers to a scenario where a speaker can perceive the success of his utterance(s) as a whole (i.e. by observing the addressee perform an action in response), but at the same time not necessarily be able to tell which parts of the utterance were particularly unclear to the hearer.[5] An answer to the above question is provided by the concept of the signaling game, formulated by Lewis (1969) almost half a century ago with the goal of modelling signal-based cooperation systems observed at all levels of biological organization, from monkeys, birds and bees to simple bacteria (see Skyrms, 2008).

2.4.1 Signaling Game

The signaling game introduced by Lewis (1969) is a general model of communication that can be used not only to represent the evolution of human language, but also the evolution of coordinational signals in basically any living organism. In its original form, the signaling game is 'played' by two agents: the sender and the receiver, both of whom are considered to be a part of the same world and who have similar interests and corresponding goals that they want to achieve. The world inhabited by these two agents is modelled by two states, which are equiprobable and are generated by the world at random. In order to obtain a payoff in either of these world states, the agents are required to perform an appropriate action, with only one such action existing for any given world state. Accordingly, in a world with two states, there are exactly two actions that can be performed by the agents.

Finally, it is defined that only one of the agents can actually observe the state that the agents find themselves in at any time and only the other agent can be in the position to perform the corresponding appropriate action, whereby the agent roles alternate randomly between successive 'rounds' of the signaling game. In order to obtain a payoff from the situation, the informed agent then needs to transmit his knowledge about the current state of the world to the actor agent, which he does by producing one of the two signals that the agents are capable of emitting. If the hearer agent performs the appropriate action, i.e. if he correctly understands the state that the agents find themselves in from observing the speaker's signal alone, both agents obtain a payoff. If the hearer performs the incorrect action, both agents are punished.

In the very simple configuration outlined above, the agents obviously have a limited number of strategies that they can follow when producing and interpreting signals, with exactly two strategies yielding maximally efficient coordination, i.e. guaranteeing success in all future interactions. In particular, maximum coordination is reached if the sender uses a different term in each world state and the receiver chooses the appropriate action for both these terms. A combination of agent production/interpretation strategies that meets these conditions was labelled by Lewis a *signaling system*. Notably, while a signaling system clearly represents a Nash equilibrium,[6] there are also other Nash

[5]The scenario is similar to a psycholinguistic study conducted by Clark and Wilkes-Gibbs (1986), in which participants were required to coordinate their language use with the goal of arranging two sets of picture cards in the same order. Here too, the interlocutors were able to estimate how successful their interactions were, but were not necessarily able to tell what exactly went wrong in case of failure.

[6]A Nash equilibrium represents a state in a game of two or more players whereby no player can increase his or her

equilibria that can emerge in a signaling game that do not achieve the highest possible payoff, i.e. if the sender produces the same signal regardless of the observed world state, the success of any interaction will be down to chance, no matter what the hearer's strategy is. Consequently, the selection of a good learning strategy in a signaling game can radically change its outcome, as will be discussed in the following section.

2.4.2 Reinforcement Learning

The most popular learning strategy applied to the signaling game in evolutionary game theory is Herrnstein reinforcement learning (cf. Roth & Erev, 1995). The underlying principle of this learning strategy is Herrnstein's matching law (see Herrnstein, 1970), which postulates that players will employ any particular strategy in proportion to the amount of payoffs obtained with the help of that strategy in the past. In its most typical form (see Barrett & Zollman, 2009), the learning strategy is initiated with an equal starting weight $q_i(0) = 1$ being associated with every possible strategy i available to the players, whereby $q_i(t)$ essentially represents a player's propensity to utilize strategy i at turn t. When a player is then required to produce a signal for a given state or perform an action in response to a perceived signal, the signal/action is selected from the pool of signalling/action strategies. Hereby, the selection probability $p_i(t)$ of a strategy i in round t is proportional to its weight, as defined by the response rule specified in equation 2.1.

$$p_i(t) = \frac{q_i(t)}{\sum_j q_j(t)} \tag{2.1}$$

After a round of the signaling game is completed, i.e. when the sender has chosen his signal and the receiver has performed what he deemed to be the correct action in response, a payoff is calculated based on the actual appropriateness of the action and indicated to both players. The value of the payoff is then added to the previous weight values of the employed strategies, as specified by the updating rule below:

$$q_i(t+1) = \begin{cases} q_i(t) + \pi(t) & \text{if strategy } i \text{ was utilized} \\ q_i(t) & \text{otherwise} \end{cases} \tag{2.2}$$

The learning process outlined above can be very intuitively represented by a simple urn model. During each round of the signaling game, the sender player is informed of the randomly selected world state and then draws a ball at random from the corresponding urn, whereby each ball in an urn has the same probability of being picked. In the case of the sender, a ball corresponds to a specific signal, which is then sent to the receiver. On hearing the signal, the receiver proceeds in a similar fashion, namely by picking a ball from the urn that corresponds to the given signal. In his case though, the picked ball specifies the action that the player will perform in response to the heard signal. If this action is appropriate for the current state of the world, both players are rewarded, in which case both players return their respective balls to the urns they were picked from, as well as each add an additional ball with the same label to the urn. In case the action performed by the receiver was not the correct one, both agents simply return the balls to the corresponding urns, i.e. no punishment is applied to incorrect strategies.

It has been shown by Skyrms (2008) with simulation that perfect signaling evolves in a population of two agents playing the 2-state/2-term signaling game. More recently, Argiento, Pemantle, Skyrms,

payoff by changing his or her strategy, assuming that the other players keep their strategies unchanged.

and Volkov (2009) have proven that the urn-based learning strategy converges to a signaling system with the probability of one when applied to the same game. However, if the number of signals or terms is even slightly increased, the chances of partial pooling equilibria developing, i.e. agents starting to produce the same signal and interpreting the same meaning regardless of any observed evidence, increase dramatically, from 0.096 for the 3-state/3-term signaling game to 0.594 for the 8-state/8-term game, if executed with the urn-based learning model, as shown by Barrett (2009) with the help of simulations lasting 10^6 turns per run. Notably, a communicative success rate of over 0.75 is reached by over 95% of the simulations of the 8-state/8-term game, meaning that a reasonably reliable language is evolving even in the extended version of the game.

In an attempt to improve on the above results, Barrett (2006) introduced punishments to the updating rule, which reduced the propensity for a particular action if it has not yielded any (or sufficient) payoff, resulting in a largely positive effect on the learnability of more complex signaling systems. However, the increase in the number of states and terms is just one of many possible extensions of the signaling game model, leaving an abundance of issues open for further investigation. In particular, considering that an 8-state/8-term signaling game in which only one state and one signal are utilized in any single play is already exhibiting quite a severe fall-off in the level of communicative success, one could only imagine how complicated was the early human's task of reliably agreeing on arbitrary lexical items in a basically unlimited and ever-changing world.

2.4.3 Lexicon Formation as a Signaling Game

In its original conception by Lewis (1969), the signaling game was meant to demonstrate how a small number of agents can organize their use of communicative signals for the purpose of improving their cooperation in some hypothetical success-based activity. At the core of such games lies the task of correctly transmitting information between the two agents, which in turn allows them to make an informed decision on the appropriate action to execute in the given state of the world. Without the information element, signals would be nothing more than meaningless sounds and would thus have no relation to the real world. The hearer could still perform an action based on an information-free signal, but unless the signal is consistently correlated with some element of the real world, there would be no reliable feedback for the agents to evaluate the correctness of their understanding of the perceived signal, making learning a conventionalized communicative system similar to a human language if not entirely impossible, but certainly even more difficult than it is already proving to be.

On the other hand, it is not sufficient for a signal to be referring to a piece of information from the real world if the information is basically worthless, i.e. if it did not help the communicators achieve anything beneficial with its help. The main reason why such signals would fail to build up a reliable a communication system is that their use would not yield any positive or negative feedback that is necessary for the early adopters to learn their appropriate use. For example, being able to express and understand that *'green'* is the colour of grass does not help one avoid a predator or discover a new source of food. Consequently, the hearer might have as well figured that *'green'* refers to the shape of the clover growing in the grass without being appropriately (or at all) punished for his misunderstanding of the signal. Accordingly, and as pointed out by Bickerton (2003), it seems rather implausible that language would have evolved from something like meaningless grooming sounds, as proposed by Dunbar (1997).

2.5 Conclusions

In summary, the signaling game presented in section 2.4 represents the arguably most important function of language – the transmission of critical information in a cooperative task-oriented setting (cf. Bickerton & Szathmáry, 2011). However, the signaling game itself is a very crude oversimplification of the real world, i.e. it does not allow for any extensions of the world's configuration beyond the number of states, signals and agents involved in the game. While this may not have been critical for Lewis in his more theoretic approach, the limitations become quite critical if one wants to effectively relate the results of such a game to the question of language evolution. Accordingly, this project is focussed on building out the model of the signaling game to an easily extensible and highly configurable workbench – the Language Evolution Workbench (LEW) – that would allow one to perform clearly comparable simulations of lexicon acquisition in a task-oriented environment with the goal of isolating the parameters that are most critical in the very early stages of the lexicon acquisition task. The LEW model is presented in detail in the following section 3.

Chapter 3

Language Evolution Workbench

3.1 Introduction

As an alternative to the approaches presented in section 2.3, Vogel and Woods (2006) have proposed and begun the development of a Language Evolution Workbench (LEW)[1] with the aim of combining the best features of the ILM and the language games while at the same time avoiding the major issues of those models like explicit meaning transmission. This model has since then been extended by Longmore (2008), Vogel (2010) and most recently by Bachwerk and Vogel (2010) within the current research project.[2] In its current state, the model is focussed on the issue of lexicon formation during the transition phase from a closed to an open, learned repertoire of communicative signs. However, owing to its highly modular structure, the LEW can be incrementally extended to address the evolution of syntax, grammar and other elements of language. Aspects of the framework that are relevant for the current research project are spelled out in the sections that follow.

Similarly to the ILM, the LEW is implemented at a relatively high level of abstraction, meaning that some features of the outside world are either modelled in a very simplistic manner or not modelled at all. While such an approach might make the model open to a certain amount of criticism regarding its validity in terms of being an acceptable representation of reality, there are two arguments that should be mentioned in defence of such an approach. First of all, a highly abstracted model of a certain system means that all its elements are independently observable and their effects well quantifiable (e.g. in the form of a mathematical analysis of the Iterated Learning Model (Smith, Kirby, & Brighton, 2003), the naming game (Baronchelli, Loreto, & Steels, 2008) and the signaling games (Argiento et al., 2009)). While a model with hundreds of parameters would certainly bring it closer to reality, one would find it extremely hard to distinguish between significant and insignificant parameters in such a model, as well as to observe the interactions between different parameters.

Furthermore, by starting with a simpler model, one can avoid the mistake of building in features that have not yet been proven or observed well enough in other disciplines to justify or necessitate their inclusion. In other terms, if one does not know the precise parameter settings for the dimensions that impinge on the problem, one should not just build in arbitrary settings as features of the model without experimenting with a range of parameter combinations first. In summary, when addressing the critiques raised in earlier sections, this workbench attempts to do so in a highly constructive and

[1] Available under the Creative Commons Attribution-ShareAlike 3.0 Unported (CC BY-SA 3.0) license at http://github.com/arski/LEW.

[2] Due to the collaborative nature of the project, this section draws on text that has been co-written by Bachwerk and Vogel (2010) and is incorporated with permission of the co-author.

meticulous manner. The remaining sections of this chapter will discuss some of the more fundamental assumptions of the model that have been selected at the beginning of the research project as sufficiently reasonable to have the rest of the model built on top of. Extensions of the model that were proposed within the project and that go beyond the more accepted notions will be then presented in the following chapters.

3.2 Model Structure

3.2.1 General Assumptions

Within the course of this research project, the LEW setup has been extended from 20 to nearly 50 adjustable parameters (see appendix A for the full list) while making as few assumptions about the agents' cognitive skills as possible. The few cognitive skills that are assumed can be considered as widely accepted (according to Jackendoff (1999) and Tomasello (2003) among others) as the minimal prerequisites for the emergence of language. These skills include the ability to observe and individuate events, the ability to engage in a joint attention frame fixed on an occurring event, and the ability to interact by constructing words and utterances from abstract symbols[3] and transmitting these to one's interlocutor.

From an evolutionary point of view, the LEW fits in with the so called faculty of language in the narrow sense as proposed by Hauser et al. (2002) in that the agents are equipped with the sensory, intentional and concept-mapping skills at the start, and the simulations attempt to provide an insight into how these could be combined to produce a communication system with comparable properties to a human language. From a pragmatics point of view, the approach directly adopts the claim that dialogue is *"the most natural and basic form of language use"* (p. 169, (Pickering & Garrod, 2004)). The consequence of adopting this claim is that, as each agent in the LEW individuates events according to its own perspective, which in most cases results in their situation models being initially non-aligned, it becomes their task to align their representations through dialogical interactions, similarly to the account presented in Pickering and Garrod (2004).[4]

3.2.2 Entities and Events

In order to be able to learn a lexicon, agents in any model need to have something to talk about. Two common approaches can be taken when constructing the referent space for agent interactions. The first of these, employed in the Language Games from section 2.3.2 and the Signaling Game presented in section 2.4.1 is to provide agents with a set of objects or basic world states that may have some properties attached to them, but with no combination of a number of such objects. The other approach, preferred by the makers of the ILM from section 2.3.1 as well as by Gong and his colleagues in the model outlined in section 2.3.3, is to generate a number of multi-referent combinations, usually with an underlying structure that can be detected by the agents, with the effect of learning a simple grammar.

[3]While such symbols are referred to as 'phonemes' within the LEW, there is no reason why they should not be representative of gestural signs. In fact, Bickerton suggests that there is no reason to pick between either speech or gesture as the primary medium used for language, as it is most likely the case that in the first instances, both were interpreted by others as meaning something. If true, this would fully justify a modelling approach neglecting this issue.

[4]This claim also implies that language must have evolved in a social setting and not in the head of a particular individual. The only (unlikely) alternative is that this individual was so intelligent that he not only developed a reasonably complex language, but was also able to teach it to others.

As has been outlined before, the focus of the LEW is currently placed entirely on the lexicon acquisition task, i.e. no capability for sentence structure or grammar learning is encoded in the agents within the model. On the other hand, it seems too strong of an oversimplification to provide agents with just one, or even the same few referents in their interactions, as it would arguably make the task of agreeing on a purely conventional lexicon much simpler than it is. Consequently, the LEW operates on a combination of the two approaches, with different multi-referent combinations, or *events*, being presented to the agents during each interaction, who at the same time are not interested in (or capable of) detecting and somehow utilising the underlying structure (see below) of such events.

Events that occur in the LEW are generated by selecting one of the allowable event types and filling it with arguments. An *event type* is identified by the number of actors or objects that are involved in it, ranging from one to three and in effect mimicking the distinction between intransitive (1 argument), transitive (2 arguments) and ditransitive (3 arguments) verbs in modern human languages and the corresponding events that these represent. The three event types are represented below, along with their internal model representations.

- R^1 [action, subject]
- R^2 [action, subject, object]
- R^3 [action, subject, direct object, indirect object]

Whenever an event needs to be generated, its head (i.e. action) is selected at random, after which an *entity* is selected for each argument slot in the event, also at random and with the only restriction that the same entity can only come up once as an argument in any individual event instance. In the original implementation of the model, entities acting as event arguments were represented via three sorts (human, animate or inanimate), i.e. there were no actual entity instances. In the current project, entity sorts have been replaced with a configurable number of entity instances. Despite this change, entities in the LEW are still represented as simple atoms, i.e. they do not currently have any properties or types associated with them

In order to represent the more complex events that can occur in the world, e.g. having an agent see another agent do something to a third agent, the object arguments in R^2 and R^3 have the possibility of being filled with other events, as well as simple entities. In effect, the event construction and argument selection process can proceed recursively with an infinite amount of sub-events being embedded inside the originally selected top-level event, thus resulting in an open-ended meaning space based on a finite number of both entities and event types. Since the dynamics of the model change significantly if a large number of events are recursive, a separation is made between experiments with 'basic' non-recursive events with entity arguments only, and experiments with unlimited recursive event embeddings allowed. In the present work, only basic events will be considered for the sake of clearer analytical accuracy, with recursive events being kept in mind for later research. To provide a more concrete example, consider the following sample event:

Event 1. `[see monkey tiger]`

In this example, `<see>` is the event head corresponding to the R^2 event type with two arguments, a subject and an object. In this particular example, `<monkey>` is the subject of the event and `<tiger>` is the object. Note that the sample event above has been presented with understandable atoms for the sake of providing a clear explanation of the event structure. Within the simulation itself, both events and entities are just randomly generated strings that cannot be easily interpreted, with an example provided below.

Event 2. [`kmsimts lurlbas xutyiis subx`]

When talking about events, <`kmsimts`>, <`lurlbas`>, <`xutyiis`> and <`subx`> are referred to as *event components* in the model. It is important to note straight away that, as event components are meant to represent facts about the simulated world, the meanings that the agents express in their utterances do not necessarily have to (though occasionally do) correspond directly to the individual components. Instead, and similarly to human language, the meanings are constructed by the agents on the basis of the different perspectives that they take on events. The way perspectives are built up and transferred onto lexical meanings in the model will be discussed in more detail in section 3.3.1 below.

3.2.3 Phonetic Inventory

As will be described in more detail below, interactions between agents are performed with the help of utterances that are composed of words, which in turn are constructed out of phonemes. While investigating the emergence and the organization of a phonetic inventory of a communication system falls outside the scope of the current research project, some technical details about words and phonemes in the LEW are nevertheless quite essential to the understanding of some of the nuances of the following sections. In particular, every phoneme in the LEW is represented as a pair of phones (denoted as sequences of letters), mimicking the onset-nucleus structure (without the coda), e.g. "*ii-j*".[5] Furthermore, when an agent needs to invent a new word to express a meaning, he always utilizes one phoneme,[6] which he constructs by generating two new phones that have not been previously used in the population.[7]

3.2.4 Agents and Lexicons

Agents in the LEW are non-physical entities (see Steels & Kaplan, 2002, and further works by Steels and his colleagues for embodied implementations) and are in principle not specialized to the question of language evolution. At the same time, what characterizes every agent in the LEW is its lexicon, with a sample agent lexicon provided below. The lexicon of an agent is represented as a set of meaning-form mappings (also referred to as words or lexical items) w for every agent $a \in \{1,..,N\}$, in a population of N agents. In such mappings, the *meaning* is a list of event components that were encountered by the agent in some event and the *form* is combination of one or more phonemes from some utterance that was either produced or heard by the agent in connection with an experienced event (see section 3.3.1 for a more detailed account of how meanings and forms are constructed and individuated within an interaction). Furthermore, every mapping w_i is associated with a *weight* q_i that is an indicator of the agent's confidence in the mapping and, subsequently, represents the propensity

[5] Note that phones in the LEW are considered completely unique and unrelated, i.e. the `ii` phone is not equivalent to the "*i-i*" phoneme.

[6] Nevertheless, multi-phoneme words are also possible in the system, as will be described in section 3.3.1.

[7] While this might seem as an unnatural restriction, especially as it assumes that the agents have heard (and memorized) all of the phones used in all of the utterances performed within a population, the motivation behind it was to eliminate any additional sources for homonymy in agent lexicons other than the interactions themselves.

of the agent to utilize it in future interactions. Depending on the experimental setup, mapping weights are adjusted according to one of the equations provided in section 3.3.2.[8]

$$Lex^a = \left\{ \begin{array}{l} w_1 = (meaning, form, weight) \\ w_2 = (meaning, form, weight) \\ \qquad \ldots \\ w_k = (meaning, form, weight) \end{array} \right\} \tag{3.1}$$

Notably, there is no underlying knowledge base available to the agents that they could fall back on when associating lexical forms with meanings. While a concept-based world model is undoubtedly a very prominent feature that is missing from the agents in the LEW model, it has been shown that humans are not the only species that are capable of organizing the world in concepts (see Hauser et al., 2002), yet they are the only ones that have a communication system as complex as human language, putting the role of an internal world model into question. Further support for the secondary role of conceptualization is that, in the LEW, linguistic conventions emerge in simple agents even without concepts, as will be seen later on in the document.

3.3 Simulation Design

3.3.1 Interactions

Building on the traditions of computer simulations of language evolution, the LEW is based on interactions between one (thinking) or two (dialogue) agents.[9] The typical interaction in the LEW occurs between two randomly chosen agents, a speaker and a hearer, whereby an agent can also end up talking to himself if he gets picked as the hearer too (language is meant for thinking as well as communicating). More specifically, the interactions follow the scenario outlined below:

Step 1: An event is constructed as described in section 3.2.2 above and fully 'displayed' to both agents. For the sake of simplicity, the sample interaction provided here will be based on the event from example 2.

Step 2: Given that an event is modelled as a list of components, the set of possible perspectives on that event is modelled with the exhaustive set of possible partitionings of that list into sublists. When building up their perspective on an event with n components, the agents have 2^{n-1} different partitions available to them, from which they select one at random.[10] For the sample event 2 presented above, the list of possible partitions, i.e. agent perspectives is provided below:[11]

- `[<kmsimts lurlbas xutyiis subx>]`
- `[<kmsimts lurlbas xutyiis> <subx>]`
- `[<kmsimts> <lurlbas xutyiis subx>]`

[8]Apart from being increased or decreased directly after an interaction, mapping weights can also decay over time if the forgetting parameter is enabled during a simulation. However, forgetting is still being experimented with by the author and was kept disabled throughout all the simulations presented here.

[9]One-to-many (monologue) interactions have also been experimented with by the author, although no conclusive results are ready to be presented on these at the moment of writing.

[10]Note that permutations of event components inside the event are not permitted as their order is meant to represent possible relations between events and entities.

[11]Note the marking of `<event components>` in an `[<event> <partition>]` as opposed to an `[unpartitioned event]`.

- [<kmsimts lurlbas> <xutyiis subx>]
- [<kmsimts lurlbas> <xutyiis> <subx>]*
- [<kmsimts> <lurlbas xutyiis> <subx>]
- [<kmsimts> <lurlbas> <xutyiis subx>]
- [<kmsimts> <lurlbas> <xutyiis> <subx>]

In effect, each perspective amounts to a view of what an agent finds useful to label and communicate about in relation to a witnessed event. Accordingly, when looking at the different partitions listed above, it becomes clear that 'meanings' do not necessarily correspond to individual parts of the world, but are rather representations of the world's parts, via perspectives. In the following, it is assumed that the speaker has selected the partition marked with an asterisk above as his perspective on the experienced event.

Step 3: The speaker assigns a lexical form, i.e. a word, to every meaning from the event partition selected during the previous step, e.g. "*ii-j*", "*i-w*" and "*p-t*"; combines the words into one continuous utterance, i.e. ("*ii-j i-w p-t*");[12] and transmits this to the hearer. The selection of words is achieved either by looking up an existing meaning-form mapping for a given meaning in one's lexicon[13] or, failing to find one, by inventing a new form as described in section 3.2.3 above.

Step 4: The hearer perceives the speaker's utterance as a continuous stream of uninterrupted sound, i.e. ("*ii-j i-w p-t*") and, first of all, needs to segment it into individual words (or decide that he regards the whole thing as a single word). If synchronous transmission is enabled (see Vogel, 2010, for a previous experiment with this condition), the hearer agent is always able to tell the boundaries between speaker's words. In the alternative case, the process of word segmentation and selection becomes very similar to the process of event partitioning, i.e. the hearer picks one of the possible 2^{n-1} segmentations at random (given an utterance consisting of n phonemes). In the sample interaction, the possible word segmentations for the utterance provided above would be as follows:

- ("*ii-j i-w p-t*")
- ("*ii-j i-w*" "*p-t*")
- ("*ii-j*" "*i-w p-t*")
- ("*ii-j*" "*i-w*" "*p-t*")

In this case, assume that the hearer picks ("*ii-j i-w*" "*p-t*") as the word segmentation of the encountered utterance.[14]

Step 5: The hearer assigns a meaning to every word he thinks himself to have encountered in the utterance. This is done in a similar way to how the speaker composes an utterance. First of all, the hearer searches his lexicon for existing meaning-form mappings that include the heard form.

[12] In the following, notation such as ("*ii-j i-w p-t*") will refer to both the full utterance, as well as its one-word segmentation.

[13] If multiple synonymous mappings exist in an agent's lexicon, then a form is randomly picked based on a probability distribution computed from the corresponding mapping weights (see below for more on weights).

[14] As can be seen, when synchronous word segmentation is disabled, hearers have the 'ability' to wrongfully segment a heard utterance and thus introduce larger words into their lexicons and subsequently, when acting as a speaker, into the lexicons of others.

Failing to find one, the hearer randomly assigns some part of the event as the meaning, whereby his perspective on the event, i.e. its partitioning into meanings, is fully independent of the speaker's perspective (cf. step 3). For the sake of completeness, assume that the hearer interpreted the encountered utterance as [`<kmsimts>` `<subx>`].

Step 6: In order for the agents to learn, an update is performed in the agent lexicons (see section 3.2.4) for every meaning-form mapping that they employed in the latest interaction. This update mechanism is described in more detail in section 3.3.2. As has been mentioned in section 3.2.4, agents in the LEW currently only store information about what lexical items they employed in past interactions, but not what events and entities they encountered along the way.

3.3.2 Learning Strategies

In the original version of the LEW, adopted at the beginning of this research project, all simulations implemented a very primitive learning strategy that essentially reinforced all meaning-form mappings utilized to either construct or decode an utterance without any consideration for their success. Based on the notation from Barrett and Zollman (2009), the updating rule for an agent's propensity $q_i(t)$ to utilize a lexical mapping i in an interaction at time t in this learning strategy could be defined for every lexical mapping i employed in the interaction as:[15,16]

$$q_i(t+1) = \begin{cases} q_i(t) + \pi(t) & \text{if mapping } i \text{ was utilized} \\ q_i(t) & \text{otherwise} \end{cases} \tag{3.2}$$

Where $\pi(t)$ represents the payoff of the interaction performed between two agents at time t and was simply fixed as:

$$\pi(t) = 1 \tag{3.3}$$

The above equations do not reflect the success of an employed strategy but rather rely solely on the historical frequency of mapping use. Consequently, there is also no way of introducing punishments for unsuccessful strategies, which have proven to be essential if one wishes to reach any kind of near-optimal solution in the more complex signaling games (see Barrett & Zollman, 2009), which the LEW clearly represents. As an improvement to the learning strategy outlined above, it has been proposed by Bachwerk (2011) to approximate the payoff $\pi(t)$ of an interaction based on its communicative success, which can range from 0 (no understanding) to 1 (full understanding). In order to do so, an additional parameter that defines the *level of minimum success* s_{min} was introduced to the LEW. This parameter allows for a more flexible payoff definition, which incorporates both reinforcement and punishment of a strategy:[17]

$$\pi(t) = \begin{cases} 1 & \text{if } s(t) \geq s_{min} \\ -1 & \text{otherwise} \end{cases} \tag{3.4}$$

While the payoff definition in equation 3.4 is surely an improvement on the completely success-

[15]Since action success in the LEW is approximated over communicative understanding, no response rule needs to be defined here. Instead, much like the Saussureans in Hurford (1989), agents make use of the same lexicon both during signal production and interpretation, meaning that the updating rule applies to both sides of an interaction.

[16]In the LEW, q_i values represent the weights and thus the usage probabilities of competing meaning-form mappings in an agent's lexicon. In general, $q_i(0)$ can be considered to be equal to 0. Note that this has no effect on the invention of new forms.

[17]With negative reinforcements, if the weight of a mapping has a non-positive value after an update has been performed, it is considered to be forever forgotten by the agent and is thus eliminated from his lexicon.

agnostic reinforcement employed previously as described in equation 3.3, it is still fixed at a certain value and is not dependant on the actual distance between the level of minimum success and the actual success rate. This shortcoming was then solved by a employing the difference between actual success and the minimum success threshold as the payoff value:[18]

$$\pi(t) = s(t) - s_{min} \tag{3.5}$$

The way communicative success of an interaction is actually defined in the LEW will be outlined in the following section 3.3.3. Finally, it should be noted at this point that the success-based learning strategies are only plausible if one accepts the premise presented in section 2.4.1, namely that communication evolved out of the need for cooperation, and that successful communication leads to clearly perceivable payoffs during such cooperation, i.e. that communicative success of agents can be employed as an approximation of the obtained payoffs in the feedback loop of the learning process. Note that, in the current implementation of the LEW, this type of implicit feedback is provided to the agents without letting them know precisely what part of the utterance it was that they succeeded in or failed to understand, thus still avoiding any kind of telepathic meaning transmission.[19]

3.3.3 Perceiving Success

While a lot of effort has been put into the avoidance of direct meaning-form transfer (i.e. telepathy) in the LEW, it is still possible to observe the levels of understanding between two interacting agents in the model. Firstly, agents themselves may be able to detect whether they have been successful in a cooperative setting where a task is at stake (as opposed to grooming) even if the utterance has not been strictly speaking understood. Secondly, an independent observer is always in a position to form theories of whether two interlocutors understood each other. When making external measurements, it has to be noted that one need not necessarily look at direct mapping-to-mapping understanding as it can be safely assumed that, in the pessimistic setting of agents individuating events differently, they will share a very limited amount of mappings. In fact, a certain level of understanding can still be attributed to the interlocutors if they have somehow reached the same meaning, even though they expressed it differently. Pragmatically, there can be some level of success if distinct symbol-meaning mappings take agents to at least some of the same meanings.[20]

In other words, there is 100% success in understanding when two agents can be observed to have denoted the same meanings, arranged in a similar order, in what was spoken and what was heard (independently of whether they segmented the sound stream identically), i.e. if they agree on what the different components of an utterance were about. In particular, for communicative agreement to be achieved, the agents need to agree on two things: what entities are being spoken about and what are the relations between these entities. The relations between the entities are considered to be understood if the ordering of the meanings is sufficiently similar between the two agents, as described in more detail in the definitions provided below.[21]

Understanding F1 (F1): The harmonic mean between precision and recall between the hearer's

[18]Having $s_{min} = 0$ in this case would make the updating rule equivalent to its basic version in Herrnstein reinforcement learning which employs just the actual payoff value.

[19]This approach is similar to that of (Skyrms, 2010) in saying that complete understanding is not a requirement for the emergence of conventionalized communicative mechanisms.

[20]If an agent wants salt, asks for salsa, and obtains salt, then some amount of communicative success has happened.

[21]The consideration of meaning arrangements within an utterance also forms a basis for the introduction of more advanced pattern recognition skills that could potentially lead to the evolution of the first syntactic structures.

interpretation of the speaker's utterance.[22]

$$F1 = 2 * \frac{precision * recall}{precision + recall} \tag{3.6}$$

Precision: ratio of meanings that were agreed on by speaker and hearer to the total number of meanings that were extracted by the hearer from an encountered utterance.

Recall: ratio of meanings that were agreed on by the speaker and hearer to the total number of meanings that were referenced by the speaker with the encountered utterance.

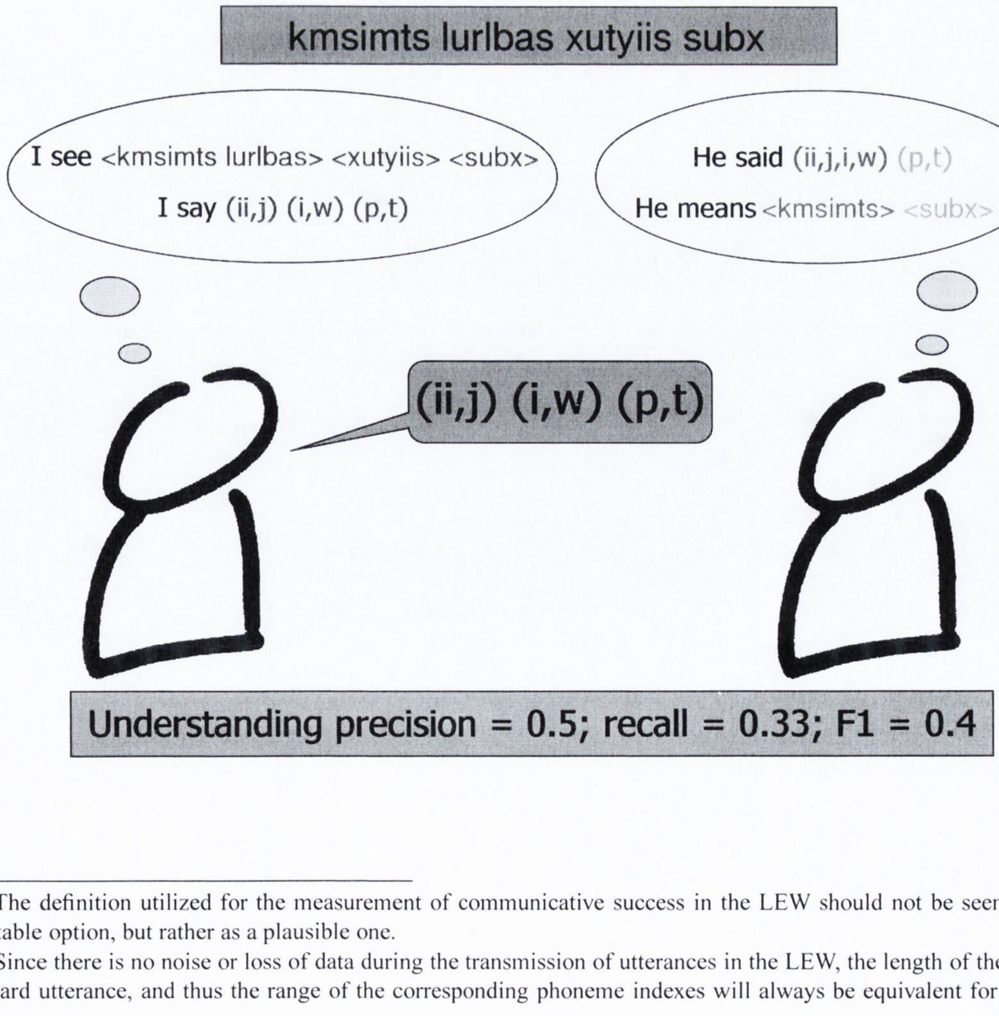
Agreement on meanings: achieved if two meanings match in positionally overlapping forms in the speaker's intended and the hearer's interpreted utterance, whereby the same form cannot be a part of more than one match.

Positionally overlapping forms: two forms from two utterances that share at least one phoneme index.

Phoneme index: the position of a phoneme inside an utterance, with the first phoneme of the first form of an utterance defined to have index = 1.[23]

Finally, figure 3.1, illustrates a sample interaction within the LEW, along with the success measures calculated from the agents' understanding of each other.

Figure 3.1: Sample interaction in the LEW



[22]The definition utilized for the measurement of communicative success in the LEW should not be seen as the only acceptable option, but rather as a plausible one.

[23]Since there is no noise or loss of data during the transmission of utterances in the LEW, the length of the spoken and the heard utterance, and thus the range of the corresponding phoneme indexes will always be equivalent for speaker and hearer.

3.4 Evaluation Measures

Apart from measuring the actual communicative success, the lexicons of the agents are also evaluated with regard to some natural language properties, as mentioned earlier in section 2.5. For instance, lexicon size indicates the range of expressible meanings and interpretable forms; the amount of synonymy and homonymy both inside the individual lexicons and across of the whole population tells us how similar the emergent languages are to natural languages, which seem to tolerate homonymy and avoid synonymy; while the amount of mappings shared by the whole population and the number of agents sharing a mapping on average are both good indicators of potential communicative success. As can be seen, these evaluation measures relate to properties of the whole system, i.e. they allow one to evaluate if one configuration leads to something more like natural language than another, which, in a way, is perhaps even more important than communicative success when attempting to simulate the emergence of human language as we know it. A complete list of the evaluation metrics, along with their exact definitions is provided below:

Lexicon Precision (LPrec): ratio of the hearer's interpreted meanings that were in agreement with the speaker's intended ones (see preceding section), when the corresponding form was present in the hearer's lexicon (see next item).

Lexicon Use (LUse): ratio of forms from the segmented utterance $U = (form_1, form_2, .., form_k)$ that a hearer was able to find in his lexicon during the i^{th} interaction.

$$LUse_i = \frac{|\{k : f_k \in U | \exists m(m, f_k) \in Lex^a\}|}{|U|} \tag{3.7}$$

Mapping Share (MS): number of agents that have a meaning-form mapping in their lexicon, averaged over all distinct mappings from the merged lexicon of the population Lex^{pop}.

$$Lex^{pop} = \bigcup_{a=1}^{n} Lex^a \tag{3.8}$$

$$MS = \frac{\sum_{i=1}^{|Lex^{pop}|} |\{a : w_i \in Lex^a\}|}{|Lex^{pop}|} \tag{3.9}$$

Agent Synonymy (Syn^a): ratio of meanings that are associated with multiple forms M_{syn} to all unique meanings M_{unique} in an agent's lexicon.

$$M^a_{unique} = \{m : (m, f) \in Lex^a\} \tag{3.10}$$

$$M^a_{syn} = \{m : (m, f) \in Lex^a | \exists f'(m, f') \in Lex^a, f \neq f'\} \tag{3.11}$$

$$Syn^a = \frac{|M^a_{syn}|}{|M^a_{unique}|} \tag{3.12}$$

Agent Homonymy (Hom^a): ratio of forms that are associated with multiple meanings F_{hom} to all unique forms F_{unique} in an agent's lexicon.

$$F^a_{unique} = \{f : (m, f) \in Lex^a\} \tag{3.13}$$

$$F^{a}_{hom} = \{f : (m,f) \in Lex^{a} | \exists m'(m',f) \in Lex^{a}, m \neq m'\} \tag{3.14}$$

$$Hom^{a} = \frac{|F^{a}_{hom}|}{|F^{a}_{unique}|} \tag{3.15}$$

Population Synonymy ***(Syn^{pop}):*** ratio of meanings that are associated with multiple forms M^{pop}_{syn} to all unique meanings M^{pop}_{unique} in the merged lexicon of the population.

$$M^{pop}_{unique} = \{m : (m,f) \in Lex^{pop}\} \tag{3.16}$$

$$M^{pop}_{syn} = \{m : (m,f) \in Lex^{pop} | \exists f'(m,f') \in Lex^{pop}, f \neq f'\} \tag{3.17}$$

$$Syn^{pop} = \frac{|M^{pop}_{syn}|}{|M^{pop}_{unique}|} \tag{3.18}$$

Population Homonymy ***(Hom^{pop}):*** ratio of forms that are associated with multiple meanings F^{pop}_{hom} to all unique forms F^{pop}_{unique} in the merged lexicon of the population.

$$F^{pop}_{unique} = \{f : (m,f) \in Lex^{pop}\} \tag{3.19}$$

$$F^{pop}_{hom} = \{f : (m,f) \in Lex^{pop} | \exists m'(m',f) \in Lex^{pop}, m \neq m'\} \tag{3.20}$$

$$Hom^{pop} = \frac{|F^{pop}_{hom}|}{|F^{pop}_{unique}|} \tag{3.21}$$

Lexicon Size (LSize): number of meaning-form mappings in an agent's lexicon.

$$LSize^{a} = |Lex^{a}| \tag{3.22}$$

Lexicon Meanings (LM): number of distinct meanings in an agent's lexicon.

$$LM^{a} = |M^{a}_{unique}| \tag{3.23}$$

Lexicon Forms (LF): number of distinct forms in an agent's lexicon.

$$LF^{a} = |F^{a}_{unique}| \tag{3.24}$$

It has been stated before in section 1.3 that the success of any particular parameter configuration of the model will be evaluated and decided on based on the comparability of obtained lexical properties to those observed in real human languages. While doing so would be clearly desirable, it has to be admitted that obtaining real language measures for some of the above properties, especially such as lexicon use and lexicon precision, is completely unrealistic. Furthermore, while dictionaries and lexical databases allow one to compute the number of different forms and meanings in a language, these measures can hardly be reasonably compared to the lexicons of the LEW agents that only have a short period of time to develop, in which only a limited set of meanings can be observed. However, two measures that can be extracted from dictionaries such as WordNet[24] and meaningfully compared to the corresponding values from the simulation results, are lexicon synonymy and homonymy.

Self-published statistics of the latest version of the WordNet database[25] suggest that there are 26896 polysemous words[26] in the English lexicon totalling 155287 distinct words, indicating a homonymy level of 0.17. Furthermore, the statistics indicate that 79450 out of the 206941 distinct senses (or meanings) represented by the English language are polysemous, implying a synonymy level of 0.38. Interestingly, these calculations appear to refute the claims made by Carstairs-McCarthy (1999) among others that synonymy is less tolerated in human languages than homonymy.[27] More importantly for the purpose of this research project, however, these values provide a reference point for comparisons of the lexicons from the presented model to those of real humans.

Finally, it should be noted that from the evaluation measures presented above, understanding precision, lexicon use and lexicon precision are computed after every interaction and, when analysed statistically or presented in a visual format, are averaged over all interactions of a simulation. Furthermore, as already mentioned above, mapping share, as well as population synonymy and homonymy are computed based on the merged lexicon of all agents and are thus by definition global measures that characterize the whole population. All other measures are calculated on a per-agent basis and then averaged over all agents in the population.

3.5 Experimental Methodology

Experiments conducted with the help of the LEW all follow the same methodological approach. First of all, a set of parameters of the model that one wishes to investigate is selected (or implemented, if missing), along with several different candidate values for these parameters. Next, for each resulting parameter combination, the workbench is configured accordingly and a batch of simulations is performed with the same configuration, but with different randomization seeds.[28] In particular, every simulation run consists of a predefined number of interactions, with the average number of interactions per agent usually kept constant across experiments (i.e. if a simulation with two agents consisted of 1000 interactions, a comparable simulation with ten agents would involve 5000).[29] The number of interactions within a simulation can vary between experiments, but usually lies at around 1000 per agent, following the reasoning that in order for a complex evolutionary trait such as language to establish itself in a population, it should provide at least some value from the very early stages, otherwise it would have been quickly abandoned by its pioneers.

Throughout the conducted simulation runs, a variety of tasks are performed at regular intervals. These tasks can include forgetting, agent elimination and addition (if these are enabled), as well as intermediate evaluations of the system that enable one to observe the evolution of different properties over time. Additionally, a full evaluation of the system is performed at the end of each simulation run, at which point the final values of the measures presented in section 3.4 above are computed. Having

[24] WordNet is an online lexical database of the English language, available at http://wordnet.princeton.edu.

[25] Retrieved from http://wordnet.princeton.edu/wordnet/man/wnstats.7WN.html.

[26] In anticipation of possible confusion, it has to be noted that 'words' refer to strings in WordNet, i.e. correspond to 'forms' in the LEW.

[27] Admittedly, the definition of polysemous senses is usually rather loose in dictionaries and does not correspond to the more strict version of the term employed in the field of linguistics.

[28] As a rule, 600 simulation runs are performed for each configuration of the model in order to guarantee statistical significance of results.

[29] Since interactions will normally occur between two agents, executing 5000 interactions in a group of ten agents would result in every agent being involved in an average of 1000 interactions.

conducted all the necessary simulations, statistical analysis is used for drawing conclusions about the experiment, with graphical representations aiding in visualising the different trends that emerge in the system. During the analysis, the results are evaluated with respect to similarity of certain observed properties to their corresponding values in human languages, some of which are outlined in the preceding section 3.4. In addition to that, outcomes of experiments conducted with different parameter configurations are compared for the purpose of deciding which of these appears to move in the right direction in terms of the different evaluation measures. More often than not, the results of an experiment will indicate possible interactions between some of the varied parameters of the system, as well as suggest further variations that might prove significant for the purpose of understanding the emergence of a communication system in the simulated population of agents. These indications are then utilized when designing and conducting future experiments with the system.

3.6 Summary

The LEW is a hybrid model of language evolution that combines aspects of the ILM, Language Games, the Coevolution Model of Language and Social Structure as well as the Signaling Game. Similarly to the above, the LEW focuses on interactions between pairs of agents who attempt to communicate about a set of topics from their environment. However, by assuming that such interactions occur in a task-oriented setting, agents in the LEW can plausibly identify the levels of their communicative success by observing the payoffs obtained from a subsequent action, thus avoiding the need for explicit meaning-form transmission. In terms of partner selection, the LEW does not model the age of agents and is hence not entirely compatible with the notion of horizontal transmission from the ILM. It does, however, allow one to configure almost any social structure within the population, which can also be dynamically adjusted by the agents themselves throughout the simulations, e.g. based on mutual understanding.

Finally, the focus of the LEW is put on the establishment of meaning-form mappings, i.e. a lexicon. While most other computational models of language evolution go beyond that and attempt to look into the emergence of syntax from the very beginning, the reasoning behind the choice of ignoring syntax at first is that it would only have been a necessity once the holistic lexicon of language inventors had reached a certain cognitive threshold, where the memorization and the processing of further lexical items, without any regularization, would have become intractable. Considering the small size of early lexicons, it would seem that such a state could not have been reached until later on in human language development.

Sticking with the popular ‘game’ metaphor employed for most models of language evolution, the simulations of the LEW are referred to as the Lexicon Acquisition Game. Keeping in mind that the signaling game has been shown to exhibit a very significant drop in success levels even for the slightest extensions of the basic 2-agent/2-state/2-signal variation, section 4.2 will begin with the simplest possible configuration of the LEW, which includes just two agents and one world state (comprising of 2-4 elements). Having performed an analysis of this very basic configuration, section 4.3 will look into stepwise extensions of the model and the corresponding effects on the different evaluation measures, which are also outlined in the following section 3. As was the case with the signaling game, it is expected that providing a formal analysis of the Lexicon Acquisition Game will become increasingly infeasible as the game is extended, in which case the results of the simulations will provide the primary fallback for model evaluation.

Chapter 4

Lexicon Acquisition Game

4.1 Introduction

In order to provide a fundamental analysis of the proposed computational framework, it is first necessary to examine the limits of lexical properties that are established based on a number of simple and intuitive constraints in a LEW-simulated population of agents. In particular, this chapter will look at interactions performed between two different agents (i.e. without self-talk) about the three types of simple, non-recursive events, with only one event available in any given simulation, in effect modelling the ability to only express simple sentences without embedded clauses. The motivation for omitting recursive events from the fundamental analysis lies in the natural complexity increase inherent to recursive and thus limitless systems. While the option of having an open-ended meaning space is still considered to be quite critical for any reliable model of language evolution and is being experimented with by the author, this chapter will look to analyse a fairly restricted version of the system to lay down a proper fundamental basis for future, more complex experiments that will be presented in chapter 5.

Within the individual sections of the chapter, an analytical analysis will be conducted for the most significant properties of a communication system, as outlined in section 3.4 before. The goal of this exercise is to determine what levels of the corresponding properties can be achieved under different constraints, as well as what are the probability distributions of the different property levels. The main constraint at the focus of the chapter is the level of minimum success, i.e. the level of lexical agreement required for a payoff to be obtained by the communicating agents. As described in section 3.3.2, the payoff of an utterance basically determines if the words utilised in it will be learned, i.e. retained for future use by the agents.

In the first section of this chapter, a very basic version of the Lexicon Acquisition Game will be discussed, which includes only two agents and one event consisting of between two and four elements. The goal of reducing the model to the simplified configuration is to establish some analytical baselines for the model, which will be useful when moving forward and examining more complex configurations of the LEW. In effect, the approach is similar to analysing the simplest possible 2-agent/2-state/2-signal version of the signaling game before looking into the possibility of extending either one or more of the limits.

The next two sections will look into an extended version of the model, with a particular focus on population size. Firstly, in section 4.3, the number of agents in the simulated population will be increased by just one agent, from two to three. While this could appear to be a very marginal extension of the model size, it actually has the potential of completely altering the dynamics of lexicon acquisition. The main investigation point here is thus if the different limits imposed on the system that were discussed in section 4.2 will simply increase/decrease as a function of the population size, or, instead, if the dynamics of the system will change altogether, owing to the presence of what could be referred to as intermittently non-participatory agents in a system where only two-agent interactions are possible.

Finally, section 4.4 will look into significantly larger populations consisting of ten agents. This number brings the model closer to the argued evolutionary state of hominid groups at the time where language first emerged. As can be imagined, in groups like this, agents are likely to find it even more difficult to agree on lexical conventions, with the possibility of full subgroup dialects emerging throughout the population. In section 4.4, the focus will remain on the general lexical properties of the individual agents as well as the overall lexicon of the population. The following chapter 5 will then look more closely into the different complex social configurations that can emerge, or that are preconfigured in the system, as well as discuss their effects on the emergent lexicon of the population. For the sake of completeness, the specific parameter settings of the LEW that were used in all of the presented experiments are presented in appendix B.

4.2 Interacting in Pairs

Since the dynamics of lexical acquisition can be very different depending on the number of actors involved in the process, the rest of the chapter will be split into three sections, focussing on populations of two, three and ten agents, respectively. From an evolutionary point of view, it would appear that a scenario involving a larger number of interlocutors seems to be more favourable, considering that most accounts of language evolution place the origins of language around the time as our human predecessors were forced to increase their group size based on a number of environmental factors. On the other hand, it does not seem implausible to assume that, even in a large group, some early form of communication might have emerged in pairs of cooperating members.

4.2.1 Lexicon

One of the most important features of a lexicon is its overall size, along with the number of unique expressible meanings and interpretable forms. In order to learn new lexical items, agents in the LEW need to experience a certain level of success first, as without that it is assumed that no payoff has been gained from the interaction and thus no motivation exists to remember the lexical mappings employed therein. Experiencing success in this case equals to reaching some agreement over what is being talked about. Depending on the minimum level of understanding necessary for an interaction to be considered a success, it may be sufficient for the agents to agree on only a subset of the meanings that were referred to.

Accordingly, if the minimum success threshold in the agents is relatively low, they may end up agreeing on just a few meanings, however, not being able to tell which meanings they interpreted correctly and which not, they would still memorize all of the lexical mappings involved in an utterance. This phenomenon results in a certain number of lexical mappings being shared across the population, with others being consistently misused, without the ability to detect which are which. A side effect of this phenomenon is that the maximum limit on the overall lexicon size of the population will grow not only proportionally to the number of expressible meanings encountered, i.e. the number and size of the events, but also to the level of minimum success.

Lexical Mappings

If M is the set of all potential meanings in a simulation, then the optimal lexicon of the population Lex^{pop} in terms of communicative potential would be equal in size to the number of elements in M, i.e. $|Lex^{pop}| = |M|$, with no meaning being represented twice in the lexicon and the lexicons of all agents being equivalent. Obviously, the agents could be so unlucky as to never guess the correct meaning of any form, thus never achieving the level of minimum success necessary to obtain a payoff and to learn a lexical mapping, resulting in an empty lexicon $Lex^{pop} = \emptyset$ with $|Lex^{pop}| = 0$. However, there is also the possibility that, in any given interaction, agents only agree on as many meanings as they need in order to surpass the level of minimum success s_{min} required for obtaining a payoff. In this case, the lexicon of every agent may end up having more than one mapping associated with some meanings. Consequently, the overall lexicon of the population can actually be noticeably larger than would be required in a given world.

To illustrate the above on a very simple example, assume that two agents inhabit a world where only one event can occur: $[\texttt{event}_1\ \texttt{entity}_1]$. Also assume that the level of minimum success employed by the agents is anything smaller than 0.5, e.g. $s_{min} = 0.4$. In this case, if the speaker produces an utterance ("*oo-u*"), the meaning of which is $<\texttt{event}_1\ \texttt{entity}_1>$, the hearer clearly needs to understand the sole meaning in order to experience any success. The interaction and the corresponding resulting lexicon are presented below:

Interaction 1.

- Event: $[\texttt{event}_1\ \texttt{entity}_1]$
- Speaker partition ($Agent_1$): $[<\texttt{event}_1\ \texttt{entity}_1>]$
- Utterance: *("oo-u")*
- Hearer interpretation ($Agent_2$): $[<\texttt{event}_1\ \texttt{entity}_1>]$
- Outcome: *success* $s = 1.0$ *and payoff* $\pi = 0.6$

Lexicon 1.

Meaning \ Agent	$Agent_1$	$Agent_2$
$<\texttt{event}_1\ \texttt{entity}_1>$	"*oo-u*" ($w = 0.6$)	"*oo-u*" ($w = 0.6$)
$<\texttt{event}_1>$	$\emptyset$	$\emptyset$
$<\texttt{entity}_1>$	$\emptyset$	$\emptyset$

Table 4.1: Agent lexicons after interaction 1 ($s_{min} = 0.4$).

However, if the speaker's intended meanings are partitioned into $[<\texttt{event}_1>\ <\texttt{entity}_1>]$, then the hearer can interpret the corresponding utterance ("*i-k*" "*f-e*") as $[<\texttt{event}_1>\ <\texttt{event}_1>]$. In this case, the level of success would be exactly 0.5, meaning that a payoff would be obtained and the agents would memorize the mappings that led them to their corresponding interpretations. The speaker agent would learn {<event$_1$> : *"i-k"*} and {<entity$_1$> : *"f-e"*}, while the hearer agent would learn {<event$_1$> : *"i-k"*} and {<event$_1$> : *"f-e"*}, as exhibited below.

Interaction 2.

- Event: [event$_1$ entity$_1$]
- Speaker partition (*Agent*$_1$): [<event$_1$> <entity$_1$>]
- Utterance: *("i-k" "f-e")*
- Hearer interpretation (*Agent*$_2$): [<event$_1$> <event$_1$>]
- Outcome: $s = 0.5$ *and* $\pi = 0.1$

Lexicon 2.

Meaning \ Agent	*Agent*$_1$	*Agent*$_2$
<event$_1$ entity$_1$>	*"oo-u"* ($w = 0.6$)	*"oo-u"* ($w = 0.6$)
<event$_1$>	*"i-k"* ($w = 0.1$)	*"i-k"* ($w = 0.1$) *"f-e"* ($w = 0.1$)
<entity$_1$>	*"f-e"* ($w = 0.1$)	$\emptyset$

Table 4.2: Agent lexicons after interactions 1 - 2 ($s_{min} = 0.4$).

Up until here, the overall lexicon of the population is still $|Lex^{pop}| = |M|$, but remember that the hearer agent from the first interaction (*Agent*$_2$) still does not have a mapping in his lexicon that would have a form corresponding to the <entity$_1$> meaning. Instead, this agent has two mappings associated with the <event$_1$> meaning. Accordingly, at a later stage, when the speaker-hearer roles of the two agents have switched, *Agent*$_2$ might want to express the meanings [<event$_1$> <entity$_1$>] himself. In this case, he would need to invent a new form for <entity$_1$>, so he might say something like (*"i-k"* *"mm-u"*). *Agent*$_1$ already has a good idea about the meaning of *"i-k"*, which is <event$_1$>, yet he does not know what *"mm-u"* means, so he might guess that the other agent was referring to <event$_1$ entity$_1$>, i.e. interpret the speaker's utterance as [<event$_1$> <event$_1$ entity$_1$>] (remember that there are no restrictions on the agents' guessing integrated at this stage). In the above example, and as demonstrated below, the success of the interaction would be equal to 0.5 and hence a payoff would be experienced by both agents for both the correctly conveyed meaning, and the misinterpreted one.

Interaction 3.

- Event: [event$_1$ entity$_1$]
- Speaker partition (*Agent*$_2$): [<event$_1$> <entity$_1$>]
- Utterance: *("i-k" "mm-u")*
- Hearer interpretation (*Agent*$_1$): [<event$_1$> <event$_1$ entity$_1$>]
- Outcome: $s = 0.5$ *and* $\pi = 0.1$

Meaning \ Agent	$Agent_1$	$Agent_2$
<event₁ entity₁>	"*oo-u*" ($w = 0.6$) "*mm-u*" ($w = 0.1$)	"*oo-u*" ($w = 0.6$)
<event₁>	"*i-k*" ($w = 0.2$)	"*i-k*" ($w = 0.2$) "*f-e*" ($w = 0.1$)
<entity₁>	"*f-e*" ($w = 0.1$)	"*mm-u*" ($w = 0.1$)

Table 4.3: Agent lexicons after interactions 1 - 3 ($s_{min} = 0.4$).

Lexicon 3.

After only three interactions, agents can thus end up with a lexicon of four mappings each, even though there are only three different meanings available in their world. On top of that, their interactions will still have a good chance of consistently performing above the minimum success threshold, meaning that the lexical mappings that are mismatched between the agents might never become rectified, resulting in consistent, yet unperceived, misunderstandings between the agents. Notably, similar scenarios are not uncommon in humans utilizing modern language. It would seem that linguistic items are clearly defined, yet one frequently experiences situations where two people seem to agree on something during conversation, but when performing a corresponding activity end up doing something quite different.

From a quantitative point of view, lexicon 3 represents the largest possible overall lexicon in a population of two agents conversing about a single two-component event and the level of minimum success set to anything below 0.5. The lexicon cannot grow any further because both agents have a form associated with every possible meaning in the world, thus eliminating the possibility of further invention. In general, the lexicon of an agent can grow if there are still meanings from the world that are not represented in it by any form. Furthermore, as described above, new forms can only be added to a lexicon if, in an interaction with another agent, a sufficient amount of agreement has been reached about the meanings involved. Accordingly, the agents need to share a certain amount of mappings in order to be able to learn further mappings. However, the lexicon of any particular agent, or the overall population, will not grow beyond $|M|$ if all of the mappings are shared.

Consequently, the limit for the largest possible lexicon in a population of agents is reached when a lexical mapping can be learned for every possible meaning in the agents' world, yet only a minimally necessary number of mappings is actually shared between the agents. The lexicon acquisition problem then becomes a matter of finding a minimum set of shared mappings that is sufficient for the acquisition of a mapping for every possible meaning in the agents' world. Since a form for a meaning can only be learned by both agents by either agreeing on the form reference directly, or by having the new form come up in an utterance together with another, shared mapping, the task of finding a minimal shared set can be represented as a graph.

Dominating Set Problem

Following the above discussion, it appears that in order to determine the maximum number of lexical mappings that can be learned by agents communicating in an isolated pair, a small detour into graph theory is necessary. Within graph theory, the problem that is most similar to determining the necessary subset of meanings with a shared mapping from the overall set of meanings is referred to as the *dominating set problem*. In graph theory terms, the task is then to find a set of nodes D for a graph $G = (V, E)$, that is a subset of V and for which it holds that every node from graph G is either in D or is connected to a member in D by some edge $e \in E$. Given this definition, a dominating set could range from having all of the nodes from the graph in it, to having just a very small subset of heavily connected nodes. Finally, a *minimum dominating set* of a graph is a dominating set with the smallest possible number of nodes in it, given graph G.[1]

In the lexical acquisition task encountered by the agents in the LEW, the set of vertices V corresponds to the set of individual meanings that can come up in any given world. As has been described in section 3.2.2, all meanings in the LEW are a part of some event, with only the meaning consisting of all components of an event appearing on its own. In order to represent this relation, an edge is added to E in the meaning graph for every pair of meanings that come up in the same partition (without having to be adjacent in it) of one of the world's events. From this graph, one can then infer the (minimum) dominating set of meanings for which the agents need to share a lexical mapping in order to be able to learn a mapping for every possible meaning in their world. Of particular interest for evaluating the process of lexicon acquisition, however, are not the different possible minimum dominating sets themselves, but rather the actual size of the minimum dominating set, also referred to as the *minimum domination number*, with respect to the size of the graph. Since the latter is defined by the number of potential meanings in an agent's world, i.e. as the number of possible partitions of an event, the size of the graph can be defined as a function of the number of events and the sizes of the corresponding events, based on equation 4.5.

$$|V| = |M| = \sum_{i=1}^{|Events|} |R^i_{meanings}| = \sum_{i=1}^{|Events|} \frac{|R^i| * (|R^i| + 1)}{2} \tag{4.1}$$

The determination of the minimum dominating set for a graph G as well as the number of vertices in such a set, i.e. the domination number, is a classical NP-complete decision problem that has been studied extensively since the middle of the 20^{th} century, as described by Hedetniemi and Laskar (1990). A minimum dominating set of a graph G with $n = |V|$ vertices can be found in time $\mathcal{O}(2^n n)$ by inspecting all vertex subsets of the graph. However, there also exist algorithms that can determine the minimum dominating set more efficiently, the fastest solution to date being in time $\mathcal{O}(1.7159^n)$ by Fomin, Grandoni, Pyatkin, and Stepanov (2008). For the sake of the current project, the crude inspection approach should suffice for determining the minimum dominating number, i.e. the minimum amount of meanings with a shared mapping in a world with a single event of one of the three types described in section 3.2.2 and a certain minimum success threshold s_{min}. The results of such inspections are provided in table 4.4 below.

Event type \ s_{min}	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
R^1 (3 meanings)	2	2	2	2	2	3	3	3	3	3
R^2 (6 meanings)	3	3	3	3	3	5	5	6	6	6
R^3 (10 meanings)	3	3	3	3	5	7	7	9	10	10

Table 4.4: Minimum domination number for the different event type and s_{min} combinations.

It has been stated above that for the lexicons of the agents to grow to their biggest possible size, the agents would be required to share the forms associated with the meanings from the minimum

[1] As implied, a graph can have more than one distinct (minimum) dominating set.

dominating set. However, this statement does not imply that the shared forms have to be the first forms learned for these meanings, nor that the agents cannot have non-shared forms associated with these meanings alongside shared ones throughout the simulations. The reason why this is important is that for certain combinations of the selected event size and the level of minimum success, meanings belonging to a minimum dominating set can come up in utterances with other members of the corresponding set, where correctly interpreting the former is not a requirement for a payoff to be obtained.

As an example, let us consider an R^2 event [event$_2$ entity$_3$ entity$_4$] in interactions with the level of minimum success set at $s_{min} = 0.25$. As can be inferred from table 4.4 above, the minimum domination number for this combination is three, with one of the minimum dominating sets (there are, in fact, three distinct ones in this case) being $D = \{$<event$_2$>,<entity$_4$>,<event$_2$ entity$_3$ entity$_4$>$\}$. It can be quickly spotted though, that given the relatively low minimum success threshold, an utterance about [<event$_2$> <entity$_3$> <entity$_4$>] does not require for both <event$_2$> and <entity$_4$> to be agreed on by the interlocutors for a payoff to be obtained – just one would suffice, resulting in $s = 0.33$ and $\pi = +0.08$. Accordingly, it is possible for the agents to fail in communicating about e.g. <entity$_4$> in a first interaction and then succeed in a later one, when the speaker-hearer roles are reversed. To make the above clear, the two interactions are presented below:

Interaction 4.

- Event: [event$_2$ entity$_3$ entity$_4$]
- Speaker partition ($Agent_1$): [<event$_2$> <entity$_3$> <entity$_4$>]
- Utterance: *("ff-d" "v-e" "aa-i")*
- Hearer interpretation ($Agent_2$): [<event$_2$> <event$_2$ entity$_3$> <event$_2$ entity$_3$ entity$_4$>]
- Outcome: $s = 0.33$ *and* $\pi = 0.08$

Lexicon 4.

Meaning \ Agent	$Agent_1$	$Agent_2$
<event$_2$ entity$_3$ entity$_4$>	$\emptyset$	"*aa-i*" ($w = 0.08$)
<event$_2$ entity$_3$>	$\emptyset$	"*v-e*" ($w = 0.08$)
<entity$_3$ entity$_4$>	$\emptyset$	$\emptyset$
<event$_2$>	"*ff-d*" ($w = 0.08$)	"*ff-d*" ($w = 0.08$)
<entity$_3$>	"*v-e*" ($w = 0.08$)	$\emptyset$
<entity$_4$>	"*aa-i*" ($w = 0.08$)	$\emptyset$

Table 4.5: Agent lexicons after interaction 4 ($s_{min} = 0.25$).

Having taken part in the above interaction 4, agents now have a shared form referring to the <event$_2$> meaning, which is one of the items in the minimum dominating set suggested above. Furthermore, $Agent_1$ now also has a form associated with another meaning from this set – <entity$_4$>, even though *Agent*$_2$ has misinterpreted the form used by *Agent*$_1$ to refer to it. The following interaction 5 will show that this does not prevent the agents from learning a shared form for the <entity$_4$> meaning when the speaker-hearer roles are reversed though.

Interaction 5.

- Event: [event$_2$ entity$_3$ entity$_4$]
- Speaker partition (*Agent*$_2$): [<event$_2$> <entity$_3$> <entity$_4$>]
- Utterance: *("ff-d" "hh-kk" "n-oo")*
- Hearer interpretation (*Agent*$_1$): [<event$_2$> <entity$_3$ entity$_4$> <entity$_4$>]
- Outcome: $s = 0.67$ *and* $\pi = 0.42$

Lexicon 5.

Meaning \ Agent	*Agent*$_1$	*Agent*$_2$
<event$_2$ entity$_3$ entity$_4$>	$\emptyset$	"*aa-i*" ($w = 0.08$)
<event$_2$ entity$_3$>	$\emptyset$	"*v-e*" ($w = 0.08$)
<entity$_3$ entity$_4$>	"*hh-kk*" ($w = 0.42$)	$\emptyset$
<event$_2$>	"*ff-d*" ($w = 0.5$)	"*ff-d*" ($w = 0.5$)
<entity$_3$>	"*v-e*" ($w = 0.08$)	"*hh-kk*" ($w = 0.42$)
<entity$_4$>	"*n-oo*" ($w = 0.42$) "*aa-i*" ($w = 0.08$)	"*n-oo*" ($w = 0.42$)

Table 4.6: Agent lexicons after interactions 4 - 5 ($s_{min} = 0.25$).

As can be seen above, after the two interactions, agents now have a shared form for two of the three meanings from the selected minimum dominating set, making it possible for them to end up learning the maximum possible number of different lexical mappings between the two, such as lexicon 6 for example (provided without mapping weights for the sake of simplicity and with shared mappings sorted on top).

Lexicon 6.

Meaning \ Agent	$Agent_1$	$Agent_2$
`<event`$_2$ `entity`$_3$ `entity`$_4$`>`	*"ee-r"* *"ee-u"*	*"ee-r"* *"aa-i"*
`<event`$_2$ `entity`$_3$`>`	*"i-ff"*	*"qq-ww"*
`<entity`$_3$ `entity`$_4$`>`	*"t-u"*	*"ll-bb"*
`<event`$_2$`>`	*"ff-d"*	*"ff-d"* *"bb-o"*
`<entity`$_3$`>`	*"bb-o"* *"ll-bb"*	*"ee-u"* *"i-ff"* *"t-u"*
`<entity`$_4$`>`	*"n-oo"* *"aa-i"* *"qq-ww"*	*"n-oo"*

Table 4.7: Agent lexicons after interactions 4 - 5 ($s_{min} = 0.25$).

Notably, during the generation of the above lexicon, the first form that was invented to refer to the <entity$_4$> meaning was actually misinterpreted by the corresponding hearer, resulting in an additional non-shared mapping being generated from a meaning that eventually ends up being a part of the shared set. What this means for the estimation of the biggest possible lexicon that can be acquired by the agents in a population of two is that M_{shared} does not necessarily have to include all the meanings from the minimum dominating set. In fact, the meanings for which a form is ultimately shared in the largest lexicon that can be established by a population can be divided into two sorts, namely *first-order* and *second-order* ones. The main difference between the two sorts is that first-order meanings will have a corresponding form shared from the very beginning, effectively introducing just a single new mapping to the agent and population lexicons. On the other hand, second-order shared meanings can have a first communication attempt about them fail, resulting in an additional distinct mapping being added to the lexicon of each agent.

Finally, it is also possible that a mapping is eliminated from an agent's lexicon if it is consistently misused and subsequently punished, e.g. if $Agent_1$ uses "*mm-u*" to refer to <event$_1$ entity$_1$> and $Agent_2$ then interprets "*mm-u*" as <entity$_1$>, resulting in communicative success $s = 0$. In the above case, it is then possible that the weights of both shared and non-shared mappings are punished (see section 3.3.2) and, if these end up having a non-positive value, eliminated from the lexicon of one or more agents. The following interaction is based on lexicon 3 from the preceding subsection.

Interaction 6.

- Event: [event$_1$ entity$_1$]
- Speaker partition ($Agent_1$): [<event$_1$ entity$_1$>]
- Utterance: *("mm-u")*
- Hearer interpretation ($Agent_2$): [<entity$_1$>]
- Outcome: $s = 0$ *and* $\pi = -0.4$

Lexicon 7.

Meaning \ Agent	*Agent*$_1$	*Agent*$_2$
<event$_1$ entity$_1$>	"*oo-u*" ($w = 0.6$) "~~*mm-u*~~" ($w = -0.3$)	"*oo-u*" ($w = 0.6$)
<event$_1$>	"*i-k*" ($w = 0.2$)	"*i-k*" ($w = 0.2$) "*f-e*" ($w = 0.1$)
<entity$_1$>	"*f-e*" ($w = 0.1$)	"~~*mm-u*~~" ($w = -0.3$)

Table 4.8: Sample agent lexicon after interaction 3 ($s_{min} = 0.4$).

If the above scenario occurs and one or both of the agents end up without a form association for one or more of the observable meanings, then the process of inventing and learning new lexical items can happen anew. In particular, there exist certain combinations for the larger event types where a non-shared mapping can be established for one of the meanings from the minimum dominating set with the help of another, which, after being forgotten in a subsequent interaction with a different meaning partition, can then have a non-shared mapping of its own being learned by the agents, resulting in both of these becoming second-order meanings, as discussed above.[2] Based on the newly defined distinction between first-order and second-order meanings with shared mappings, an updated table can be constructed that better reflects the conditions which can result in the maximum possible lexicon size in a population of two agents, which is provided below.

Event type \ s_{min}	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
R^1 (3 meanings)	2+0	2+0	2+0	2+0	2+0	3+0	3+0	3+0	3+0	3+0
R^2 (6 meanings)	2+1	2+1	2+1	2+1	1+4[3]	5+0	3+2	6+0	6+0	6+0
R^3 (10 meanings)	2+1	2+1	2+1	1+3	1+4	3+4	3+4	9+0	10+0	10+0

Table 4.9: Number of first-order and second-order meanings with a shared mapping required to learn a form for every meaning in the world.

Population Lexicon

The reason why such a sizeable portion of the section has been dedicated to the determination of the number of first-order and second-order meanings with a shared mapping that are required to learn a form for every meaning in the world is that knowing these allows one to compute the limits of the population lexicon, as described in section 4.2.1 above. In particular, any meaning which does not have a first-order shared mapping associated with it will be represented in every agent's lexicon by two different forms. Accordingly, the maximum agent lexicon size in a population of two agents, regardless of the number of conversations in the limit, can be defined as follows, with the set of

[2] For an example, take the same [event$_2$ entity$_3$ entity$_4$] event discussed above and $s_{min} = 0.4$. In this configuration, interpreting one of two meanings results in a payoff, while interpreting one in three yields a punishment. Accordingly, agents can agree on <event$_2$> and <entity$_4$> while misinterpreting <entity$_3$ entity$_4$> and <event$_2$ entity$_3$>, respectively, then delete the shared mappings associated with the two single-component meanings in failed interactions based on the [<event$_2$> <entity$_3$> <entity$_4$>] partition and then happen to agree on <entity$_3$ entity$_4$> and <event$_2$ entity$_3$> while wrongly interpreting <event$_2$> and <entity$_4$>, respectively. In effect, only one meaning (<event$_2$ entity$_3$ entity$_4$>) should be a first-order shared one, while the four mentioned above can be second-order shared and <entity$_3$> – completely non-shared.

[3] Also 2+1, which results in the same population-level maximum (17), but a lower agent-level maximum (10 vs 11).

first-order shared meanings referred to as M'_{shared} and the set of second-order shared meanings – as M''_{shared}.

$$\max|Lex^a| = |M'_{shared}| + (|M| - |M'_{shared}|) * 2 \tag{4.2}$$

Plugging the values for the minimum number of shared lexical items between agents from table 4.9 into equation 4.2 results in a table of the maximum agent lexicon size in the population of two agents, presented below.

Event type \ s_{min}	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
R^1 (3 meanings)	4	4	4	4	4	3	3	3	3	3
R^2 (6 meanings)	10	10	10	10	11	7	9	6	6	6
R^3 (10 meanings)	18	18	18	19	19	17	17	11	10	10

Table 4.10: Maximum agent lexicon size in simulations with two agents.

When considering the joint lexicon of the population, this will contain a single mapping for each of the meanings with first-order shared mappings, three for meanings with second-order shared mappings (one that is shared and further two that have different forms referring to the meaning) and four for the meanings without a shared mapping (two for each of the communication attempts). The reason why failed interpretations result directly in two additional mappings on the population level, as opposed to one in the agent lexicon, is that forms that were meant to refer to one meaning and were interpreted to mean something else will by definition be assigned to different meanings in the lexicons of the two agents, thus creating two new mappings in the population. Based on the above, the total maximum number of distinct mappings in the joint lexicon of a population of two agents is:

$$\max|Lex^{pop}| = |M_{shared}| + |M''_{shared}| * (2+1) + (|M| - |M'_{shared}| - |M''_{shared}|) * (2+2) \tag{4.3}$$

The following table 4.11 presents the maximum population lexicon sizes computed with the help of equation 4.3 and values from table 4.9 for the 30 different event type and minimum success threshold combinations, in simulations with two agents.

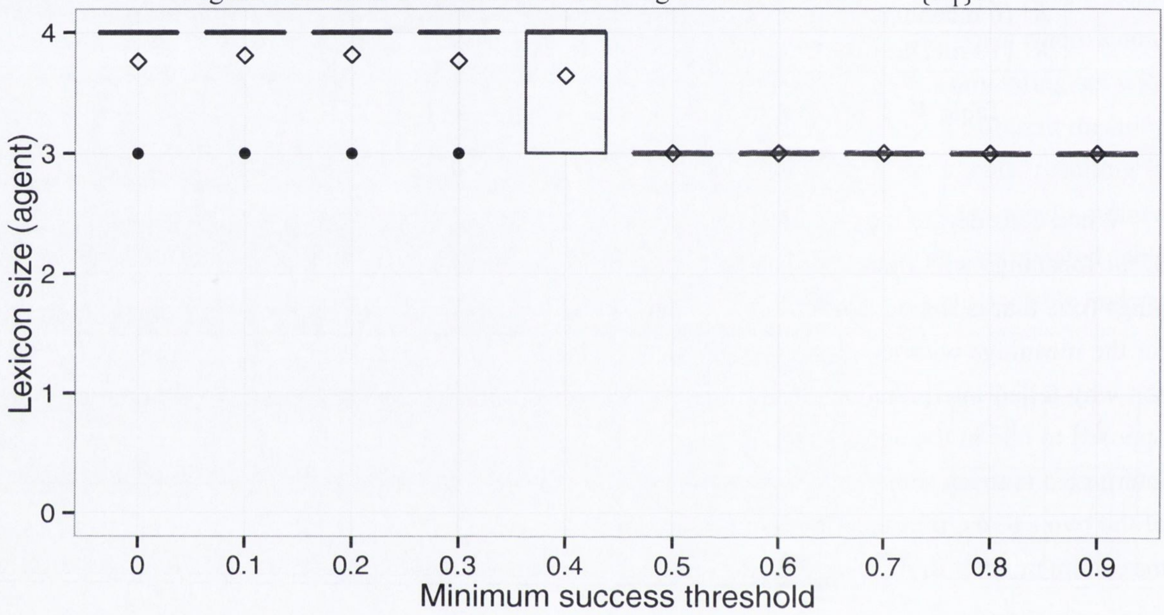
Event type \ s_{min}	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
R^1 (3 meanings)	6	6	6	6	6	3	3	3	3	3
R^2 (6 meanings)	17	17	17	17	17	9	13	6	6	6
R^3 (10 meanings)	33	33	33	34	33	27	27	13	10	10

Table 4.11: Maximum population lexicon size in simulations with two agents.

As will be the case for each of the observed properties, simulations have been conducted to determine the actual distribution of the number of unique mappings acquired by the agents based on the type of the solitary event occurring in their world and the level of minimum success imposed on them. These distributions are then plotted by a combination of several box plots, allowing the reader to quickly identify the range of values taken up by the given property under each of the conditions, as well as the mean, represented by the diamond icon. The same 500 simulations for each of the parameter combinations will be examined throughout this section, each consisting of 1000 interactions between the two agents in the simulated population.

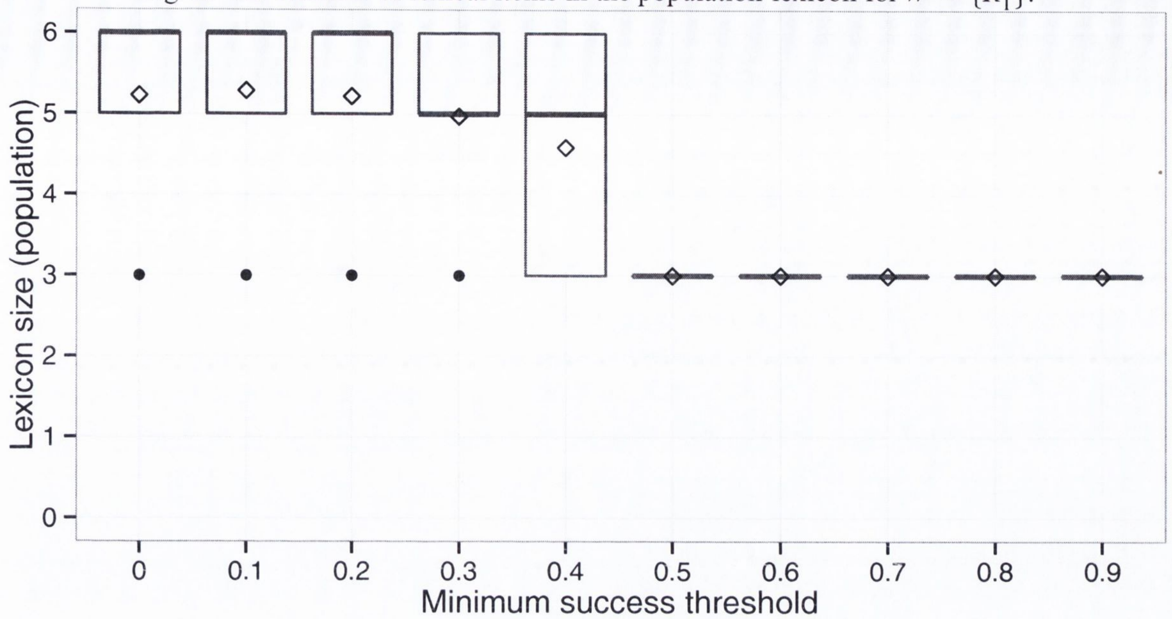
Figure 4.1 shows that in a world with one event of type R^1, for any level of minimum success $s_{min} \leq 0.4$, i.e. where the minimum set of first-order shared mappings required to learn a lexical item for every potential meaning is equal to two and thus max $|Lex^a| = 4$, the agents actually predominantly end up employing the maximum possible number of mappings, as opposed to the optimal three that would ensure perfect communicative success. For $s_{min} \geq 0.5$, the minimum set of meanings with a shared mapping M'_{shared} becomes equivalent to M and hence the agents cannot learn more than three mappings, which is reflected in the graph.

Figure 4.1: Number of lexical items in agent lexicons for $W = \{R_1^1\}$.



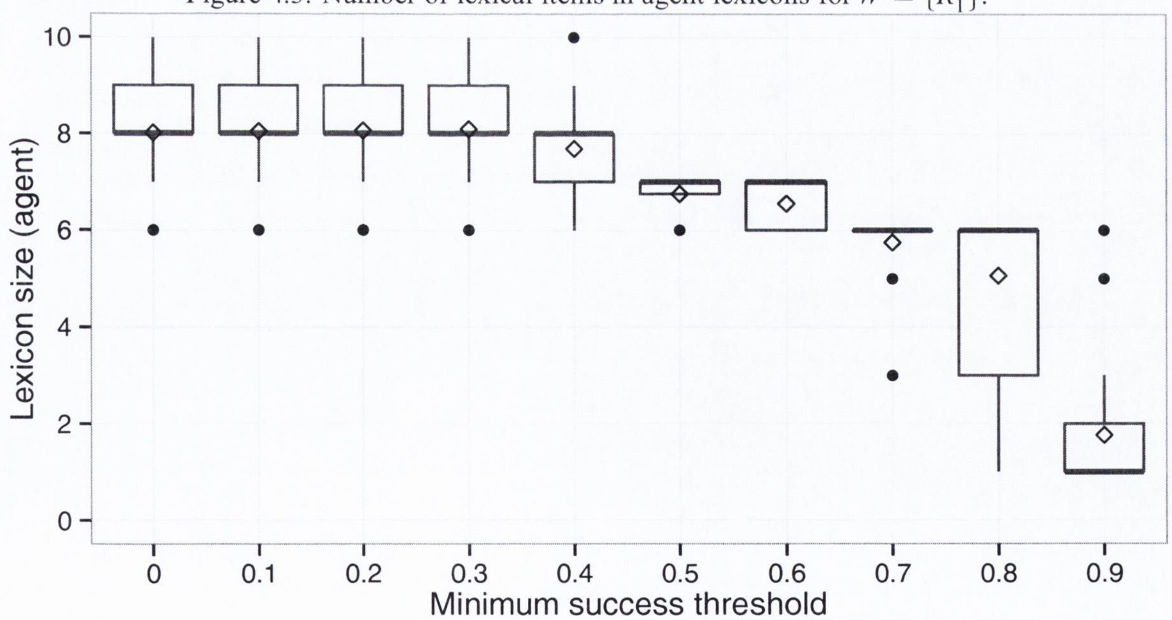
Given that the majority of the agents in experiments conducted with an R^1 event end up having the maximum four items for $s_{min} \leq 0.4$, it is only a logical consequence that the joint lexicon of the corresponding populations will have a total of at least five items, in the case where both agents have four different mappings in their lexicons. This is indeed confirmed by figure 4.2, although the suggestion is that in about a quarter of the simulations, the agents actually end up agreeing on the form for the meaning that is not a part of the two-item minimum dominating set at the second attempt, resulting in a population lexicon size of five. Inevitably, for $s_{min} \geq 0.5$, all mappings are shared between the agents and hence the population lexicon (and thus its size) is equivalent to the lexicon of any individual agent from the corresponding population.

Figure 4.2: Number of lexical items in the population lexicon for $W = \{R_1^1\}$.



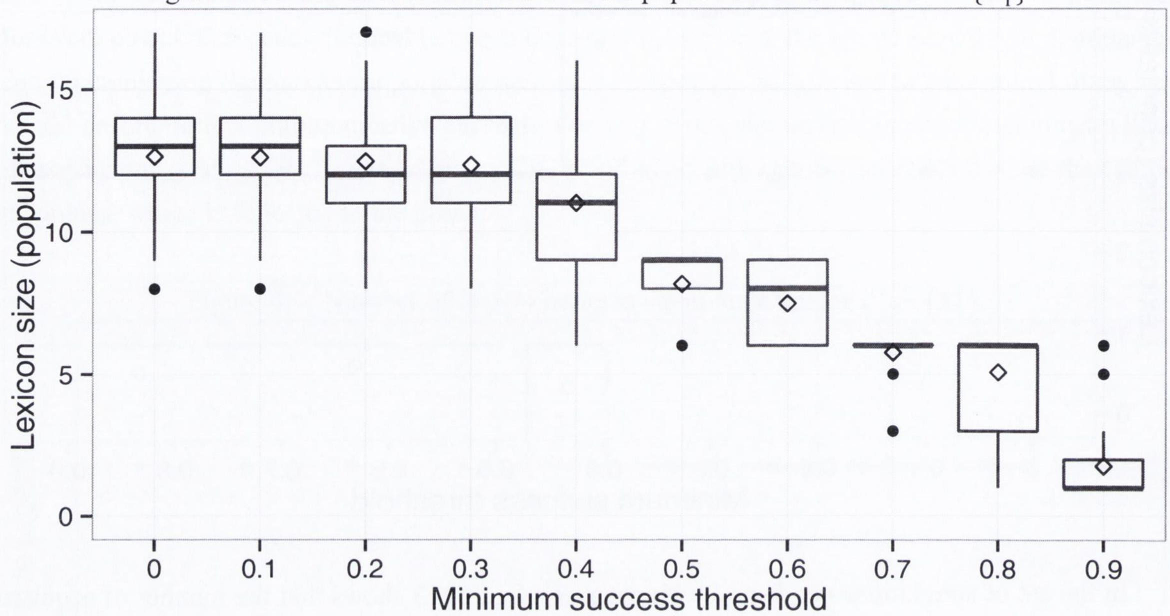
In the set of simulations conducted with event R^2, figure 4.3 shows that the number of acquired lexical items is predominantly concentrated around eight out of the maximum ten for $s_{min} \leq 0.4$ and the maximum seven mappings 7 for $0.5 \leq s_{min} \leq 0.6$). As was the case in the experiment conducted with event R^1 above, it can be clearly seen that for $s_{min} \geq 0.7$ the agents could not learn more than six mappings due to the minimum success requirement.

Figure 4.3: Number of lexical items in agent lexicons for $W = \{R_1^2\}$.



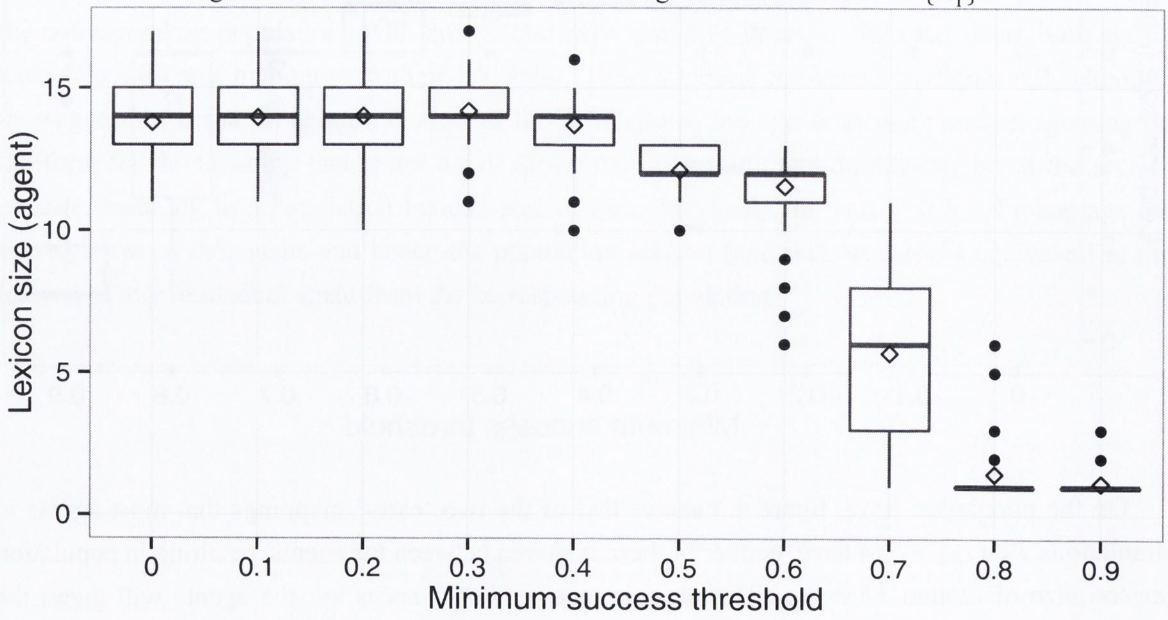
On the population level, figure 4.4 shows that of the two 'extra' mappings that most agents in simulations with $s_{min} \leq 0.4$ have, neither of these is shared between the agents, resulting in population lexicon size of around 12 items. However, it appears that chances are the agents will guess the meaning for at least one of the forms not referring to a meaning from the minimum dominating set, as the population lexicon never reaches the maximum size of 18. On the other side, for $s_{min} \geq 0.7$, all of the lexical items are shared between the two agents, regardless how few of these are actually learned by the agents, resulting in a population lexicon size equivalent to that of the individual agent lexicons.

Figure 4.4: Number of lexical items in the population lexicon for $W = \{R_1^2\}$.



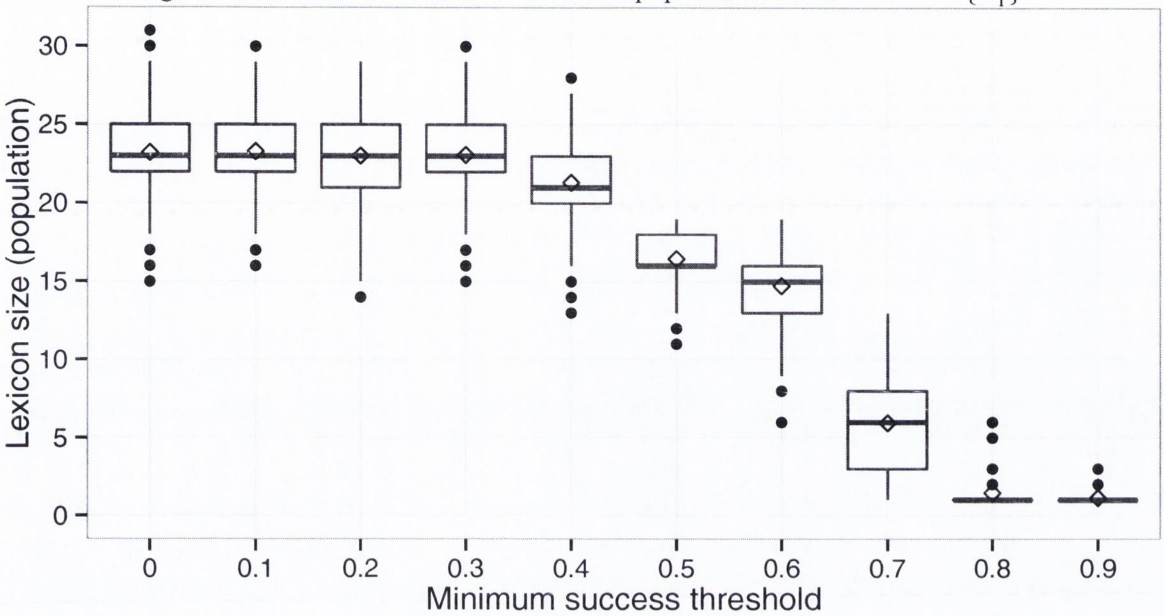
A very similar picture develops in experiments conducted with event type R^3, with the majority of the agents slightly short of the maximum 18 mappings that can be learned for $s_{min} \leq 0.3$, the 15 for $s_{min} = 0.4$ or the 9 for $0.5 \leq s_{min} \leq 0.6$. Similarly to the experiments with a two-argument event, a very pronounced drop can be observed for simulations with $s_{min} \geq 0.8$, where the agents quite clearly fail to learn over 50% of possible words after 1000 interactions, again dropping to just one for $s_{min} = 0.9$ as an effect of the very high success threshold and the much lower chances of guessing all of the meanings from an encountered utterance.

Figure 4.5: Number of lexical items in agent lexicons for $W = \{R_1^3\}$.



In simulations with an R^3 event, figure 4.6 shows that again, out of the four 'extra' mappings that each of the agents learns on average for $s_{min} \leq 0.3$, only one is likely to end up being shared, resulting in population lexicons consisting of 23 items. A similar trend continues throughout all levels of minimum success imposed on the agents, with clearly a very low chance of the agents correctly guessing the meanings of the forms utilized by one another during interactions.

Figure 4.6: Number of lexical items in the population lexicon for $W = \{R_1^3\}$.



The main conclusion that can be drawn from the figures depicting the number of distinct mappings that two agents can learn within 1000 interactions is that the time frame is certainly not sufficient if a high requirement for success is imposed on the communicative process. The reason for that is that the chances of getting the interpretation of the forms utilised in different partitions in successive interactions completely right are fairly low. In addition to that, the small payoffs and the large punishments issued in the corresponding simulations result in a low tolerance for failure, meaning that a single partially failed interaction in between several successful ones can take the agents right back to square one. However, the following sections will show that while the lexicons of such agents might be small, some other features of these lexicons may actually be superior to those of agents who are under less pressure for communicative success.

Unique Meanings

The number of unique meanings that the agents can learn in the LEW is bound by the events that can be observed in their world. In the model, events are represented as combinations of one event atom and one or more event arguments, which are meant to represent different entities such as persons and objects involved in the event. In general, an event R can be considered to have a certain length, which is defined as the sum of its arguments and the event head atom.

$$|R| = |R^{args}| + 1 \tag{4.4}$$

As described in section 3.2.2, agents individuate events by partitioning them into individual meanings comprised of one or more event components each. The partitioning process has no restrictions imposed on it other than that event components cannot be repositioned after an event has been defined, i.e. when putting the partitioned meanings back together, they should result in an identically ordered copy of the original event. This follows from modelling any perspective on an event as an element of the set of possible subsequences of the event. According to the above definition of event generation and individuation, the number of unique meanings $|M|$ that can be potentially generated from an event R of length $|R|$ is defined by the number of possible subsequences that can be extracted from it and can thus be written down as follows:

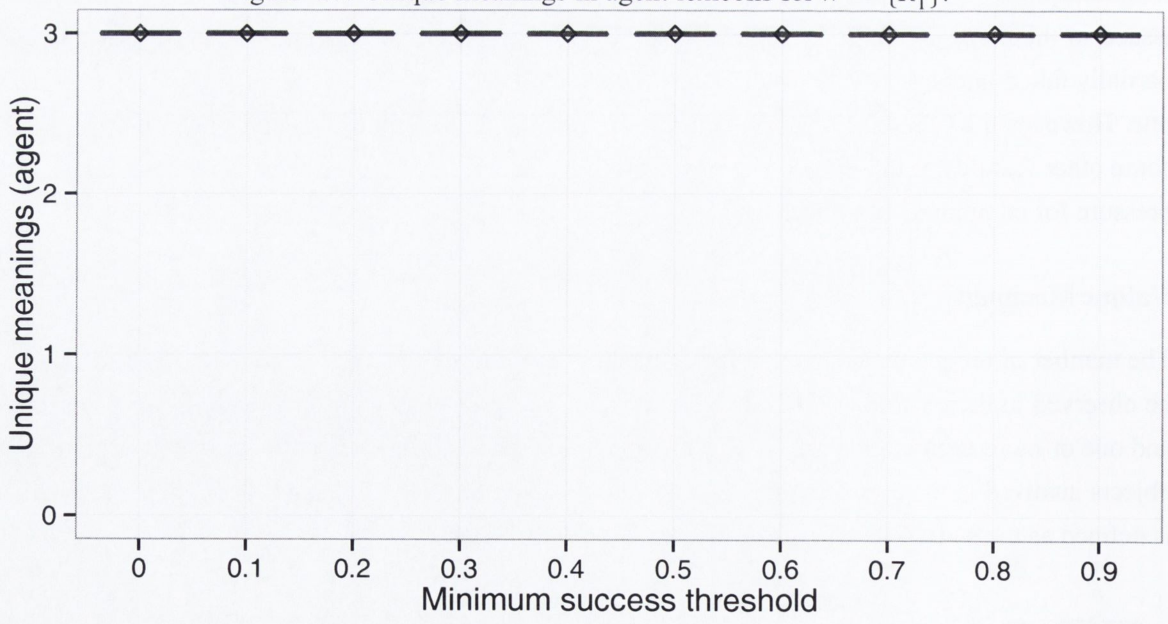
$$|M| = \sum_{i=1}^{|R|} |R| - i + 1 = \sum_{i=1}^{|R|} i = \frac{|R| * (|R| + 1)}{2} \tag{4.5}$$

Since the agents cannot invent any meanings on their own, the number of unique meanings in their lexicons, as well as the population lexicon, is at most the sum of the number of unique parts that can be extracted from the w events in the world $W = R_1, R_2, .., R_w$ that they are inhabiting.

$$\max |M^a_{unique}| = \max |M^{pop}_{unique}| = \sum_{i=1}^{w} |M_i| \tag{4.6}$$

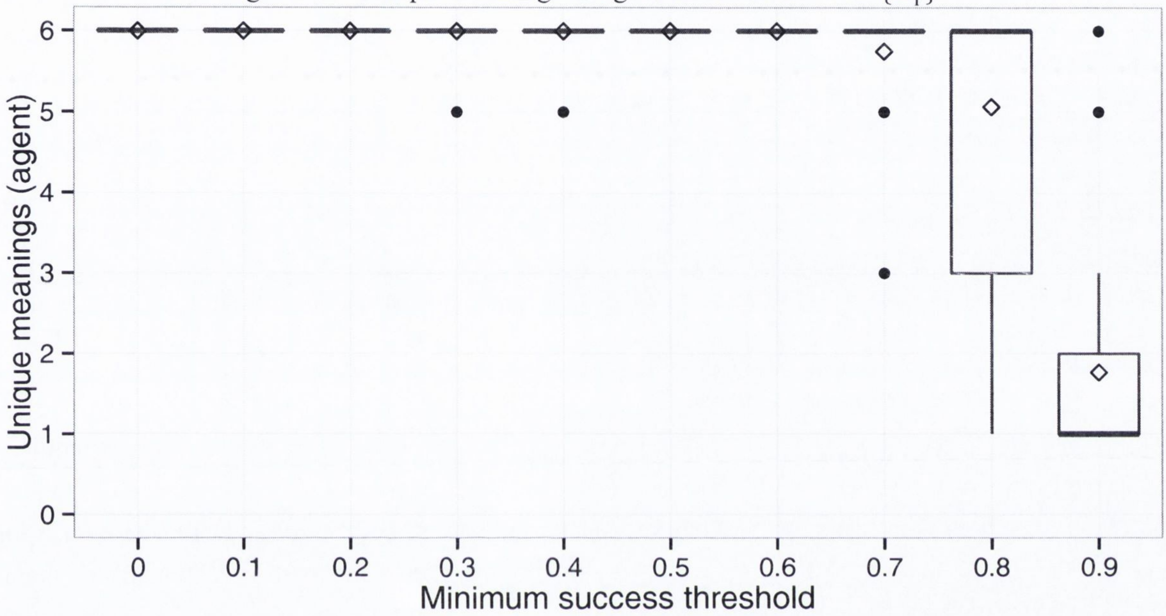
Accordingly, in a world with just one event with fixed arguments, the number of meanings that the agents can potentially learn depends solely on the length of that event, i.e. on the number of arguments that the event takes. The following graphs show the distribution of the number of unique meanings with a form associated with them at the end of the conducted simulations. Figure 4.7 shows that, for the event of the smallest type (R^1) that takes one argument and has thus only three potential meanings that can be extracted from it, the agents end up learning a form for every one of those meanings after 1000 interactions regardless of the level of minimum success imposed on their interactions. Logically, if the agent lexicons always end up having a form for every meaning in the agents' world, then the population lexicon, by extension, will have at least one form associated with every meaning as well.

Figure 4.7: Unique meanings in agent lexicons for $W = \{R^1_1\}$.



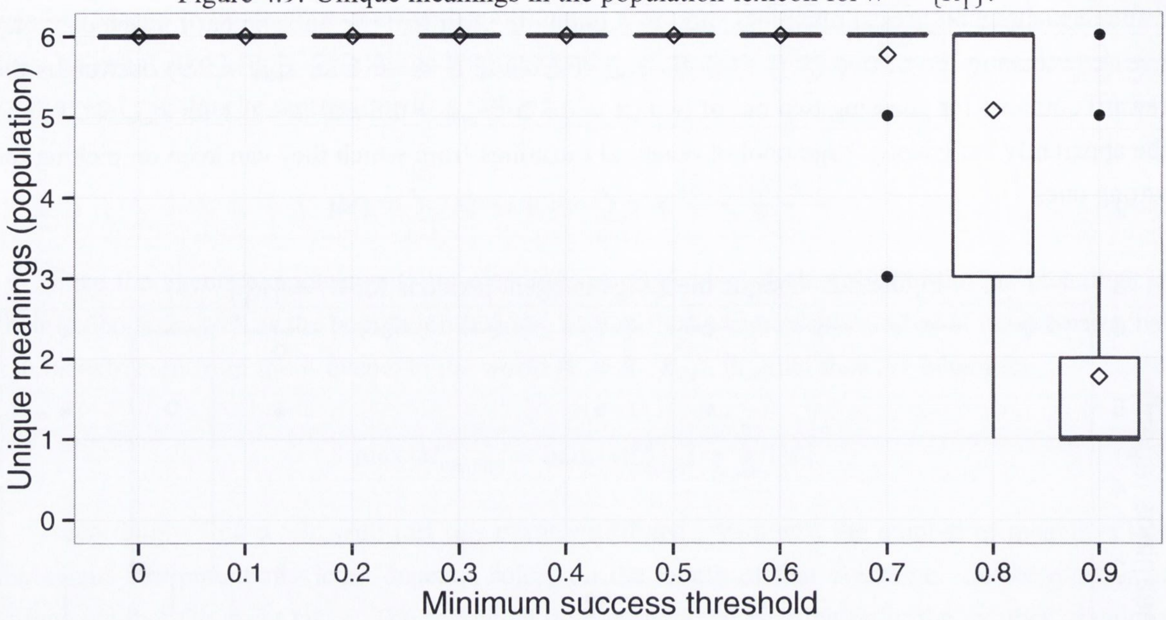
The following figure 4.8 shows that increasing the number of arguments by just one to two, i.e. by populating the world with one event of type R^2, it is no longer guaranteed that the agents will be able to associate a meaning with each of the six potential meanings that can be extracted from the event within the 1000 interactions. For $s_{min} \leq 0.7$, the agents still overwhelmingly learn a mapping for every possible meaning, with the exception of a handful of agents were not quite able to associate a form with each of the available meanings. However, for $s_{min} \geq 0.8$ agents actually find it extremely difficult to agree on lexical mappings, not least due to the fact that the punishment for one correctly guessed meaning out of two ($\pi = -0.3$ for $s_{min} = 0.8$ and $\pi = -0.4$ for $s_{min} = 0.9$) outweighs the reward obtained for guessing two out of two ($\pi = 0.2$ for $s_{min} = 0.8$ and $\pi = 0.1$ for $s_{min} = 0.9$) and the apparently sufficiently large pool of potential meanings from which they can keep on picking the wrong ones.

Figure 4.8: Unique meanings in agent lexicons for $W = \{R_1^2\}$.



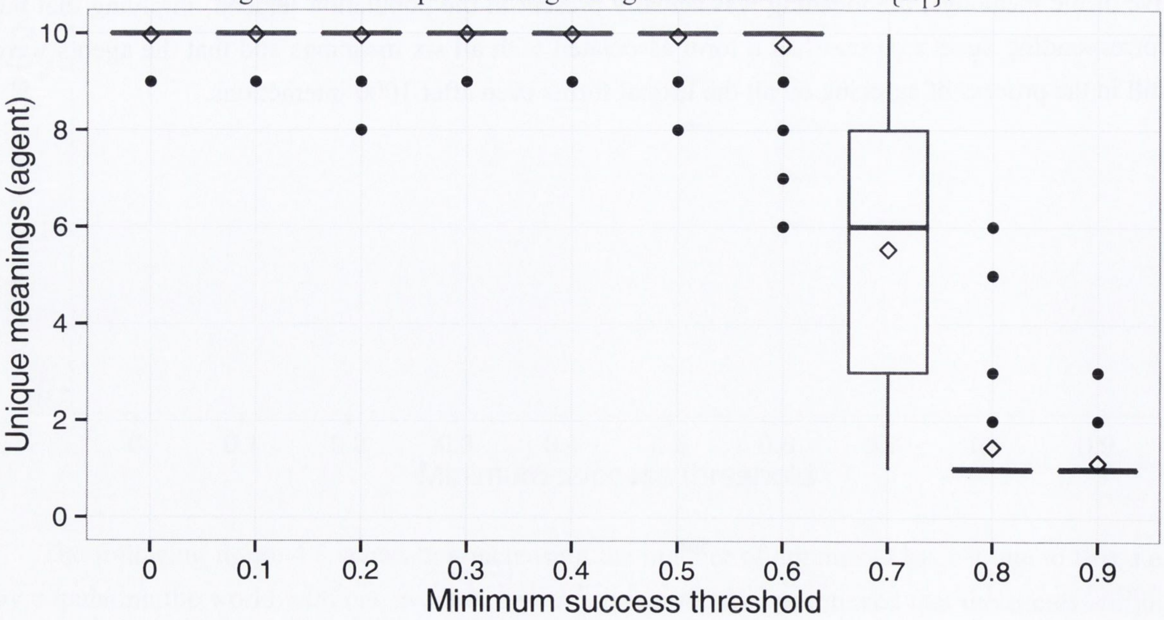
Since there were cases in which agents did not learn how to express each of the six available meanings, here it makes sense to take a look at the number of meanings that were actually represented in the combined lexicons of all agents, i.e. the so called population lexicon. These numbers, depicted by figure 4.9, indicate that in the two cases where one of the agents only had a form associated with five of the meanings, the meaning was actually present in the population lexicon, implying that the corresponding agent's partner had a form associated with all six meanings and that the agents were still in the process of agreeing on all the lexical forms even after 1000 interactions.

Figure 4.9: Unique meanings in the population lexicon for $W = \{R_1^2\}$.



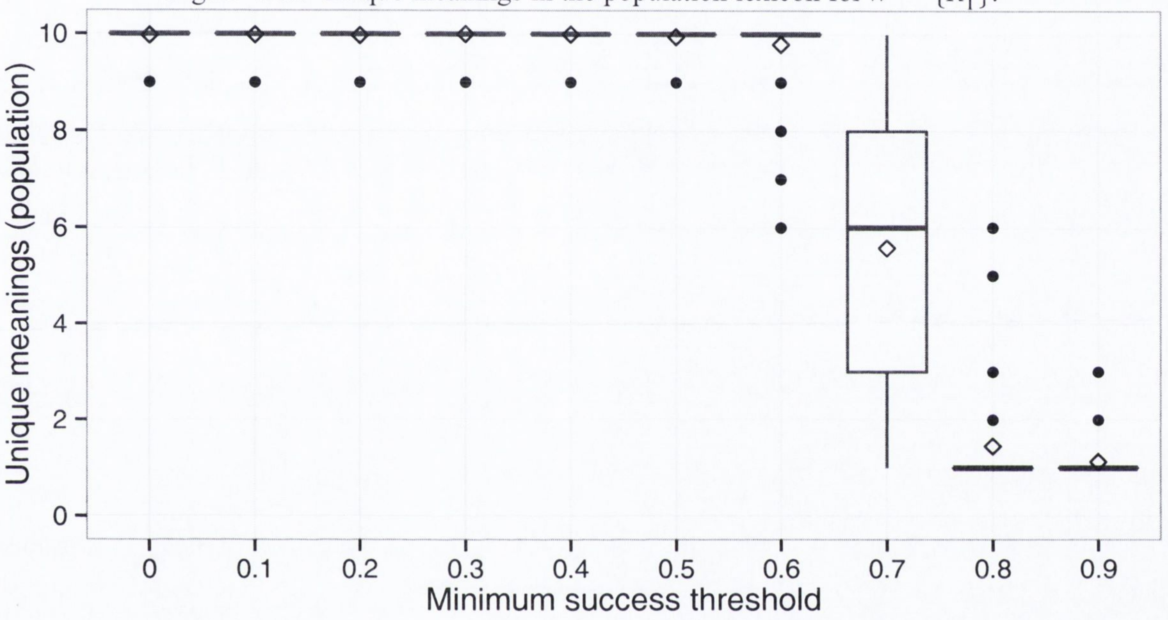
The picture worsens further with the three-argument R^3 as the only event in the agents' world, as can be seen from figure 4.10. Hereby, not even the smallest level of minimum success required from the agents guarantees that 500 out of 500 times both agents will learn to express each of the ten potential meanings with some form. The limit of having the vast majority of agents know a form for every meaning shifts slightly from $s_{min} = 0.7$ to $s_{min} = 0.5$ when compared to agents talking about R^2. Furthermore, for $s_{min} = 0.9$ the overwhelming majority of the agents actually ends up knowing a word to express just one meaning - usually the joint partition including all four components of the event, which, being only used in isolation, is not dependent on any other meaning being interpreted correctly and is thus impossible to unlearn having been correctly interpreted at least once.

Figure 4.10: Unique meanings in agent lexicons for $W = \{R_1^3\}$.



On the population level, the number of distinct expressible meanings, depicted by figure 4.11, suggests that in the majority of the cases, if one of the agents only knows how to express a fraction of the available meanings, then the other agent in the population will not know how to express any more meanings. This is quite logical, given that the agents end up having overwhelmingly the same lexicons. However, while the set of distinct forms in the agents' lexicons will indeed always be the same, as will be discussed in the following subsection, agents do have the possibility of assigning a form to different meanings, in particular in simulations with a low level of minimum success. An example of this happening was already observed above in the experiment with an R^2 event and is also present in this experiment for $s_{min} = 0.2$ to $s_{min} = 0.5$, where there is an agent each with eight meanings only in his lexicon, yet the corresponding population lexicon having at least nine meanings expressed in it.

Figure 4.11: Unique meanings in the population lexicon for $W = \{R_1^3\}$.



In summary, the results of the simulations performed with the LEW model suggest that agents are by no means guaranteed to acquire a mapping for even as few as six or ten meanings that are available to them over a period of 1000 interactions. For lower levels of s_{min}, the minimum dominating set, i.e. the minimum set of mappings that need to be agreed on by the agents is still relatively small, which means that the agents only need to agree on a fraction of the meanings in order to learn a mapping (shared or not) for every other meaning. However, as the level of minimum success is increased and with it the number of mappings that need to be shared, agents find it increasingly more difficult to reach the minimum that would allow them to learn to express every meaning in their world.

Unique Forms

There are certain implicit restrictions imposed on the learning of new forms by agents in the LEW. In particular, when acting as a speaker, agents will occasionally invent new forms when they have none assigned with a meaning that needs to be expressed. When this happens, the agent will pick a form that has not been used within the simulation yet, based on the assumption that he will want to employ a new lexical item that no one has potentially used in a different context yet. When acting as hearers, agents always check their lexicon first when encountered with a set of forms from an utterance. If a form can be found in the agent's lexicon, he will select the associated meaning as the interpretation of the form, regardless if the meaning is actually present in the context, i.e. the observed event.[4] Accordingly, when considering the lexicon of an agent, the number of unique forms in it will be equivalent to the number of mappings in it, since an agent has no way of assigning multiple meanings to a single form, either when speaking or when listening.

$$\max |F^{a}_{unique}| = \max |Lex^{a}| \tag{4.7}$$

In terms of the overall population lexicon, it can clearly happen that a form is assigned to a different meaning across different agents. For this to happen, the hearer needs to simply incorrectly guess the meaning of one of the forms uttered by the speaker, while at the same time guessing a sufficient number of forms correctly so as to end up above the minimum success requirement and obtaining a payoff for all of the involved mappings. However, based on the observation that the weights of mappings with corresponding forms will always be equivalent between the two agents, it can be concluded that both agents will always have the exact same set of unique forms in their lexicons. It then follows that the set of unique forms in any individual agent's lexicon is equivalent to the set of unique forms known by both agents, which is further equivalent to the set of unique mappings known by each agent. Accordingly, in a population of two agents, the maximum number of unique forms both in every agent's lexicon and the joint lexicon of the population follows from equation 4.2:

$$\max |F^{a}_{unique}| = \max |F^{pop}_{unique}| = |M'_{shared}| + (|M| - |M'_{shared}|) * 2 \tag{4.8}$$

4.2.2 Synonymy & Homonymy

When estimating both agent and population lexicon synonymy and homonymy limits, it is assumed that agents have reached a state in which they have assigned a form to each of the world's meanings. This assumption is made for the sake of keeping limit estimates within plausible and meaningful bounds. For example, it is possible that during intermediate states of lexicon acquisition, agents will both know a certain set of forms, but not share the meanings associated with any of these, resulting in a population homonymy level of $Hom^{pop} = 1.0$, i.e. have more than one different meaning associated with every form in the population lexicon. While the potential states in which agents can find themselves throughout the lexicon acquisition process are equally of interest, the focus of the current analysis will remain on the properties of the relatively finalized lexicons, which would allow one to appropriately compare like with like.

Agent Synonymy

Both synonymy and homonymy levels are clearly highly dependent on the amount of shared and non-shared lexical items between agents, as well as the overall lexicon size (which in turn is also based on the number of shared mappings). Intuitively, it would seem that agent synonymy levels should not be dependent on the amount of mappings that an agent shares or does not share with the other agent(s). One could imagine a scenario where the agents both share several dozens of mappings all referring to just a handful of different meanings, resulting in high synonymy, as well as one where they do not share a single mapping, yet do not have multiple forms referring to any of the meanings within their

[4] Experiments in which having the interpreted meaning appear in the observed event is imperative are being conducted by the author as part of a separate project.

individual lexicons. However, the above scenarios are not necessarily possible in the LEW. In fact, as will be described below, for an agent to have multiple forms associated with a single meaning in the model, i.e. for the agent's lexicon to have synonyms in it, he cannot share all of his mappings with the other agent(s).

The reason for the above is that generating new synonyms on the agent level is a capacity that is only open to the hearer who is not constrained by the principle of contrast, i.e. is not prevented from mapping additional forms to a meaning that already has one or more forms associated with it (cf. Clark, 1987), and can thus, on hearing an unknown form, interpret it to stand for any existing meaning, even one already associated with a different form previously. Accordingly, if both agents already shared the mapping with the previous form, then the speaker agent would have used this form, and not invented a new one when referring to the corresponding meaning. The only option that remains for a synonym to be implanted in the hearer agent's lexicon is for him to incorrectly interpret a form intended to stand for a different meaning, for which the agents do not yet agree on a form for.

The first conclusion that can be instantly drawn from this is that in a population of two agents, agent synonymy can only be above zero if the lexicons of the agents are not completely synced, i.e. if there are forms with unmatched meanings in the agents' lexicons. Conversely, if agent lexicons are not identical, then the forms from non-shared or second-order shared mappings will have to be associated with some meanings in their corresponding lexicons, for which they already have (a potentially shared) word. Taken together, it can be stated that $Syn^a > 0$ iff $|Lex^{pop}| > |M|$ or $|Lex^a| > |M|$, i.e. the lexicon of at least one agent (and thus also the combined lexicon of the population) is larger than the number of distinct meanings in the world.

In such lexicons where there is at least one form more than there are distinct meanings, the 'extra' form needs to be associated with some meaning, resulting in at least one meaning having synonymous forms. Regardless of the number of 'extra' forms, there is no reason why the agent could not associate all of these with the same meaning, which implies that the minimum number of meanings associated with multiple forms does not change with the size of the lexicon or the number of 'extra' forms. Since the number of unique meanings in the agent's lexicon M^a_{unique} is also fixed at a certain maximum level, the minimum level of agent synonymy is fixed as follows (in cases where there is at least one synonymous form):

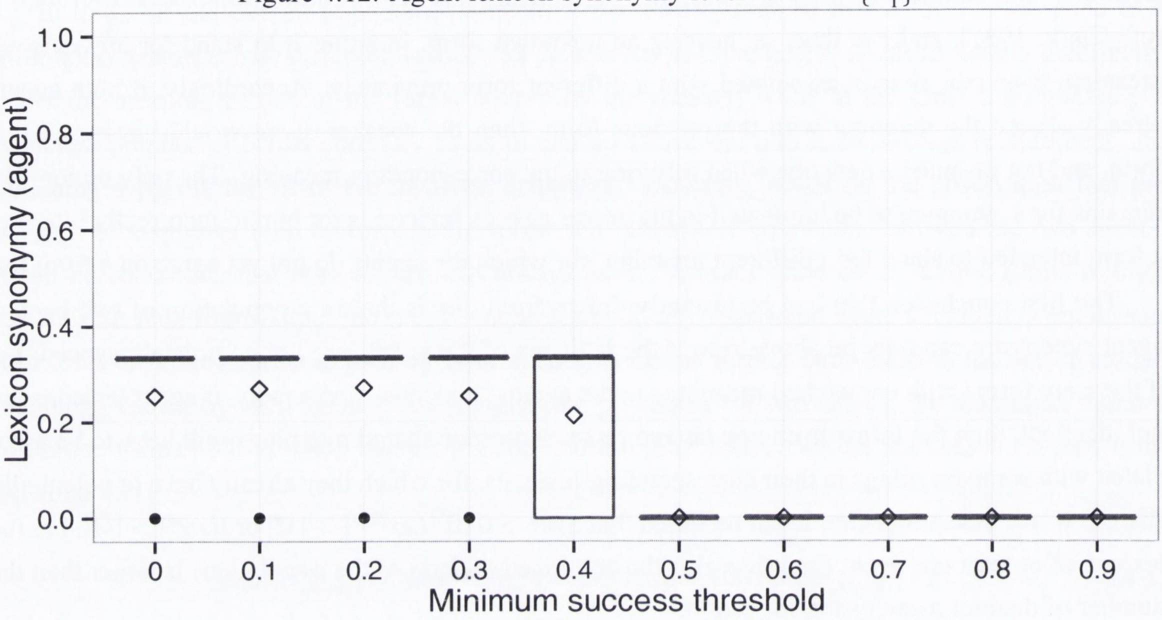
$$\min Syn^a = \frac{1}{|M^a_{unique}|} \tag{4.9}$$

Analytically, the maximum level of agent lexicon synonymy, i.e. $Syn^a = 1.0$ is reached when *every* meaning in the agent's lexicon has two or more forms associated with it. However, this level cannot be reached by two agents as the number of distinct forms that the agents can learn is $|F^{pop}_{unique}| < 2 * |M|$. This equation is based on the premise that two forms can only be learned via non-shared mappings; however, since the two agents will always need to share at least one mapping in order to learn a form for every meaning in the world, the number of meanings without a shared form will at most be $|M| - 1$. Accordingly, the highest level of lexicon synonymy that can be reached would result from an agent assigning every 'extra' form to a different meaning, effectively acquiring as many synonyms as he has non-shared (or second-order shared) mappings, or put formally:

$$\max Syn^a = \frac{|M| - |M'_{shared}|}{|M|} = 1 - \frac{|M'_{shared}|}{|M|} \tag{4.10}$$

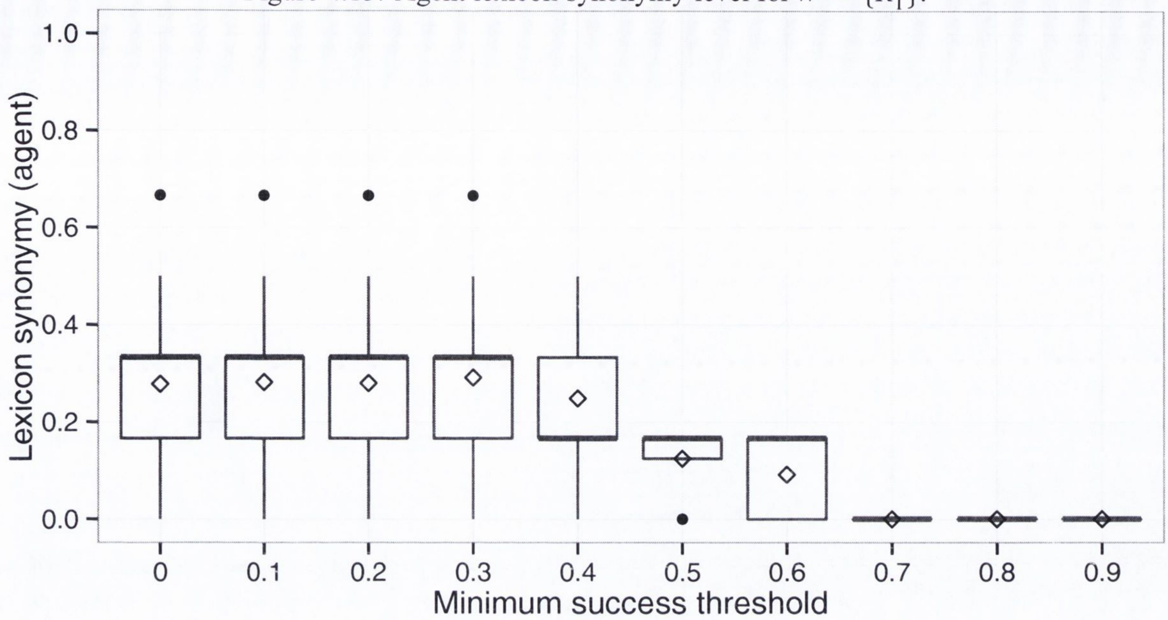
The results of the 500 simulations conducted for each level of minimum success depict the actual distribution of agent synonymy within its limits. In a world with one event of type R^1, agents following $s_{min} \leq 0.4$ can end up having one non-shared mapping, resulting in an 'extra' form being assigned to some meaning in the lexicon of both agents. In such a scenario, $\max Syn^a = \min Syn^a$, meaning that the level of agent synonymy becomes a direct reflection of the number of unique forms known by the agents with no adjustments to the distribution possible, as confirmed by figure 4.12.

Figure 4.12: Agent lexicon synonymy level for $W = \{R_1^1\}$.



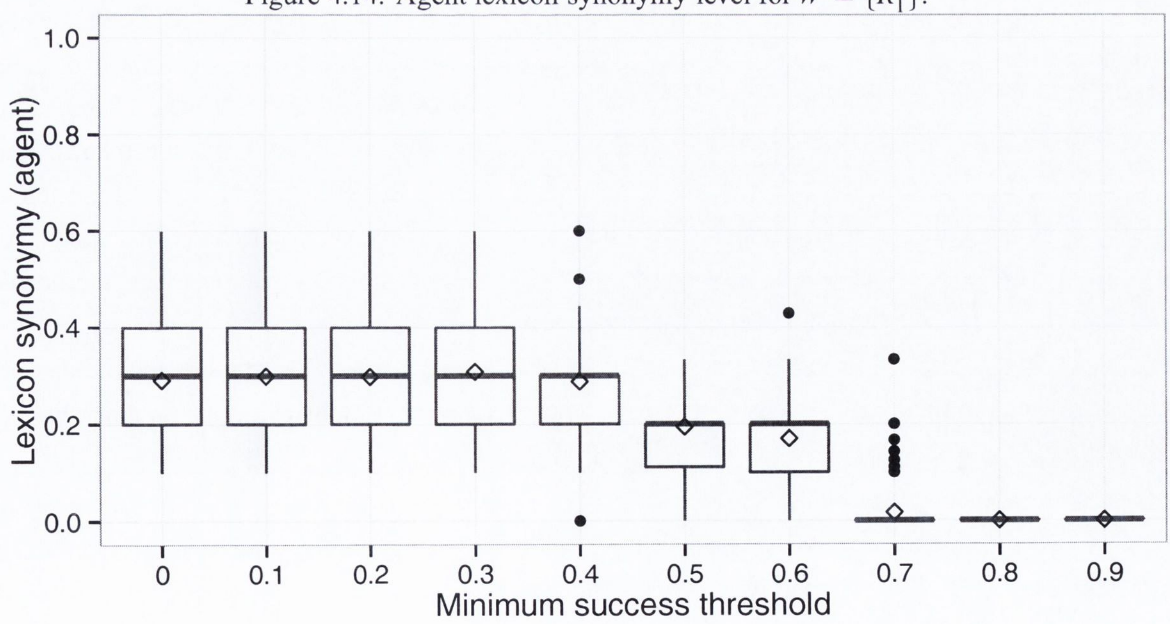
For experiments with event of type R^2, there is significantly more room for the agents to assign the forms from non-shared mappings to different meanings. Figure 4.13 shows that agent synonymy is still closely tied to the number of unique forms (and hence the number of 'extra' forms) in the agents' lexicons, with the two extra forms being evenly distributed between the six meanings on average, resulting in an agent synonymy level concentrated around one in three for $s_{min} \leq 0.4$. For $0.5 \leq s_{min} \leq 0.6$, there is usually just one 'extra' form, leaving $\frac{1}{|M|} = \frac{1}{6}$ as the dominant level of agent lexicon synonymy. Finally, for $s_{min} \geq 0.7$, agents do not learn more forms than there are meanings and also never associated more than one form with the same meaning (due to the fact that obtaining a payoff for such high level of s_{min} requires full agreement over the utilized mappings), which naturally results in the lack of agent lexical synonymy.

Figure 4.13: Agent lexicon synonymy level for $W = \{R_1^2\}$.



A similar trend can be observed in simulations with an R^3 event, as depicted by figure 4.14, although with slight adjustments. In particular, for $s_{min} \leq 0.4$, the four extra forms end up being distributed between two and four different meanings, resulting in average agent synonymy level between 0.29 and 0.31. For higher levels of minimum success threshold the results are largely the same, with agent synonymy nearly non-existent for $s_{min} \geq 0.7$ due to the same logic as described above.

Figure 4.14: Agent lexicon synonymy level for $W = \{R_1^3\}$.



In summary, agent synonymy levels clearly depend on the number of 'extra' forms that agents end up learning throughout the simulations. Logically, since there are then more forms than meanings, some meanings will end up have multiple forms associated with them, thus introducing agent synonymy. The actual amount of synonymy, i.e. if the extra forms are all assigned to the same meaning, or to different ones, or somewhere in between is then simply decided by the chances of picking a particular meaning combination for these forms. Interestingly, in most experiments with the lower levels of s_{min}, levels of synonymy observed are actually very similar to the WordNet-based estimation 0.38 in the English language. However, it is too early to tell if this is just a chance occurrence or a real trend.

Population Synonymy

Population synonymy follows slightly different general principles to lexicon synonymy. In particular, the minimum number of meanings with multiple forms associated with them in a population lexicon is actually 2 for any lexicon in which there is at least one non-shared mapping. This is due to the fact that the form from this mapping, by definition, needs to be assigned to two different meanings by the two agents and that the meaning that each agent assigns such a form to will either be a part of a shared mapping, in which case it will now have an extra form associated to it, resulting in a new agent and population lexicon synonym; or have two forms associated with it between the agents (once the other agent has invented a new form for the meaning, or assigned a form from another utterance to it), again resulting in a new population lexicon synonym. The number of further non-shared form associations is, however, equally irrelevant here, since both agents can theoretically assign all of the forms from non-shared mappings to the same two meanings.

$$\min Syn^{a} = \frac{2}{|M^{a}_{unique}|} \tag{4.11}$$

There is a slight difference in the maximum number of meanings with synonymous forms associated to them across the agent lexicons, when compared to individual lexicons. As has been outlined in the preceding section, an agent can only end up having as many meanings with multiple forms as there are 'extra' forms that he might end up distributing across these meanings. When considering the joint lexicon of the population, the same principle applies as well. However, even if both agents will have the same number of 'extra' forms; it does not imply that they have to assign these to same meanings. So, for example, $Agent_1$ can assign three 'extra' forms to meanings 1-3, while $Agent_2$ can end up associating these same forms with meanings 4-6. In addition to that, every meaning that is not from the first-order shared set, i.e. every meaning that has allowed for a new 'extra' form to be added to an agent's lexicon, may have either one (in the case of second-order shared meanings) or two different forms assigned to it between the two agents. While this does not result in additional lexicon synonymy on the agent level, it certainly contributes to population lexicon synonymy.

Accordingly, the joint lexicon of the population Lex^{pop} can at most have the following number of meanings with multiple forms mapped to them (the number cannot be higher than the total number of distinct meanings):

$$\max |M^{pop}_{syn}| = min((2 * |M''_{shared}| + 3 * (|M| - |M'_{shared}| - |M''_{shared}|)), |M|) \tag{4.12}$$

The maximum level of population synonymy in a population of two agents is then

$$\max Syn^{pop} = \frac{min((2 * |M''_{shared}| + 3 * (|M| - |M'_{shared}| - |M''_{shared}|)), |M|)}{|M|} \tag{4.13}$$

Since there are numerous scenarios in which less than half the meanings have a shared form, especially for a low minimum success threshold, it is quite possible that population lexicon synonymy level reaches 1.0 in a population of two agents, as confirmed by the following three figures. Figure 4.15 further shows that the level of population lexicon synonymy falls precisely between the minimum and the maximum levels, provided there are 'extra' forms available, with the majority of simulations ending with just two of the three meanings in the world having multiple forms associated with them.

Figure 4.15: Population lexicon synonymy level for $W = \{R_1^1\}$.

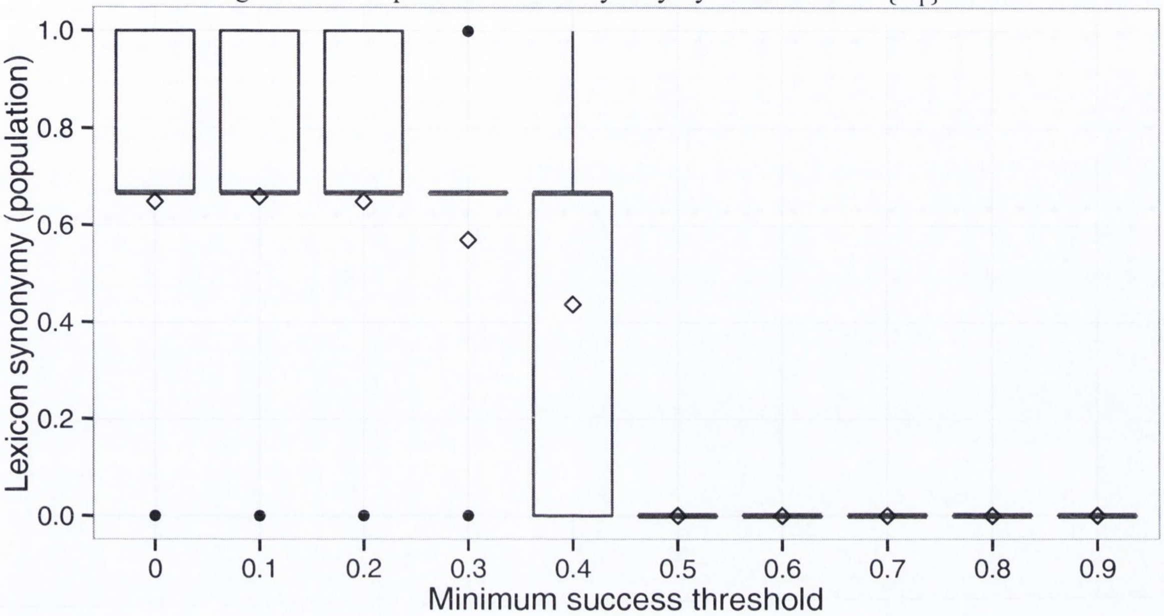
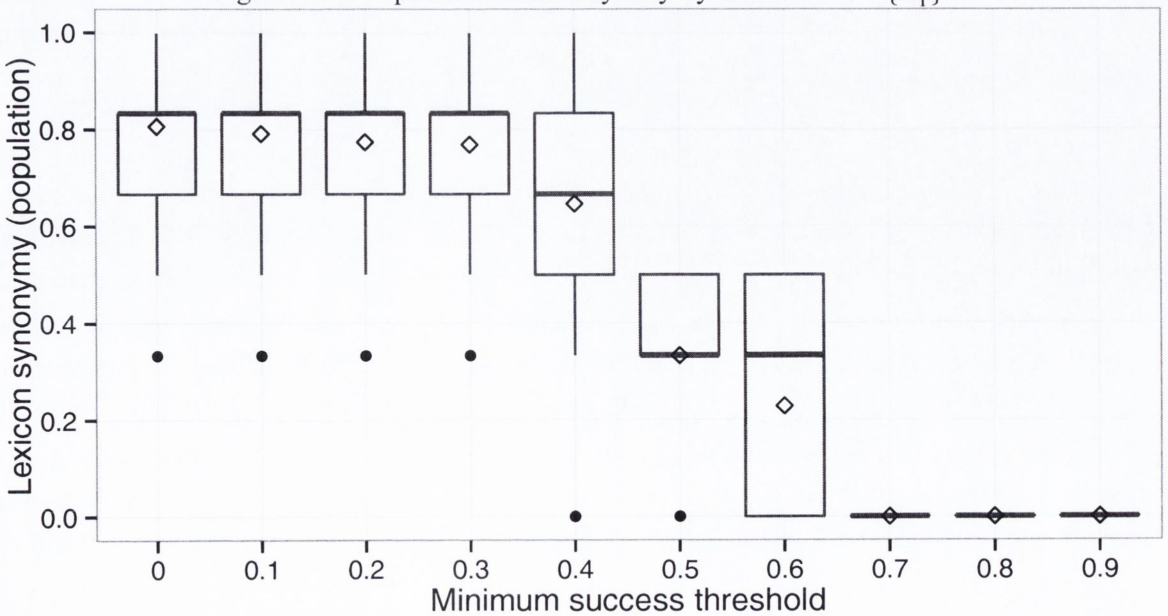


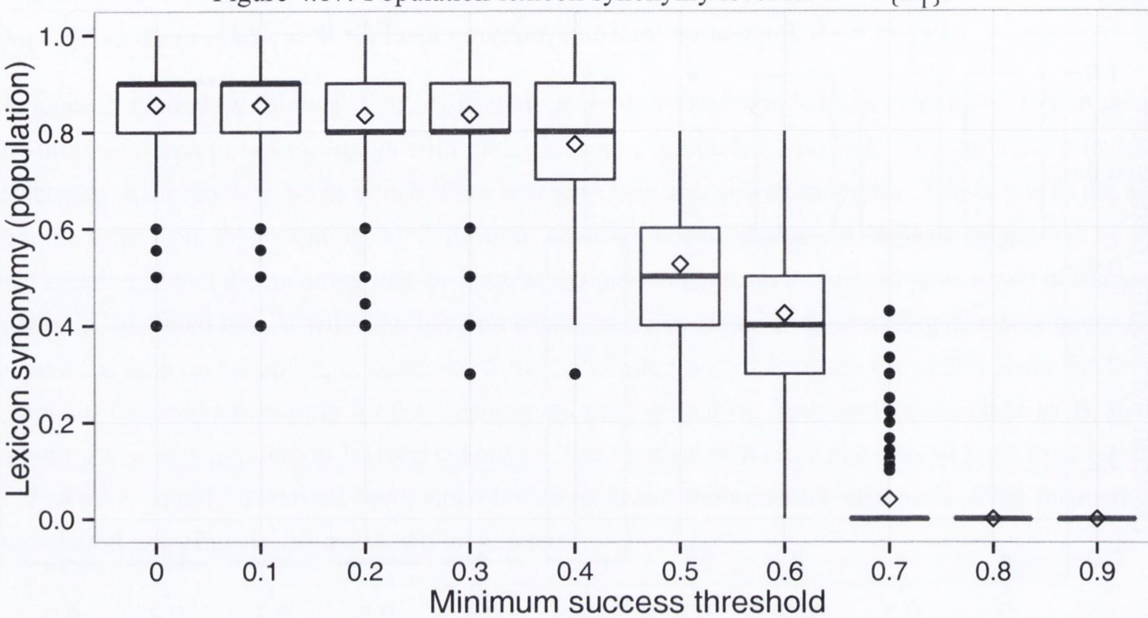
Figure 4.16 shows that in simulations with event R^2, the two agents tend to end up having multiple forms associated with five out of six possible meanings in the majority of simulations with $s_{min} \leq 0.3$. Such high levels of population lexicon synonymy can be clearly attributed to the very low minimum number of first-order shared mappings (2) and, consequently, a relatively high number of distinct forms learnable by the agents (10). For slightly higher levels of minimum success $0.5 \leq s_{min} \leq 0.6$, agents only end up having one meaning with synonymous forms, while for $s_{min} \geq 0.7$ agents naturally have no population synonymy as all mappings are shared between the agents.

Figure 4.16: Population lexicon synonymy level for $W = \{R_1^2\}$.



In experiments with event R^2, the levels of population synonymy vary slightly more between the different levels of minimum success, as seen on figure 4.17. However, the distributions here follow the same underlying principles as above and should thus require no further interpretation.

Figure 4.17: Population lexicon synonymy level for $W = \{R_1^3\}$.



To round up, population synonymy levels are directly dependent on the agent lexicon sizes as are the individual lexicon synonymy levels. The contrast between the two lies in population synonymy growing at two to three times the rate of agent synonymy in a population of two agents, or N+1 in a population of N agents, due to forms from non-shared mappings being by definition assigned to different meanings between the agents. However, population synonymy levels are not as straightforward an indicator of potential communicative success as population homonymy, as it is possible for a particularly successful form referring to a meaning to accumulate a much higher weight than other, less used synonymous forms. In this case, the form will be almost always used by the agents when referring to the said meaning, thus allowing them to successfully communicate about it even with an apparently low overlap between their lexicons. Finally, contrary to the trends from individual agent lexicons, population synonymy levels appear to be at their closest to the ones observed in human languages for $0.5 \leq s_{min} \leq 0.6$.

Agent Homonymy

Homonymy arises when an agent assigns a form to represent multiple meanings. In theory, there is a possibility of such an event happening both during production and interpretation of an utterance. For a homonym to be established during production in the LEW, the speaker agent must first of all have no form associated with a meaning that he wants to express. If this is the case, the speaker will try to invent a new form as a representation for the target meaning, which theoretically might end up being a form previously associated by the agent with a different meaning, thus resulting in a homonymy relationship being established. However, the current implementation of the LEW assumes agents to be able to keep track of the forms that they have in their lexicon, in effect forcing them to actually come up with a completely new form when resorting to invention and subsequently avoiding producing a homonym during utterance production.

The second way of introducing homonyms to the lexicon of an agent is by assigning an alternative meaning to an already known form during the interpretation of a heard utterance. However, the configuration of the LEW model specifies that if an agent has a meaning associated with a known form, it will select that meaning as the corresponding interpretation, thus eliminating the possibility of a homonym being introduced to its lexicon during interpretation.[5] In summary, agent homonymy is currently inadmissible in a population of two agents talking to each other.

Population Homonymy

It has been shown in section 4.2.1 above that the set of distinct forms will always be equivalent in the lexicons of agents in a population of size two. Accordingly, determining the population lexicon's homonymy levels can be reduced to estimating the maximum number of forms that could potentially end up being associated with different meanings between the two agents. By definition, the number of such globally homonymous forms will be equivalent to the number of distinct forms with non-shared meaning associations between the two agents, i.e. the total number of unique forms in the population lexicon minus the shared ones from first-order and second-order shared mappings. Consequently, the population homonymy level in a population of two agents is defined as follows:

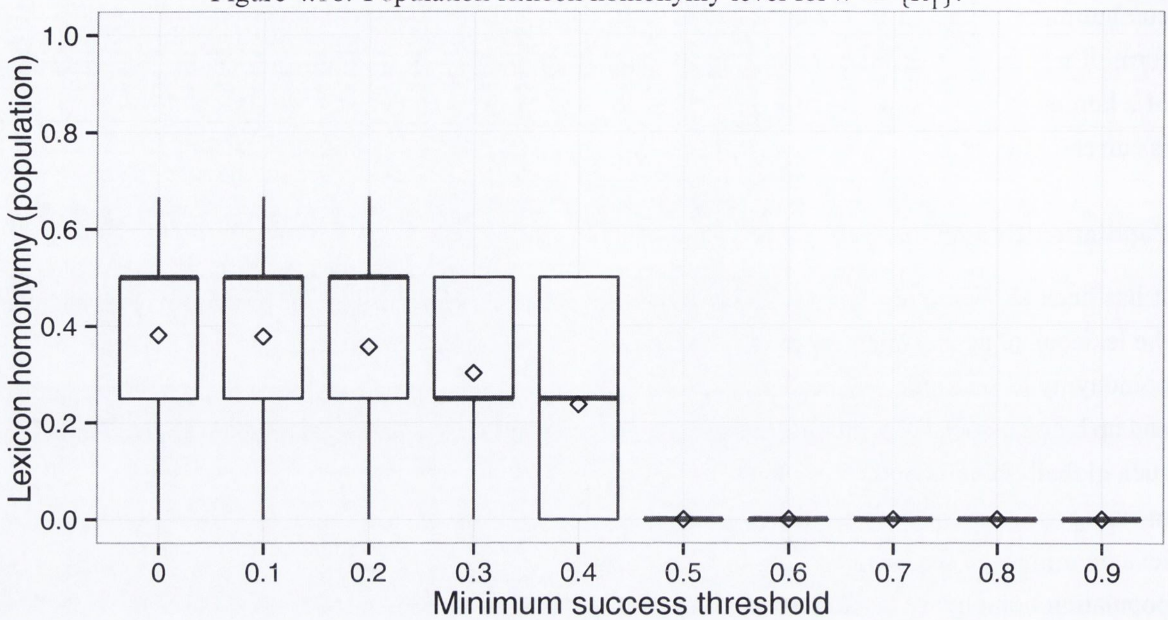
$$
\begin{aligned}
Hom^{pop} &= \frac{|F^{pop}_{unique}| - |M'_{shared}| - |M''_{shared}|}{|F^{pop}_{unique}|} \\
&= \frac{|M'_{shared}| + (|M| - |M'_{shared}|) * 2 - |M'_{shared}| - |M''_{shared}|}{|M'_{shared}| + (|M| - |M'_{shared}|) * 2} \\
&= \frac{(|M| - |M'_{shared}|) * 2 - |M''_{shared}|}{|M| * 2 - |M'_{shared}|}
\end{aligned}
\tag{4.14}
$$

As can be seen, for a given set of events, i.e. a fixed $|M|$, the level of population homonymy is inversely proportional to the number of shared items and directly proportional to the overall size of the lexicon. Note that while the equation for the level of population homonymy is similar to that of the *maximum* level of population synonymy, the former value will always be defined by the equation, given $|M'_{shared}|$, $|M''_{shared}|$ and $|M|$, while population synonymy will vary from simulation to simulation between the minimum and maximum levels.

Figure 4.18 depicts that the majority of agents talking about an event of type R^1 and $s_{min} \leq 0.4$ who have one meaning without a shared form, i.e. who have four forms in their corresponding lexicons, end up associating the 'extra' forms with different meanings between themselves, resulting in an average population lexical homonymy level between 0.24 and 0.38. For agents with shared mappings only, i.e. for $s_{min} \geq 0.5$, the population lexicon homonymy level is naturally equal to zero.

[5] Remember that, in the present configuration, the agent will not mind if the meaning was not actually present in the observed event (i.e. is not part of any of the possible event partitions).

Figure 4.18: Population lexicon homonymy level for $W = \{R_1^1\}$.



The following figures 4.19 and 4.20 show the levels of population homonymy levels for experiments conducted with event R^2 and R^3 respectively. These follow the same probabilistic distribution depending on the number of shared mappings, the total number of forms acquired by either agent and the chances of the agents not arriving at the correct meaning of a form the first time around (yet still obtaining a payoff due to low level of minimum success) but actually inferring the correct meaning when the speaker-hearer roles are reserved.

Figure 4.19: Population lexicon homonymy level for $W = \{R_1^2\}$.

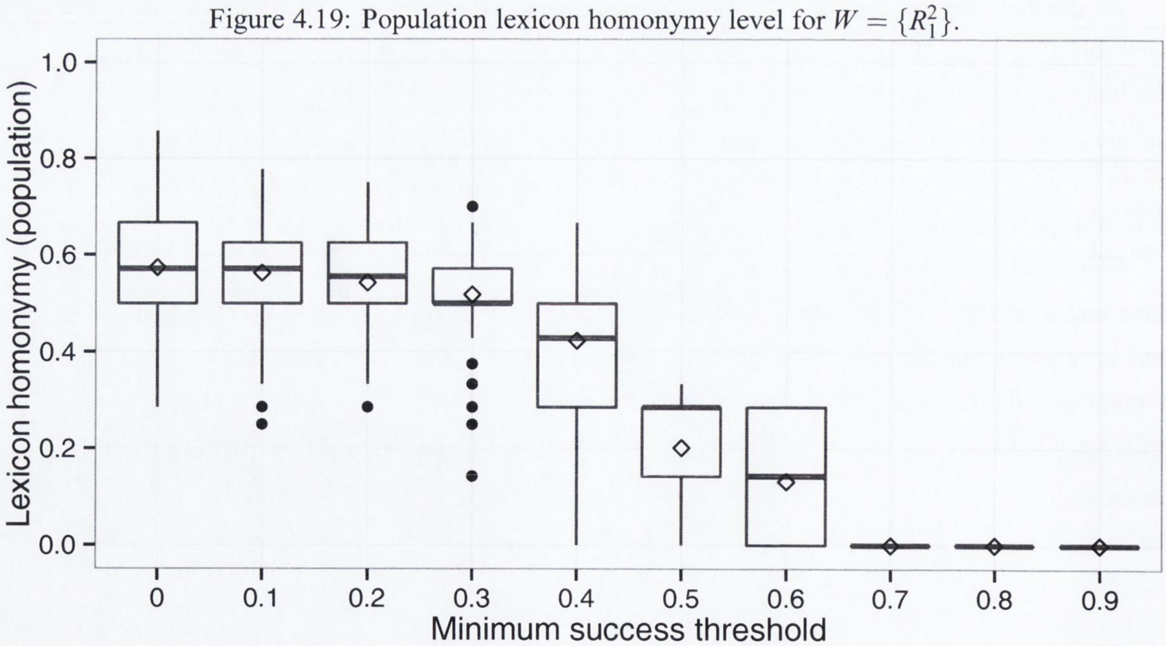
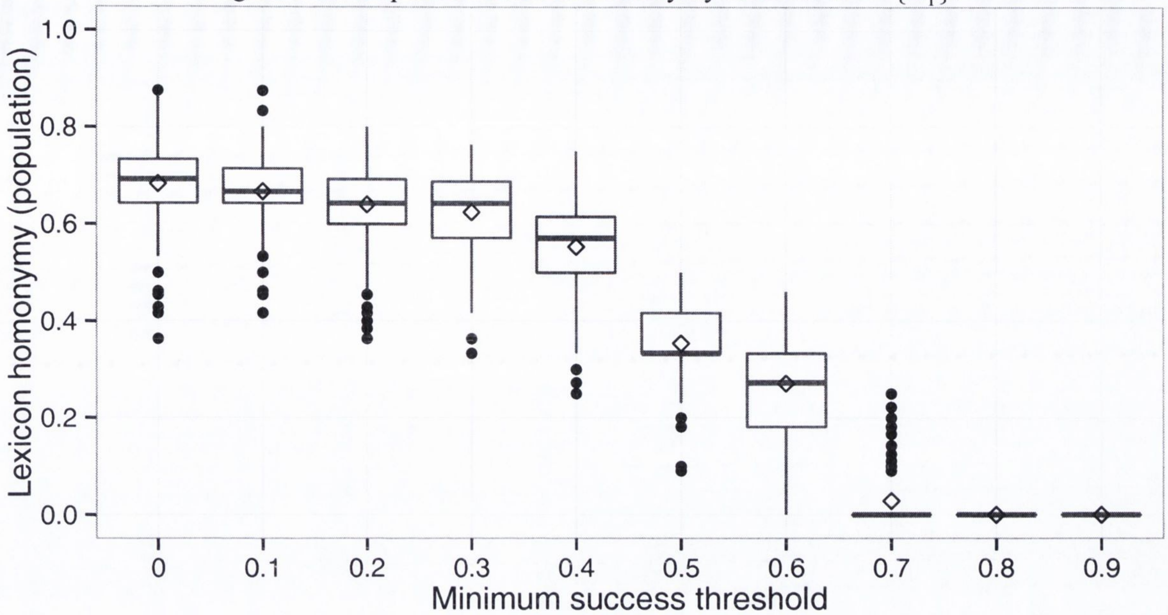


Figure 4.20: Population lexicon homonymy level for $W = \{R_1^3\}$.



The three figures above provide a confirmation that population homonymy levels are very closely and naturally tied to the size of the agents' lexicons and in effect depict the overlap between the lexicons of the different agents. In effect, the population homonymy levels indicate the chances of a form being misinterpreted in an interaction between the two agents (disregarding the different weightings) due to the fact that it is assigned to different meanings in the corresponding lexicons. Finally, just as was the case for population synonymy, homonymy levels in the population lexicons are closest to real language values for $0.5 \leq s_{min} \leq 0.6$, indicating a very strong correlation in the early simulations already.

4.2.3 Communicative Success

When communicating with one another, agents clearly have two options available to them when dealing with either a meaning that needs to be expressed or a form that needs to be interpreted. The first option is to make use of one's lexicon, i.e. a mapping in the lexicon with the corresponding meaning or form. If that option is unavailable, the agents will have no other choice than resort to either inventing a new form for the given meaning, or guessing the meaning of the encountered form. Consequently, the overall communicative success Und of agents will be a combination of the accuracy of their lexicon $LPrec$, when an appropriate mapping could be found in their lexicon, and the chances of correctly guessing the meaning of a form P_{guess}, when no useful mapping was deemed to be found in the lexicon. In the following equation, $LUse$ represents the proportion of individual form interpretations that hearers were able to attempt with the help of their lexicon.

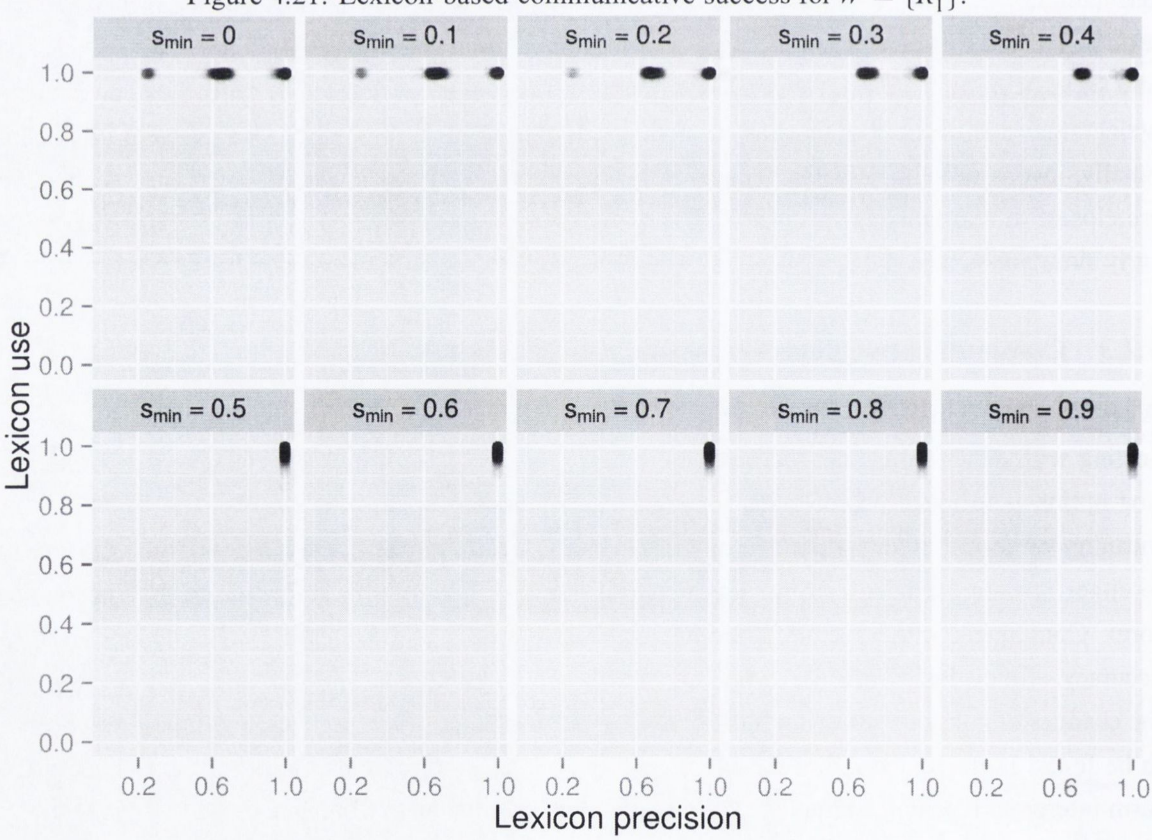
$$Und = LPrec * LUse + P_{guess} * (1 - LUse) \tag{4.15}$$

In the above equation, P_{guess} will be a constant for a given set of events, since the chances of selecting any particular meaning from all possible event partitions do not depend on any property of the agents. In effect, P_{guess} represents the expected level of communicative success that agents in the LEW would achieve if they never made use of their lexicon, i.e. if they never remembered any meaning-form associations even if they obtained positive feedback and thus always resorted to guessing the meanings in all subsequent interactions.

The two measures that can and will vary significantly based on the agents' capability of detecting, memorizing and reusing correct interpretations of utterances are *LUse* and *LPrec*. Logically, the level of lexicon use can only be high if the agents know at least as many forms as there are possible meanings in the world. However, being able to locate every possible form used by other agents does not guarantee that a hearer agent will correctly interpret it. The observed probability of agents correctly interpreting forms with the help of their lexicons is represented by the *LPrec* measure. Note that, a high level of *LPrec* alone can also be quite insignificant, e.g. in case where agents actually almost never use their lexicon. Ultimately, agents should be ranking high in both these measures, if they were to be considered successful communicators.

A final note that should be made at this point is that, in a configuration where agents do not make mistakes when segmenting each other's utterances into words, the composition and thus the length of both the spoken on the heard utterance will be exactly the same. Since precision and recall measures only differ in the relation to the size of either the source or the target utterance, if both these are of the same length, then precision and recall will also be equivalent for a given hearer's interpretation of the speaker's utterance.

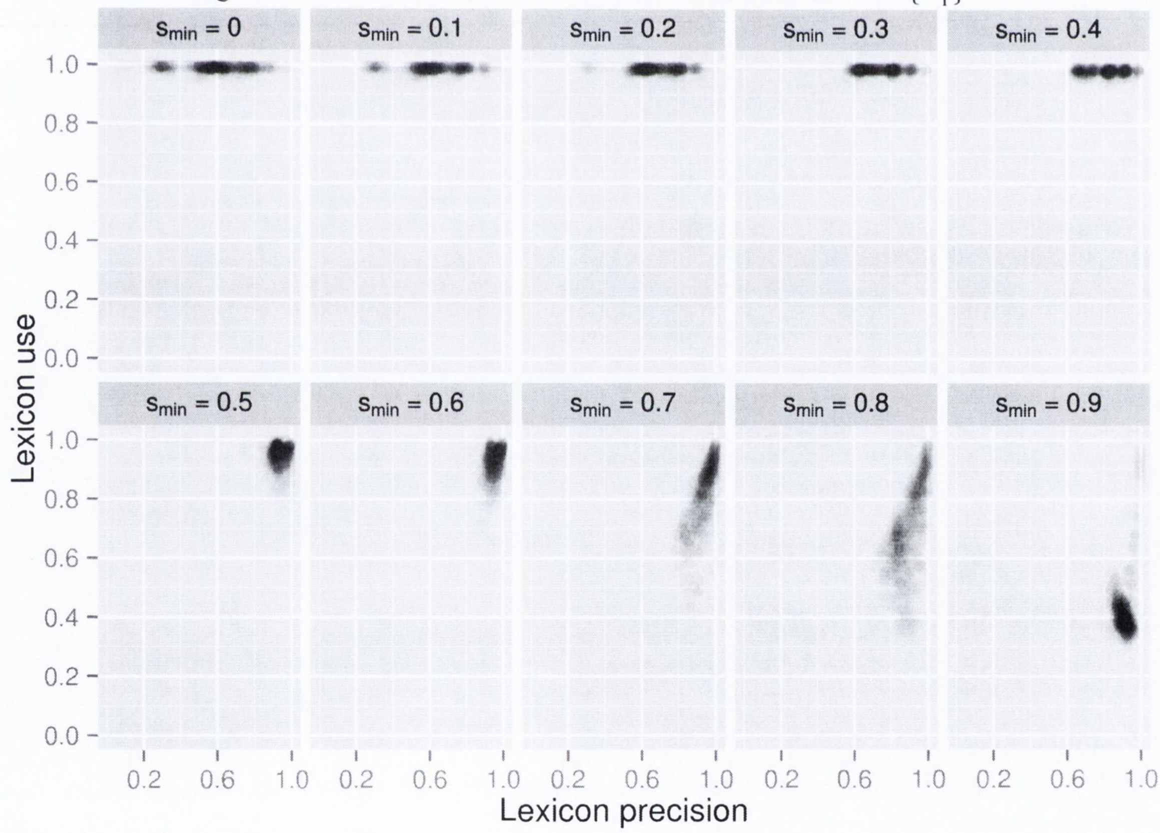
Figure 4.21: Lexicon-based communicative success for $W = \{R_1^1\}$.



The above figure 4.21 depicts the interplay between lexicon use and lexicon precision in simulations performed with two agents and one R^1 event. As can be seen, there is a striking distinction between simulations with $s_{min} \leq 0.4$ and those with $s_{min} \geq 0.5$. In simulations with the lower minimum success requirement, agents appear to quickly learn the maximum possible number of forms and then always use these in future interactions, though admittedly with varying success. In simulations with $s_{min} \geq 0.5$, on the other hand, agents are only ever able to learn mappings that are guaranteed to be correct[6], which results in their lexicon precision being fixed at 100% once the mappings have been learned. The small tail of lexicon use below 1.0 is a reflection of the failed interactions that it took the agents to get to that point.

Looking further into the simulations with $s_{min} \leq 0.4$, it appears that there are three maxima that the agents can end up in, with lexicon precision at around 0.25, 0.66 and 1.0, respectively. The latter maximum is obviously the desired global maximum, while the former two are local maxima. In order to reach the global maximum, agents need to make sure that they either fully understand or misunderstand each other in interactions. While the above might seem illogical, especially the part about misunderstanding, the rationale behind this requirement is that, if the agents finish an interaction with partial success, an incorrect mapping will be inserted into the lexicons of both agents. However, given the very low level of s_{min} in these simulations, these mappings might never get eliminated,[7] thus keeping the agents stuck in local maxima with limited communicative success. On the other hand, if agents do not experience partial understanding, it might take them longer to learn the correct mappings, but at least they will not insert any incorrect ones into their lexicons. Finally, the distinction between the two local maxima lies in the number of 'extra' forms in the population lexicon, which can vary between one and two – hence the two local maxima.

Figure 4.22: Lexicon-based communicative success for $W = \{R_1^2\}$.



In simulations with R^2 and R^3 events, as depicted by figures 4.22 and 4.23, respectively, the situation is quite similar for $s_{min} \leq 0.4$. However, given the more dynamic nature of payoffs in interactions with a larger possible number of components, agents seem to be taking up a range of different values, with the average lexicon precision level shifting from around 0.59 to around 0.81 in

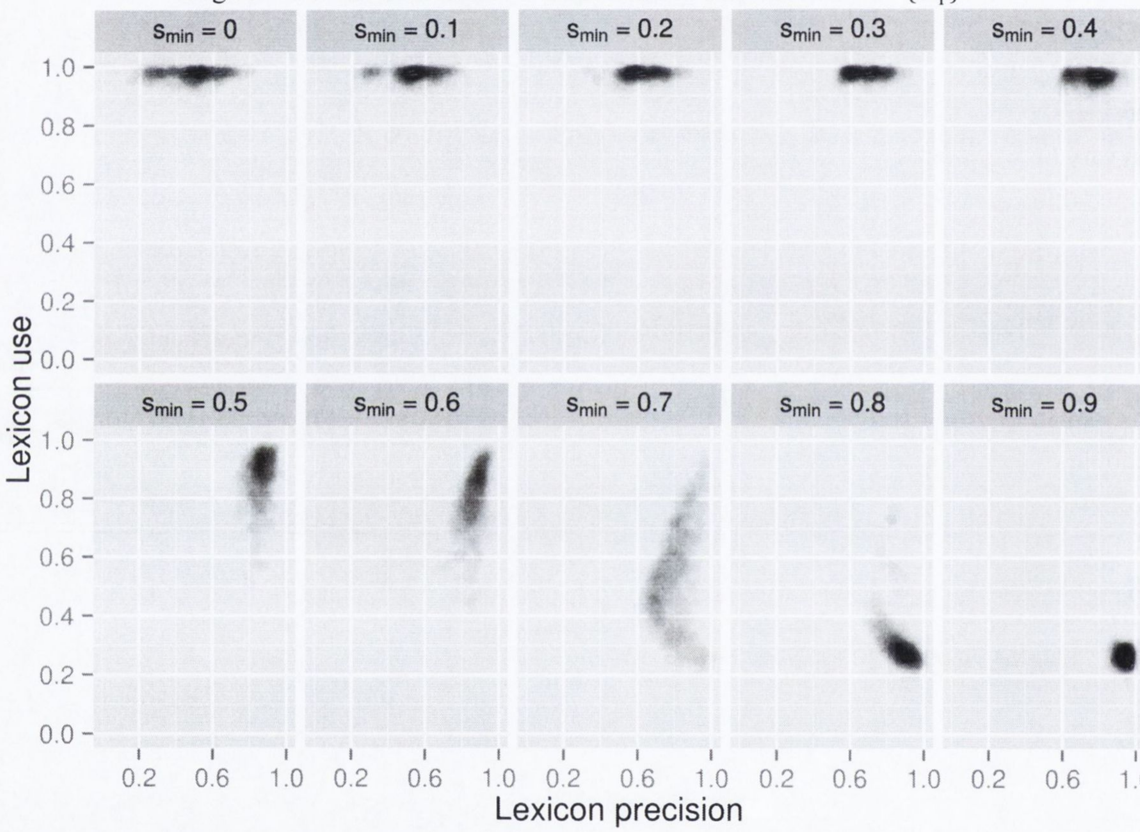
[6]This is due to utterances in simulations with one R^1 event having a maximum of two components, meaning that both of these need to be interpreted correctly for a success level above 0.5, which would warrant a payoff.

[7]In fact, for $s_{min} = 0$, agents will never eliminate any mappings, as there are no punishments in this scenario.

experiments with an R^2 event and from 0.49 to 0.75 in simulations with an R^3 event.

For both event types, simulations with $0.5 \leq s_{min} \leq 0.6$ exhibit the largest number of agents finishing closest to the intersection between constant lexicon use and perfect lexicon precision. As the level of minimum success is increased further, however, one can observe a dramatic drop in lexicon use all the way down to the 40% mark for R^2 and even lower to around 25% for R^3. Notably, the falloff from high lexicon use rate to a very moderate one is accompanied by only a very slight increase in lexicon precision, clearly suggesting that there is an optimal level of s_{min} somewhere between 0.4 and 0.6.

Figure 4.23: Lexicon-based communicative success for $W = \{R_1^3\}$.



In summary, it can be said that the bottom five levels of minimum success, i.e. $s_{min} \leq 0.4$, do not guarantee correct interpretations in any of the three event conditions, even though the lexicon is almost always used in these cases, thus fully eliminating the guessing element. For top five levels of minimum success, the lexicon is normally a very reliable reference point, with the exception of some transitional phases observed for the two bigger event types. However, the possibility of using one's lexicon in an interaction is significantly reduced in simulations with the highest levels of minimum success threshold and the two bigger events. In this case, one could argue that not being able to use one's lexicon for the majority of one's communicative needs does not justify for trying to acquire such a lexicon in the first place. Accordingly, the level of minimum success with the most potential is arguable $s_{min} = 0.5$, which exhibits both extremely high chances of lexicon use, as well as quite reliable precision across all experiments.

4.3 Interacting in Triads

Having gone thoroughly over the lexicon acquisition task in populations of exactly two agents, this section will evaluate the effects of adding just one additional agent on the overall dynamics of the population. In particular, in a population of three or more agents, any agent can interact with any other agent, with both the speaker and the hearer roles being assigned by random. As in the section on pairs of agents, the analysis will begin with the estimation of the limits on the size of the acquired lexicon, following with an evaluation of lexical synonymy and homonymy and concluding with an analysis of communicative success measures.

4.3.1 Lexicon

Lexical Mappings

If the simulated population consists of more than two agents, then the limit of the maximum number of lexical items learnable within the population (and without intermediate mapping elimination) will grow accordingly for a given set of events and minimum success threshold. As was the case for a group of only two agents, the largest overall population lexicon is reached under the condition that as few mappings are shared across agents as possible. In a population of over two agents, this condition applies not only to the number of mappings shared between any two agents, but also to the number of agents sharing any particular mapping.

To illustrate the above, imagine a population of multiple agents sharing some set of mappings required to associate a form with every meaning in the world. In such a scenario, the maximum lexicon size would increase by the number of meanings without a shared mapping for each additional agent. However, every agent in the population need not share the same set of mappings with every other agent in the population in order to be able to learn a form for every possible meaning in the world. In fact, it is possible that, if acting as a speaker first, every additional agent will agree on a shared set of lexical items required for him to learn the whole lexicon that is completely new to the population.[8] Note that it is not possible for agents to evolve a local lexicon within every possible pairing, as once they have associated at least one form with every meaning, they would always use these forms in the future, unless deemed unhelpful and eliminated from their lexicon.

In the case where mappings can be eliminated after unsuccessful interactions have been punished, the lexicon size dynamics in groups of over two agents differ quite dramatically from what was observed in a two-agent population. As has been described in section 4.2.1, two agents who always correctly interpret each other's words will have their mappings rewarded and punished evenly, which means that if a new mapping is learned, it is learned by both these agents, and if a mapping is eliminated, it is eliminated from the lexicon of both agents simultaneously. The most important consequence of this lexical synchronicity between the two agents is that it is impossible for one agent to run out of mappings for a particular meaning and invent a new one without the other agent also running out of the same mappings, which in turn results in a form being completely wiped from the population lexicon before a new lexical item can be learned, thus eliminating the possibility of learning additional lexical items.

[8] The reason why the third agent would need to act as a speaker first, in order to add to the lexicon of the two agents who are already using some (partially shared) lexicon is that only an agent without any mappings can invent new forms. Accordingly, if the first two agents started talking to the third one, he could only end up learning lexical items that were already in one of the other two agents' lexicons, thus not introducing any new shared mappings to the population lexicon.

Since pairs of interlocutors in a group of more than two agents will alternate between interactions, it no longer holds that mapping weights are constant across the whole population. Accordingly, when a mapping is used in an interaction conducted by two agents, it will only be rewarded or punished in the lexicons of the two agents involved in the interaction, meaning that other agents might end up with a different weight assigned to the mapping, or end up not knowing forms that others do or vice versa.

For the estimation of the limit on lexicon size, this property of multi-agent group interactions has a quite drastic consequence. What becomes possible in a group with at least three agents is that one of the agents acts as an inventor and the other two either always understand or always misunderstand him, respectively. In such a scenario, when the inventor agent speaks to the agent who understands him, a new mapping will be introduced to the lexicon of both these agents. However, the inventor agent now has the option of talking to the misunderstanding agent and thus eliminate lexical items from his lexicon without them also disappearing from the population lexicon (the understanding agent will hold on to these, at least as long as he hasn't misused them himself, which requires for him to actually act as a speaker at some point). Having forgotten all of the lexical mappings that the inventor agent originally had associated with a meaning, he can now go ahead and invent a new mapping again and have it reinforced by the understanding agent.[9] There are no bounds on this loop, which means that there is also no upper limit on the lexicon size of the population of over three interacting agents.

The above scenario implies that one of the agents needs to be constantly speaking, and the other two only act as hearers, for the lexicon to grow without a limit. However, the same can also happen even if the agents were forced to take turns in speaking/listening, as long as all three agents don't agree on too many shared lexical items, thus allowing them to keep on inventing new mappings. The question if the agents actually end up agreeing on a partially shared lexicon and stop inventing new forms after a certain period of time, is partially answered by simulation, although since there is no exact number of interactions that would guarantee agents ending up in a local or a global maximum, the simulation results only show the probability distribution after the corresponding number of interactions.

Population Lexicon

Having gone through the abstract technical analysis of limit cases, it is now time to evaluate the results of the corresponding simulations and see whether they perform within expectations, starting with the overall lexicon size of the simulated populations. As has been discussed above, the number of distinct forms and thus the number of mappings in an agent's lexicon can grow infinitely in a population of three or more agents. The following graphs should show if and to what extent that is allowed to happen within a simulation consisting of 1500 interactions, i.e. 1000 interactions per agent.

[9] In a way, the described scenario is not dissimilar for the parent-child vertical transmission model utilised by Kirby (2002) and Smith (2002) among others.

Figure 4.24: Number of lexical items in agent lexicons for $W = \{R_1^1\}$.

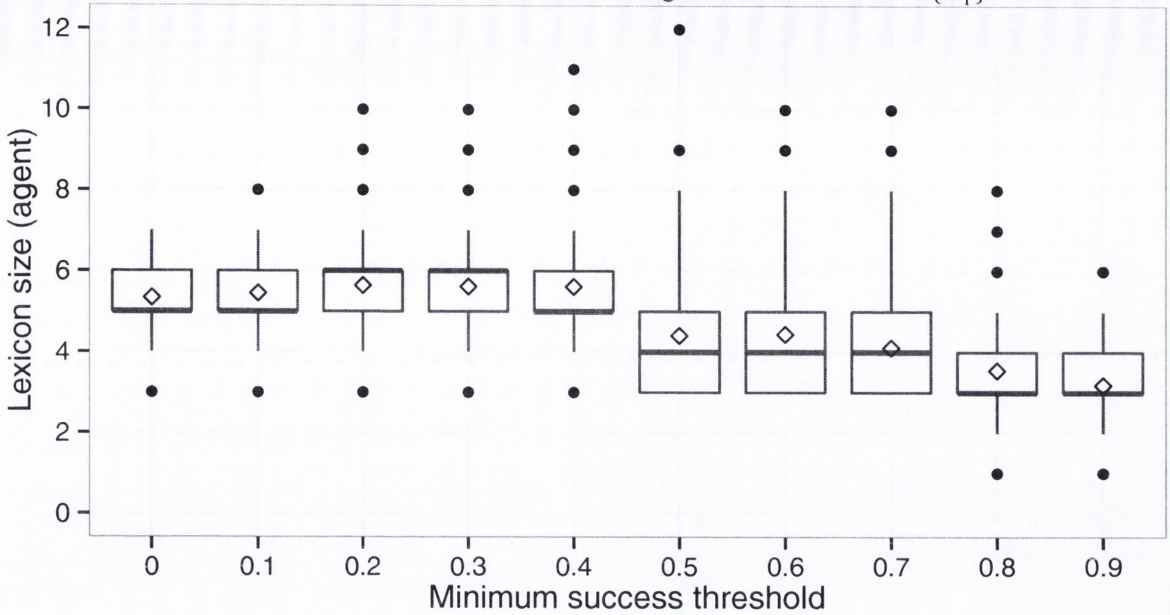
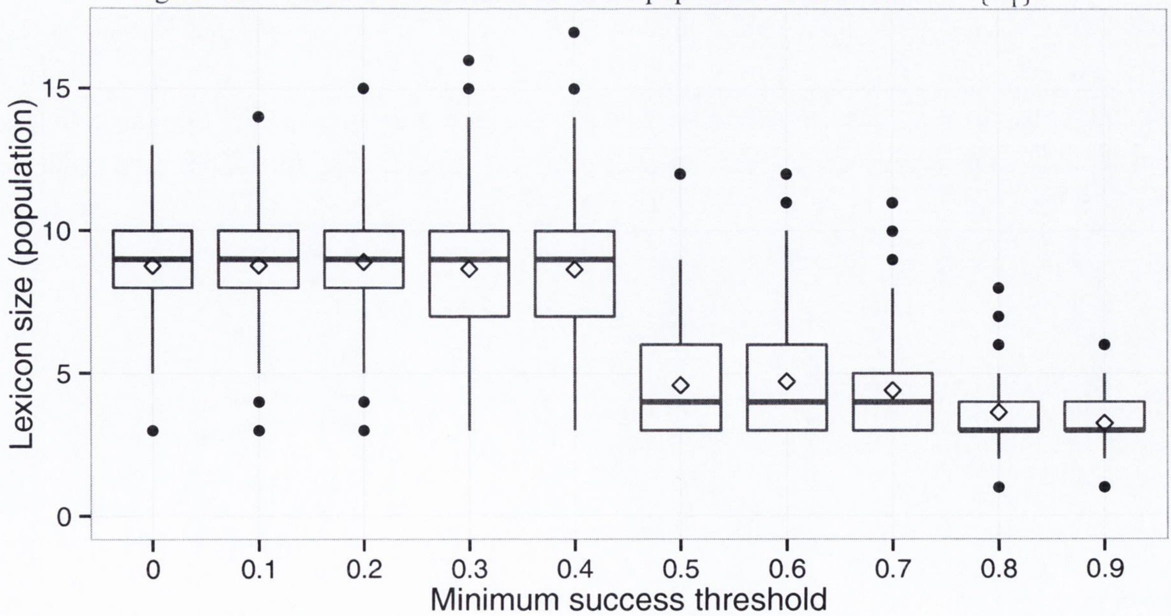


Figure 4.24 plots the distribution of the number of lexical mappings accumulated by the individual agents' lexicons in the conducted simulations in groups of three. The first observation that can be made straight away is that for $s_{min} \leq 0.7$ the majority of the agents, and thus also the populations they inhabit, end up utilizing more than just the three lexical items required to express all of the world's meanings. In fact, for the mid-range levels of minimum success, the number of learned words reaches as much as four times the amount of distinct meanings.

Figure 4.25: Number of lexical items in the population lexicon for $W = \{R_1^1\}$.



Furthermore, figure 4.25 shows that, indeed, the size of the population lexicon is significantly higher than what was observed in a population of two agents, reaching on average nine mappings for $s_{min} \leq 0.5$, four for $0.6 \leq s_{min} \leq 0.7$ and only settling at the required level of three for $s_{min} \geq 0.8$. In particular, the lexicon size of the population in a world with just three meanings reaches as many as 17 lexical items in some simulations, which clearly indicates that some amount of looping, in sense described in the preceding paragraph, occurred between the agents in these simulations. Additionally, it should be noted that for all levels of minimum success except the highest two, the population lexicon always includes at least three forms, i.e. the minimum number required to be able to express all meanings.[10]

Figure 4.26: Number of lexical items in agent lexicons for $W = \{R_1^2\}$.

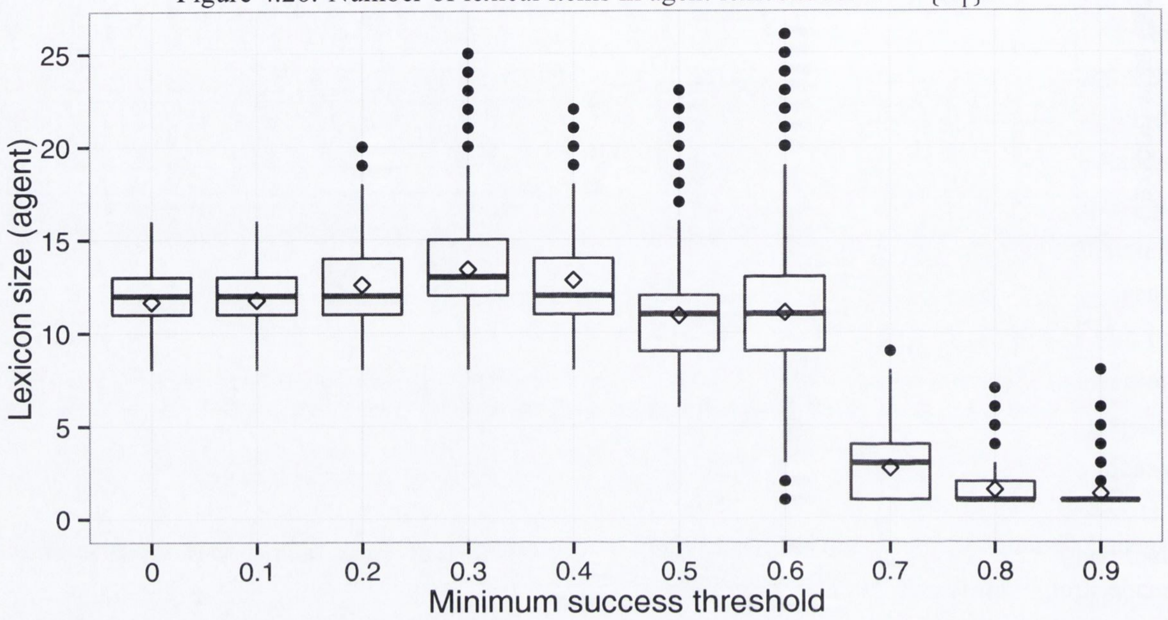
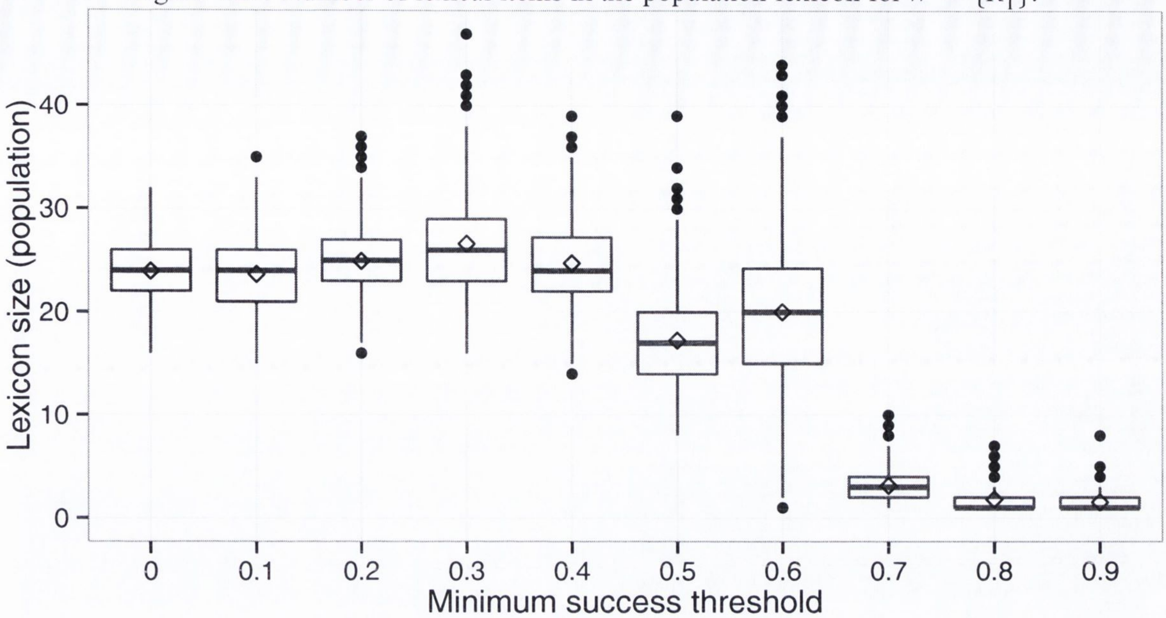


Figure 4.26 suggests that increasing the size of the topic event to include two arguments does not seem to significantly influence agents adhering to a level of minimum success $s_{min} \leq 0.6$ who end up consistently learning at least double the amount of lexical items required to express the six different meanings available to them, with their lexicons extending up to 26 different lexical mappings, i.e. over four times the necessary amount. However, for $s_{min} \geq 0.7$, i.e. where interpreting one of two or two of three forms from an utterance is no longer sufficient for a positive payoff, agents actually struggle to learn and agree on enough mappings to be able to express every possible meaning. In fact, for $s_{min} \geq 0.8$, the majority of the agent's lexicons only contain one lexical item, associated with the partition consisting of all of the event's components.

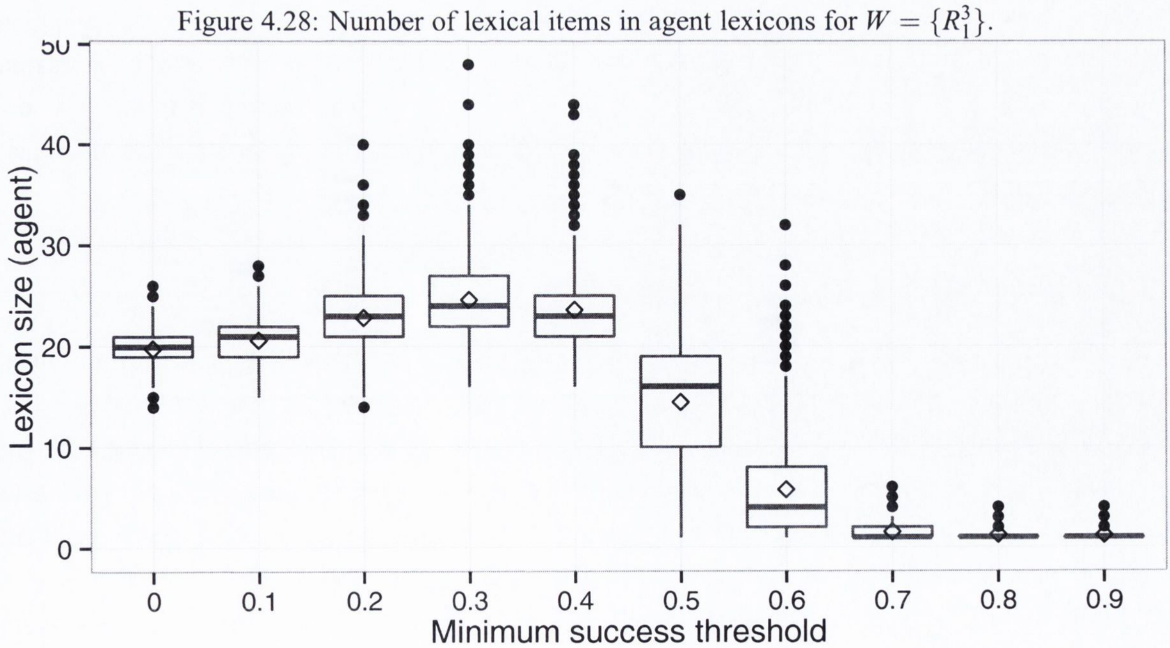
[10]Note that this does not necessarily imply that every (or any) agent knows how to express each of the meanings, only that there are three distinct mappings in the group language, which could possibly all refer to the same meaning and only be known by one of the agents. The discussion of unique forms and meanings in agent and population lexicons will provide more insights into the question if this is indeed the case.

Figure 4.27: Number of lexical items in the population lexicon for $W = \{R_1^2\}$.



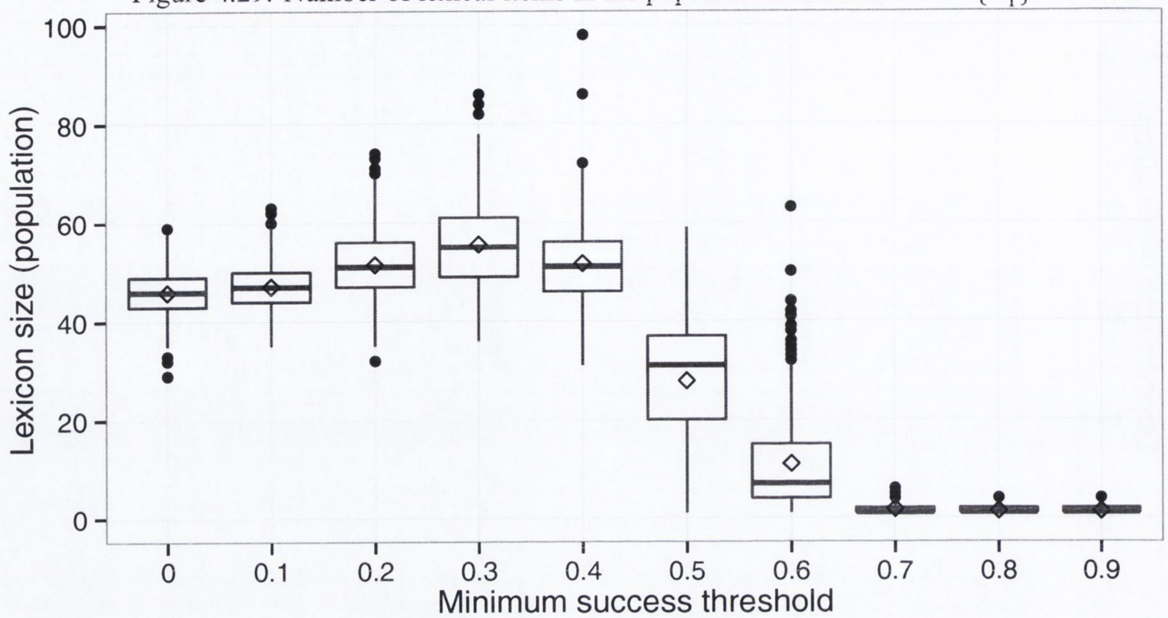
The size of the population lexicon grows even further in simulations performed with an R^2 and an R^3 event, as depicted by figures 4.27 and 4.29, respectively. In simulations with an R^2 event, the population lexicon has over 10 different mappings for $s_{min} \leq 0.4$, nearly reaching 50 in some simulations, while in experiments with an R^2 event, the overall lexicon has over 30 entries for $s_{min} \leq 0.4$, going up to as high as 100 in some cases. In both these experiments, the population lexicon is significantly smaller for $s_{min} \geq 0.7$, in most cases actually just containing a single lexical mapping.

Figure 4.28: Number of lexical items in agent lexicons for $W = \{R_1^3\}$.



In simulations with an R^3 event, the number of distinct mappings in agents' individual lexicons follows the same tendencies as those observed in preceding simulations. Figure 4.28 shows that for $s_{min} \leq 0.5$, agents still mostly acquire double the number of words (10) needed to express all meanings, with the lexicons of some agents having close to 50 lexical items. On the contrary, for $s_{min} \geq 0.7$, agents again fail to learn a sufficient number of words.

Figure 4.29: Number of lexical items in the population lexicon for $W = \{R_1^3\}$.



The general decrease in the population lexicon size for the higher levels of minimum success threshold can be clearly explained by the need to experience nearly perfect success in order for a payoff to be obtained, which would allow a lexical mapping to be stored in the agents' lexicons. However, even if the two of the agents experience full success on occasion, there is still a high chance of one of the agents trying to utilize the previously successful mappings with the third agent who might not share these, which will often result in failure, and thus the elimination of the mappings even from the lexicon of the agent who had used them successfully in the past. The high level of minimum success also results in severe punishments, which can basically wipe out a lexical mapping from the lexicons of agents after one failed interaction, even if it had been used with success on numerous occasions in preceding interactions. All of the above clearly contributes to the very low number of mappings in the population and, as will be seen in the following sections, the individual agents' lexicons.

Unique Meanings

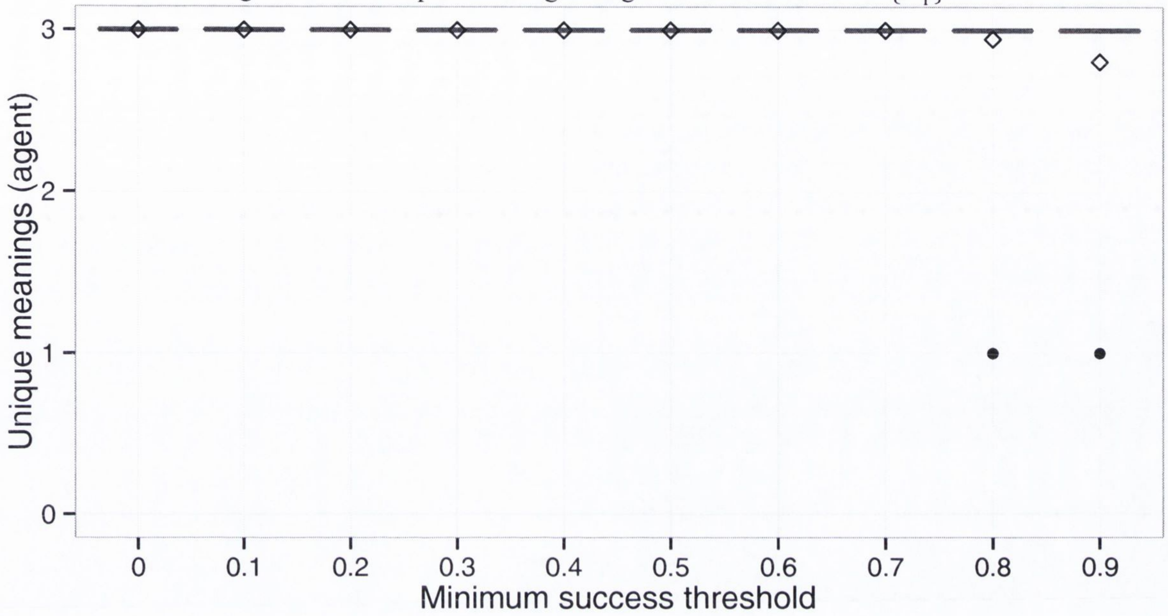
The effects described in the preceding section have no effect on the maximum number of meanings that three or more agents can end up learning in a system with a fixed number of non-recursive events. Regardless of how good or bad the agents may be at agreeing on forms that correspond to the different meanings, there is still no way for the agents to invent new meanings beyond the ones extractable from the given set of world events. Accordingly, the maximum number of unique meanings in agent lexicons remains fixed at the value defined by equation 4.6.[11]

Similarly to the experiments conducted with two agents, a world with one single-argument event does not pose a significant challenge to the agents even in a population of three. As can be seen from figure 4.30 below, the participants almost always end up knowing how to express every possible meaning in the world after 500 epochs of three interactions each, i.e. after participating in 1000 interactions on average. The level of minimum success imposed on the agent conversations appears

[11] As has been admitted before, the limited meaning space is clearly a quite unnatural restriction imposed on the agents. However, this restriction is necessary when conducting an analysis of the model's fundamental dynamics.

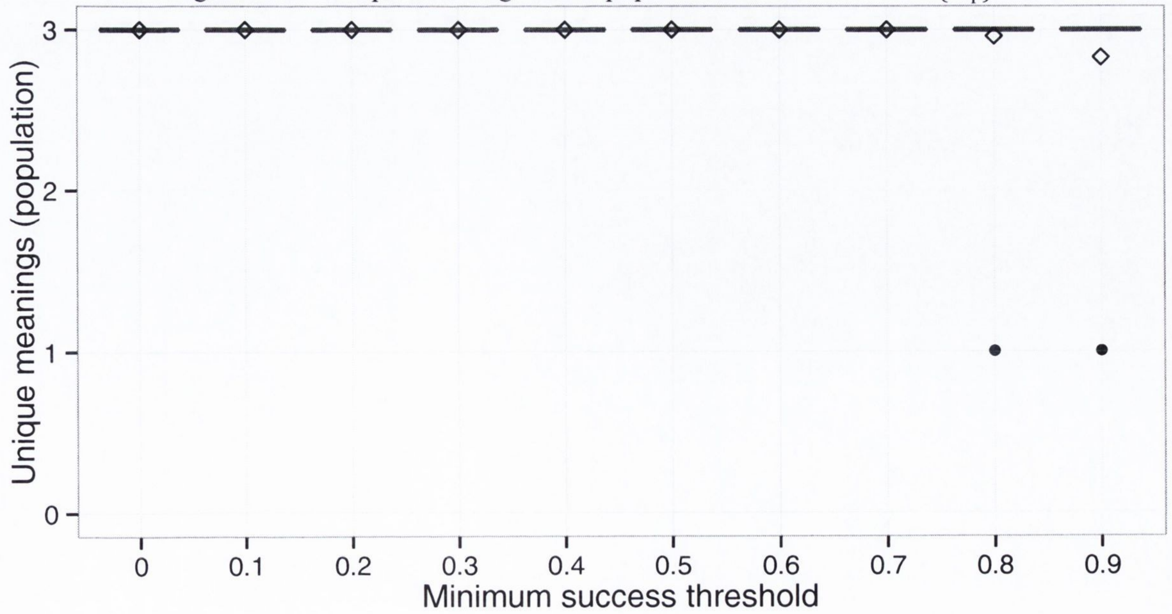
inconsequential in this simple world, with only a handful of individual cases of agents not being able to acquire a form for one of the meanings, observed for the two highest level of minimum success threshold.

Figure 4.30: Unique meanings in agent lexicons for $W = \{R_1^1\}$.



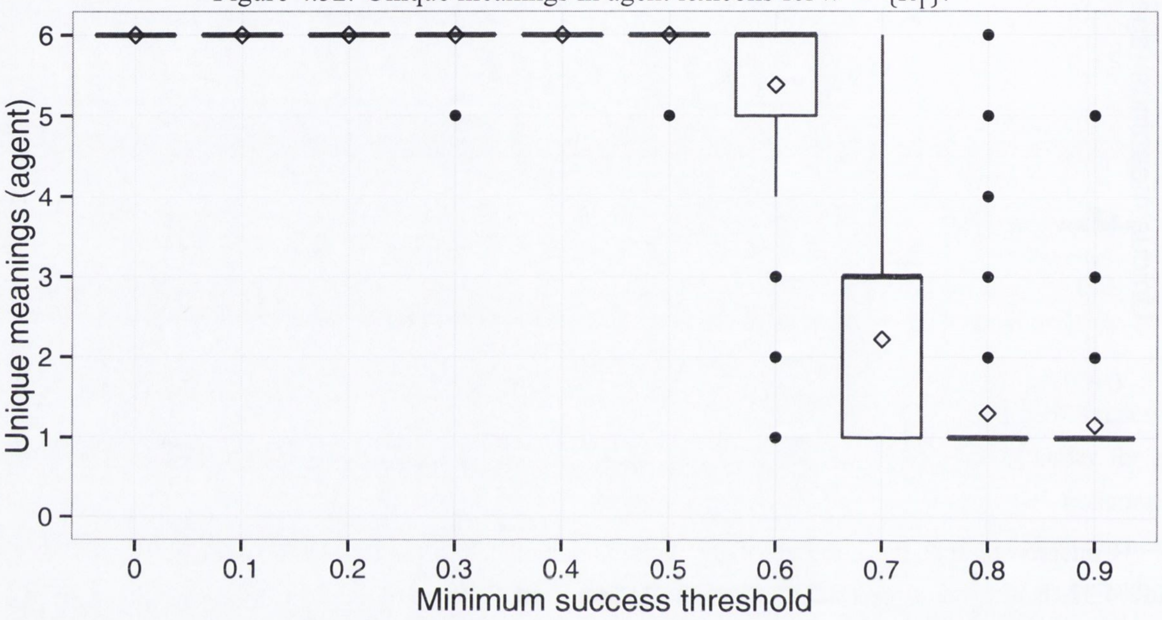
Considering that the number of distinct meanings in the population lexicon, depicted by figure 4.31, has to be at least as high as in individual agent lexicons, the only question that remained open here was if, in the few cases where the agents were able to learn a word for only one of the forms, the same applied to both agents or not. The answer is clearly that in those few cases, it was both of the agents who got stuck at the one meaning and could not agree on a form for the remaining two.

Figure 4.31: Unique meanings in the population lexicon for $W = \{R_1^1\}$.



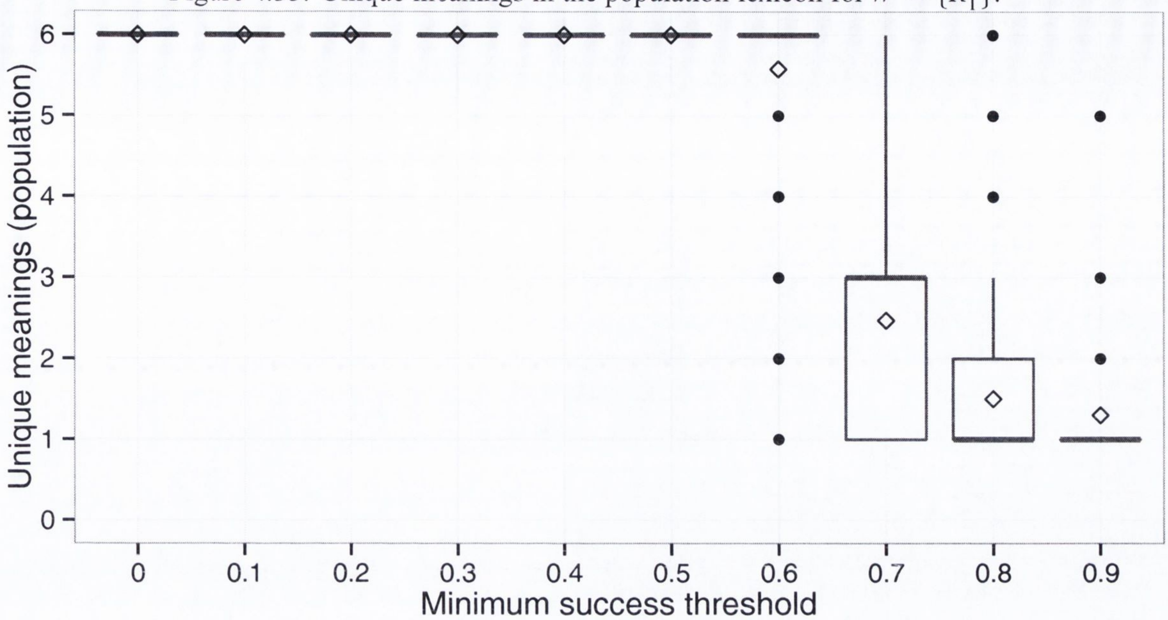
Increasing the number of arguments allowed in the single world event to two has a deteriorating effect similar to the one observed for two agents in figure 4.8. In particular, as depicted by figure 4.8, the three agents keep on reliably learning a lexical item for all six possible meanings up to $s_{min} \leq 0.5$, two tenths of the level lower than what was observed for two agents. For $s_{min} = 0.6$, agents still predominantly learn to express all meanings, though not reliably so. For $s_{min} = 0.7$, the majority of the agents only manages to associate a form with half of the available meanings, while simulations performed with $s_{min} \geq 0.8$ result overwhelmingly with just one meaning being a part of the agents' lexicon after 1000 interactions per agent.

Figure 4.32: Unique meanings in agent lexicons for $W = \{R_1^2\}$.



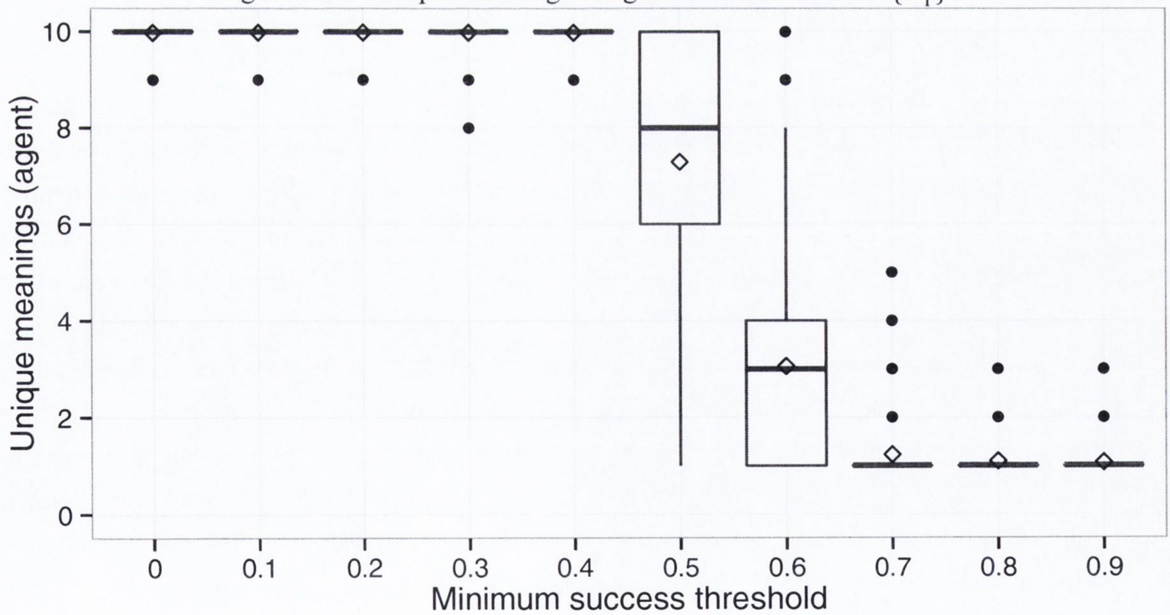
When observing the number of unique meanings in the population lexicon, as plotted by figure 4.33, it appears that the few agents that could only learn a form for five of the meanings for $s_{min} \leq 0.5$ always had at least one member in the population who knew a word for all six of the meanings. A similar tendency if observed for the higher levels of minimum success where the number of distinct meanings in the population lexicon is often slightly higher than in the individual agent lexicons. What is observed here is a consequence of one pair of agents having agreed on certain mappings, but not being able to transfer these to the third agent just yet.

Figure 4.33: Unique meanings in the population lexicon for $W = \{R_1^2\}$.



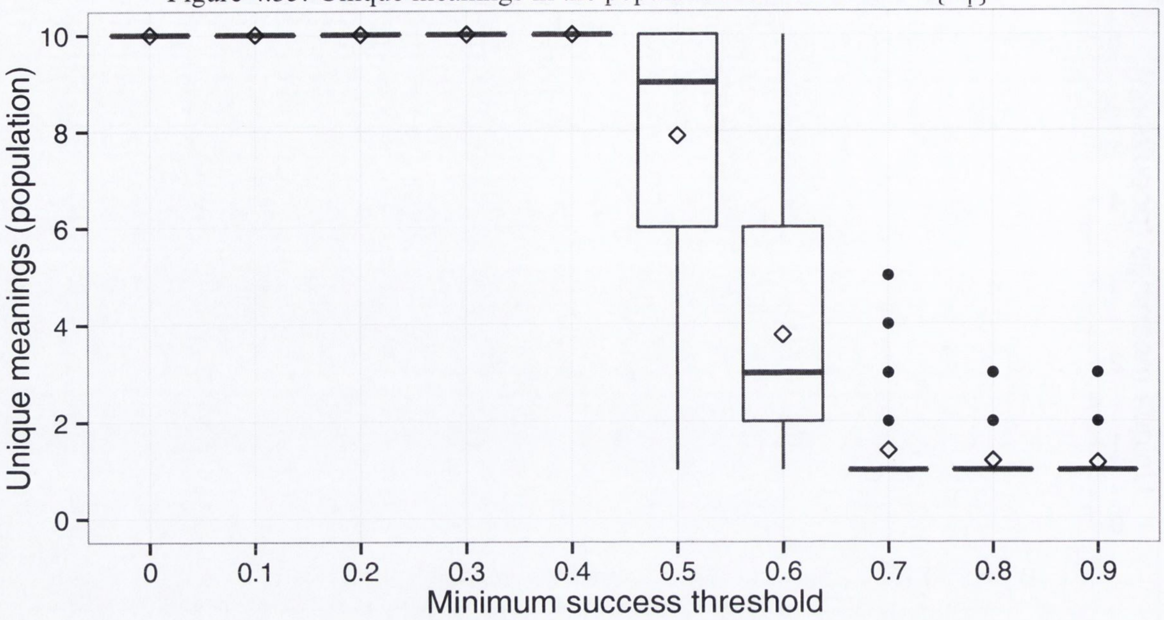
In a world with a single three-argument event, the overall trends remain largely the same. Figure 4.34 shows that the transition from having a mapping associated with every meaning ($s_{min} \leq 0.4$) to largely just being able to express one meaning ($s_{min} \geq 0.7$) moves down by one tenth of the level of minimum success imposed on the agents, as compared to simulations with a single R^2 event.

Figure 4.34: Unique meanings in agent lexicons for $W = \{R_1^3\}$.



The population lexicons in experiments with an R^3 event always contain a mapping for each of the ten meanings for $s_{min} \leq 0.4$, even if there are numerous individual agents who have one or two of the meanings missing from their lexicons, figure 4.35 shows. A slight increase in the number of meanings expressible on the population level is also observed for the higher levels of minimum success threshold. Both observations are again caused by two agents agreeing on a mapping for some meanings, for which the third agent in the triad has not yet been able to pick a corresponding lexical form.

Figure 4.35: Unique meanings in the population lexicon for $W = \{R_1^3\}$.



The conclusion that can be drawn based on the results presented above is very much in line with what was said about the simulations performed with two agents. As was the case there, there is certainly no guarantee that the agents will be able to learn a lexical mapping for every meaning in the world after as little as 1000 interactions each. On top of that, having an additional agent certainly does not help the cause, as the good understanding between one pair of agents can be easily undone by a third, 'uncommunicative' agent, with significant effects on the lexicons of all three agents.

Unique Forms

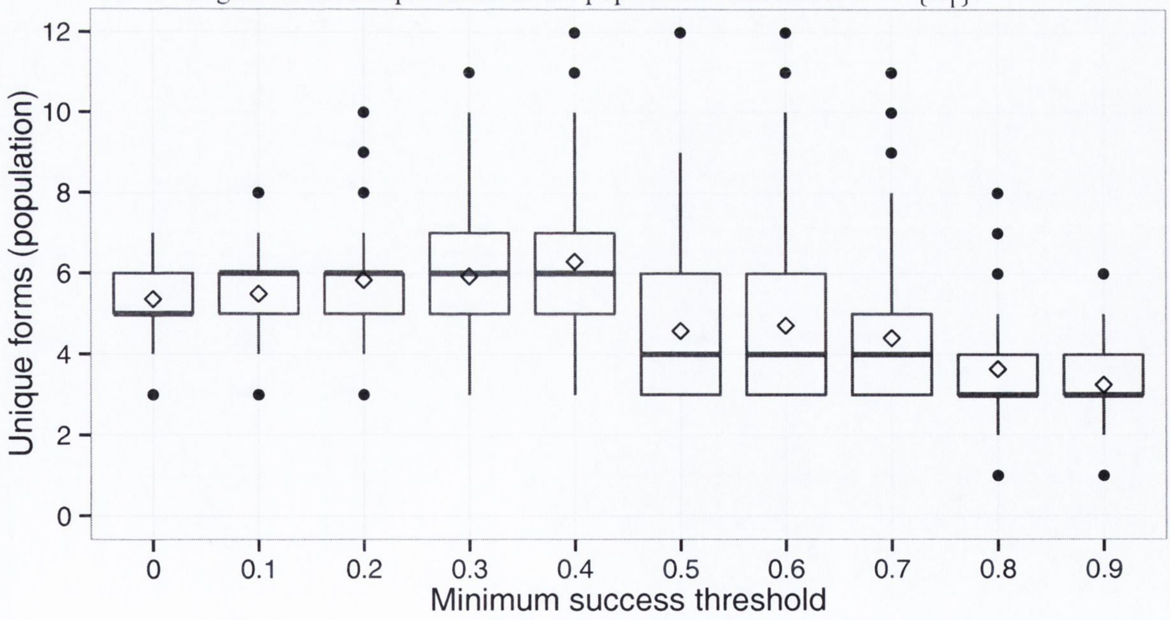
When analysing simulations performed with two agents, an interesting observation that was made was that the number of distinct lexical items in the agents' individual lexicons will be equivalent to the number of distinct forms. As mentioned in section 4.2.1, this is a direct consequence of agents not being able to assign a form to multiple meanings. Furthermore, it was observed that the number of distinct forms in the combined lexicon of the population (i.e. the two agents) will be also equivalent to the number of distinct forms in any individual lexicon. This latter feature was due to the fact that, synchronous word segmentation assumed, the two agents taking part in an interaction will always use the same forms and since the reinforcement or punishment of an individual form used in an interaction is not defined by the success of *that* form, but of the combined success of the whole utterance, which will be the same for both agents, the weight associated with any form will be always adjusted by the same amount in the lexicons of both the speaker and the hearer, meaning that both agents will both add and eliminate forms to/from their lexicon simultaneously.

The equal payoff received by both agents participating in any individual interaction clearly remains in populations of multiple agents, as the number of agents sitting by the side without interfering does not affect the two interacting agents. Similarly, the fact that the agents cannot assign multiple meanings to the same form is not altered either, as this does not even depend on the number of agents involved in an interaction, but solely on the way form generation and lookup are defined in the LEW. In particular, the latter implies that the number of unique forms will remain equivalent to the number of lexical mappings in any agent's lexicon, as well as that agent homonymy remains impossible in a population of three or more agents.

What does change, however, is the capability of an agent, having just associated a form with a certain meaning in an interaction with one agent, to engage in an interaction with another agent and have that mapping eliminated straight away and then possibly replaced by a different meaning-form association involving the same form. Consequently, the number of distinct forms in agent lexicons can vary significantly in a population of multiple agents. In addition, given that the different agents can end up with lexicons containing very different forms, the total number of unique forms present in the population lexicon will also be no more restricted by the size of agents' lexicons. In fact, the number of unique forms in the population lexicon will be higher than the number of distinct forms in any agent's lexicon as long as there is at least one form that the agent with the highest number of unique forms does not have in his lexicon.

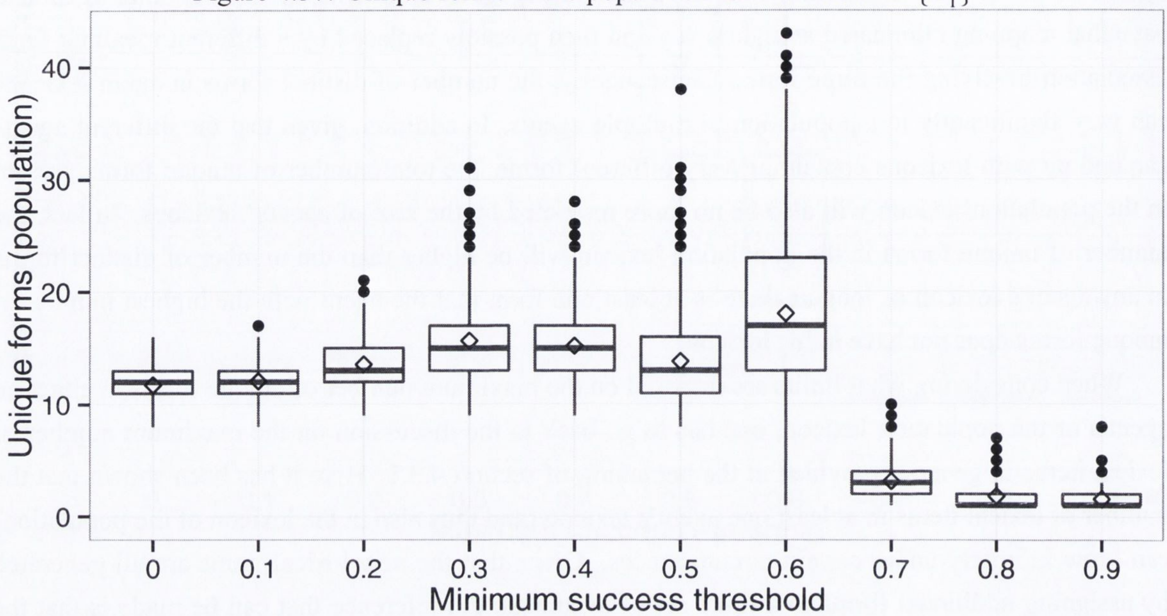
When considering what limits are imposed on the maximum number of unique forms in either an agent's or the population lexicon, one has to go back to the discussion on the maximum number of lexical items in general, provided at the beginning of section 4.3.1. Here it has been shown that the number of lexical items in at least one agent's lexicon (and thus also in the lexicon of the population) can grow infinitely under certain circumstances. Given that the new lexical items are all generated by assigning additional forms to a finite set of meanings, the inference that can be made is that the number of distinct forms and the number of lexical items in both the agent lexicons and the overall population lexicon will change synchronously with any increase or decrease in lexical mappings. Accordingly, there is no maximum limit on the total number of unique forms on either the agent or the population level in a population of three or more agents.

Figure 4.36: Unique forms in the population lexicon for $W = \{R_1^1\}$.



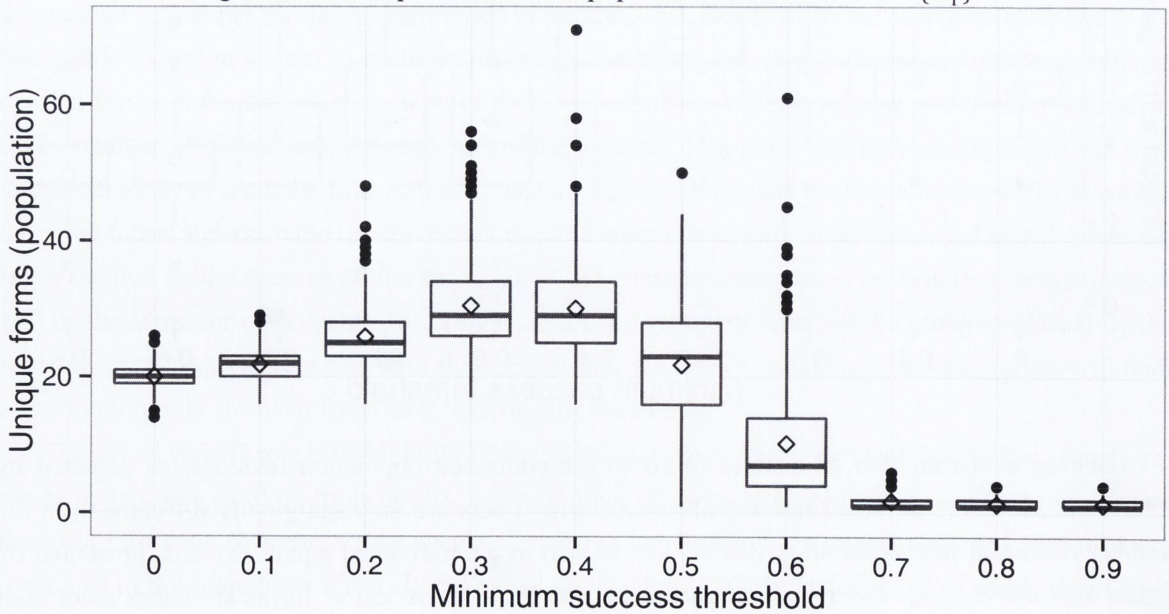
Looking at the number of distinct forms in the combined population lexicons, as depicted by figure 4.36, it has to be noted that the observed form counts are only marginally different from the number of lexical items (and thus also distinct forms) in the individual agent lexicons, presented by figure 4.24 above. This observation suggests that the majority of 'extra' forms are either associated with the same meanings between the agents, or are only present in one agent's lexicon. This conclusion can be made based on the intuition that, if every agent had a large number of individual lexical forms that were not a part of shared mappings, then the number of distinct forms in the combined lexicon of the population would significantly outweigh that of the individual agents.

Figure 4.37: Unique forms in the population lexicon for $W = \{R_1^2\}$.



In experiments with an R^2 event, the picture is slightly different from what was observed in simulations with the R^1 event, figure 4.37 suggests. Here, the number of distinct forms in the overall population lexicon is at times over 50% higher than the corresponding number in any individual agent's lexicon (see figure 4.26). What can be concluded from this is that agents start ending up with non-shared mappings with distinct forms that quickly inflate their lexicons, without them being actually reliably usable in interactions with other agents. No significant differences in the number of distinct population forms can be observed in simulations with an R^3 event when compared to experiments with an R^2 event, as shown by figure 4.38.

Figure 4.38: Unique forms in the population lexicon for $W = \{R_1^3\}$.



In summary, given a low to moderate threshold of minimum success, agents in a population of three tend to end up learning quite a large number of redundant forms that they would not be able to utilize in future conversations anyway, as most of these forms only remain in the lexicon of one of the agents. This stands in stark contrast to simulations with just two agents where the lexicon item and also the distinct form counts were limited by the combination of possible world meanings, allowable event partitions and the level of minimum success. Furthermore, there appears to be no clear optimal level of minimum success that would result in the agents learning just the right amount of forms for the given set of referable meanings. The figures in this section suggest that there is a certain level of minimum success threshold, below which most agents end up learning significantly more forms than they need to and above which the majority of them does not manage to learn a sufficient number of forms to express all meanings. However, there is also a reasonably large number of simulations where agents have learned just the right amount of words. These findings underline the difficulty, i.e. the low chances of a reliable communication system being developed from nothing in a very short period of time that our predecessors were up against.

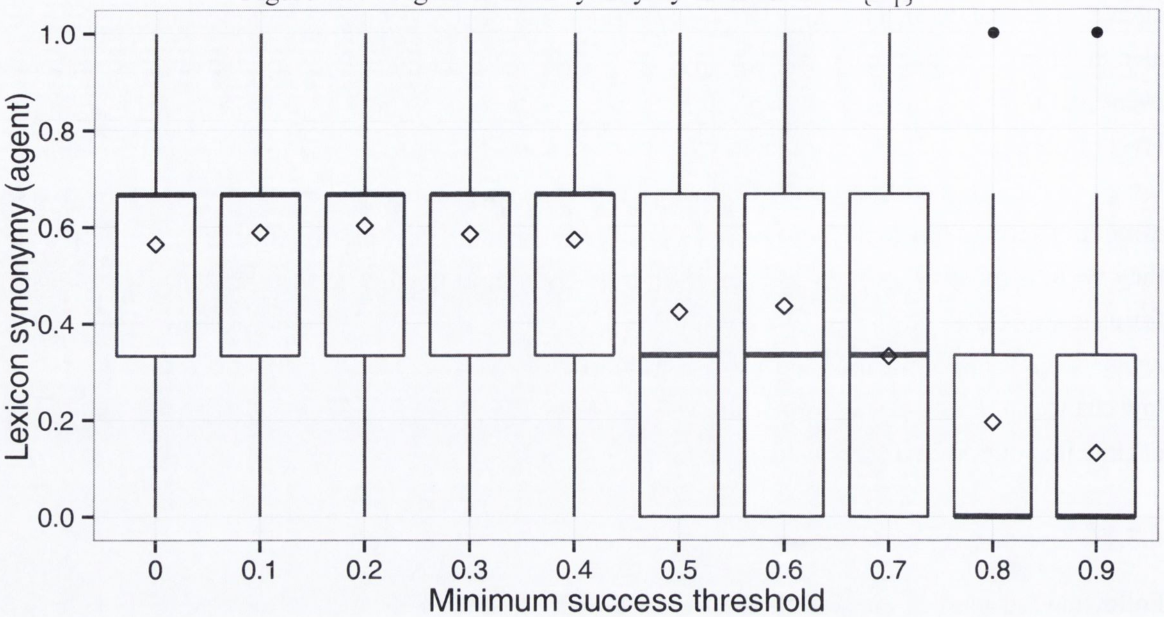
4.3.2 Synonymy & Homonymy

Following the approach in section 4.2, the analytical parts of the following subsections will be based on the assumption that the agents have reached a state where they have at least one mapping associated with every meaning in the world. Naturally, it can happen that after 500 epochs of a simulation run, the agents have not yet reached that state, meaning that the experimental results may not necessarily reflect the analytical limits.

Agent Synonymy

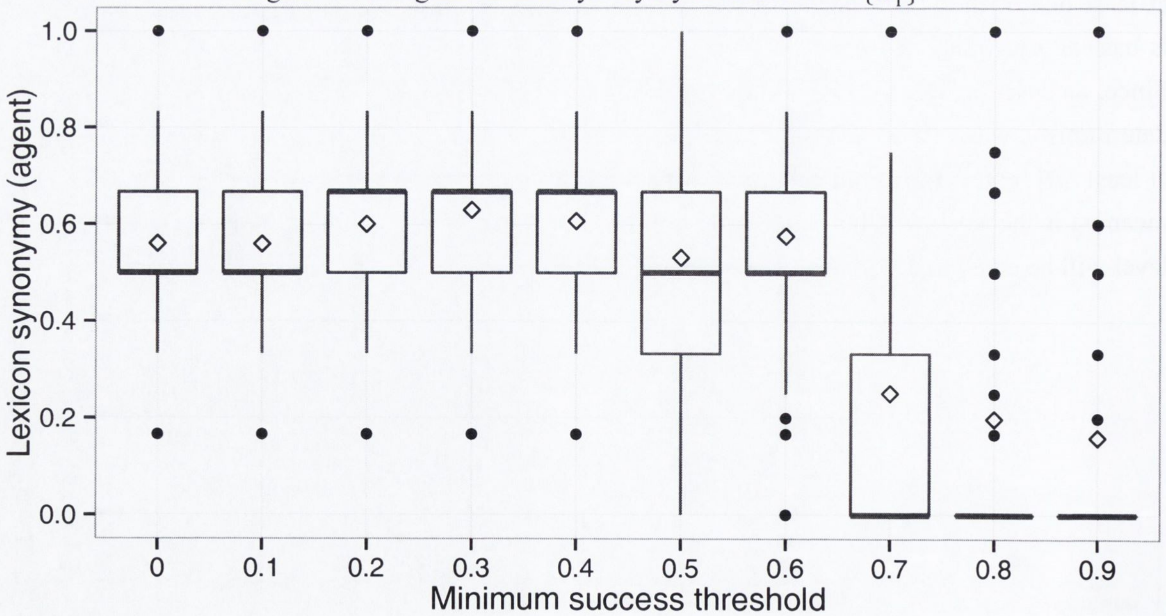
Agent synonymy levels in a population of three agents adhere to the same major principles as with just two agents involved. In particular, synonyms are created when a newly encountered form is assigned to a meaning that already had another form associated with it. As this can only occur when at least one of the two forms involved was unsuccessfully employed in an interaction, synonymy is basically restricted by the number of forms from non-shared mappings between the two agents. Since, as described above, the number of forms in at least one agent's lexicon is not limited in any deterministic way in a population of three or more agents, there is every chance that, having acquired at least $|M|$ 'extra' forms, an agent will assign all of these to different meanings. In this case, every meaning in the world will have at least two forms associated with it and the agent's lexicon synonymy level will be equal to 1.0.

Figure 4.39: Agent lexicon synonymy level for $W = \{R^1_1\}$.



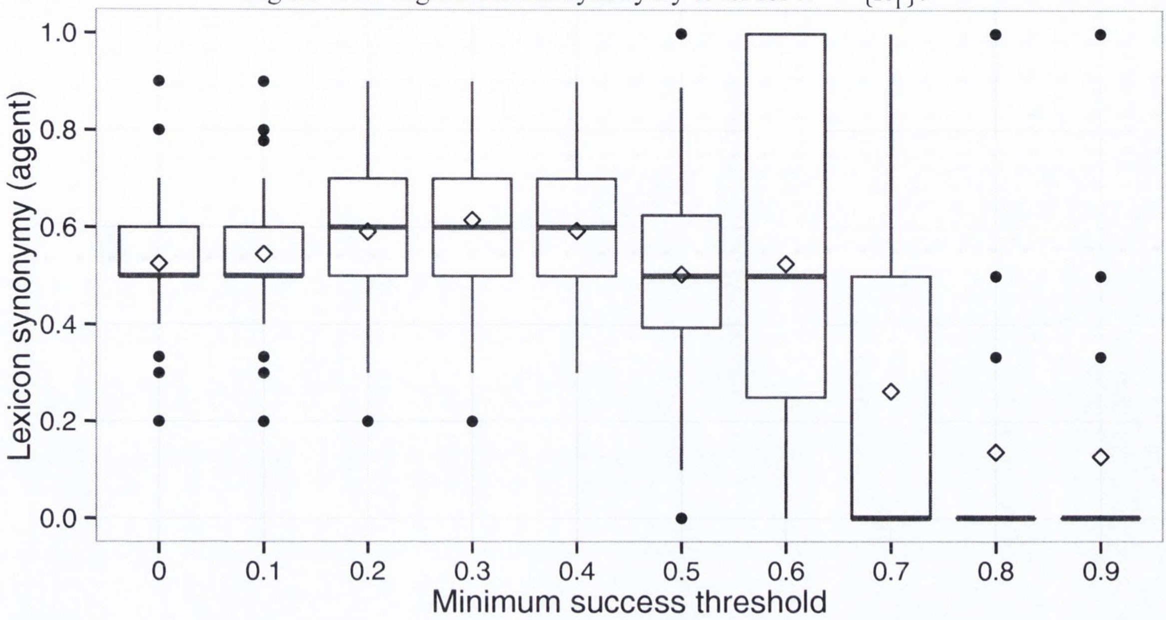
The discussion of agent synonymy limits outlined above is directly reflected by figure 4.39, in which the synonymy levels of agent lexicons are shown to take up every possible value for every level of minimum success experimented with. Furthermore, it is only for $s_{min} \geq 0.8$ that not having synonymous forms in one's lexicon is the more dominant outcome of the simulations. However, even for such high levels of minimum success imposed on the agents in a world with just the one R^1 event, it can be seen that some agents still end up having multiple forms assigned to every meaning in their lexicon.

Figure 4.40: Agent lexicon synonymy level for $W = \{R^2_1\}$.



There are a few small differences to the distribution of agent synonymy levels that can be detected when examining the results of simulations with an R^2 event, as depicted by figure 4.40. The major observation is that for $s_{min} \leq 0.4$, every agent in all of the 500 simulations runs executed for each of the corresponding levels of minimum success ends up having at least one meaning with synonymous forms in his lexicon. At the same time, agents who do not acquire any synonyms after 1000 interactions are in the majority from $s_{min} \geq 0.7$ already, even if just narrowly for $s_{min} = 0.7$.

Figure 4.41: Agent lexicon synonymy level for $W = \{R_1^3\}$.



It is interesting to note that, even if agents in simulations with an R^3 event and $s_{min} \leq 0.4$ tend to end up learning between double and five times the amount of forms needed to express the ten meanings available in their world, as was illustrated previously by figure 4.28, it turns out that none of these agents actually ends up assigning multiple forms to each of the 10 different meanings, as suggested by figure 4.41. Similarly to simulations with an R^2 event, none of these agents manages to end his 1000 interactions without having a synonym in his lexicon either. In fact, every agent in simulations with $s_{min} \leq 0.4$ ends up connecting multiple forms with at least two different meanings.

Population Synonymy

Since the population synonymy level is at least as high as the maximum level of synonymy in the lexicons of the agents involved, it is trivial to conclude that population synonymy will reach a level of 1.0 as soon as one agent assigns at least two forms for every meaning in the world. It will not matter if none of the other agents will have any synonyms in their lexicons since all of the agents inhabit the same world with the same set of meanings. Accordingly, since it can happen that at least one agent reaches ‘full’ synonymy before agreeing on lexical mappings and stopping its lexicon growth, population synonymy can (and is in fact very likely to) reach the analytical maximum of 1.0 in a population of three or more agents.

Figure 4.42: Population lexicon synonymy level for $W = \{R_1^1\}$.

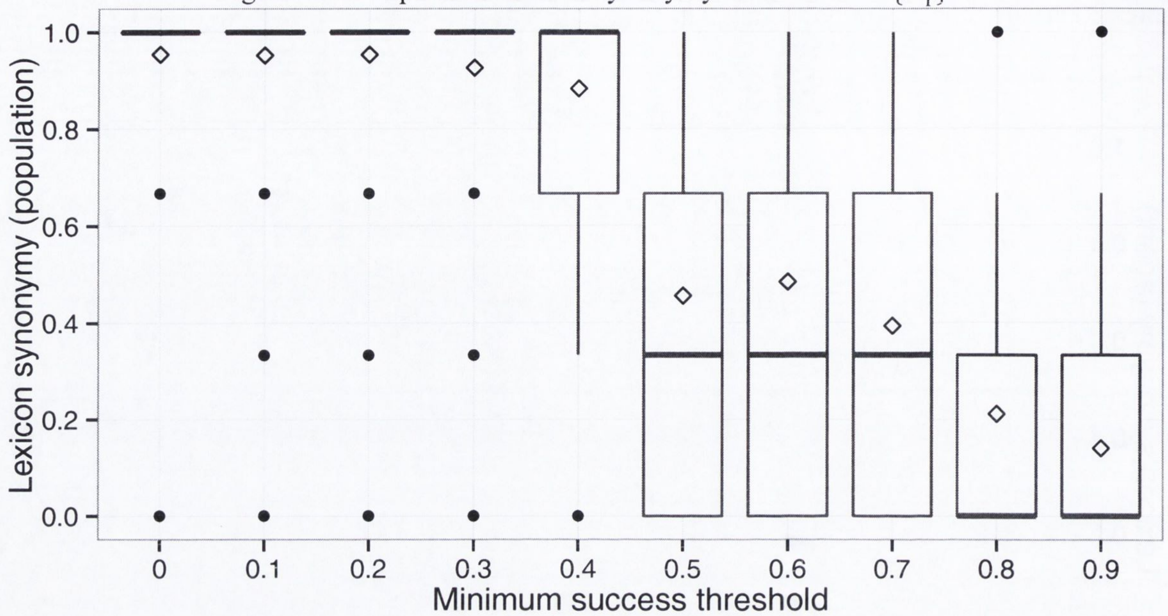
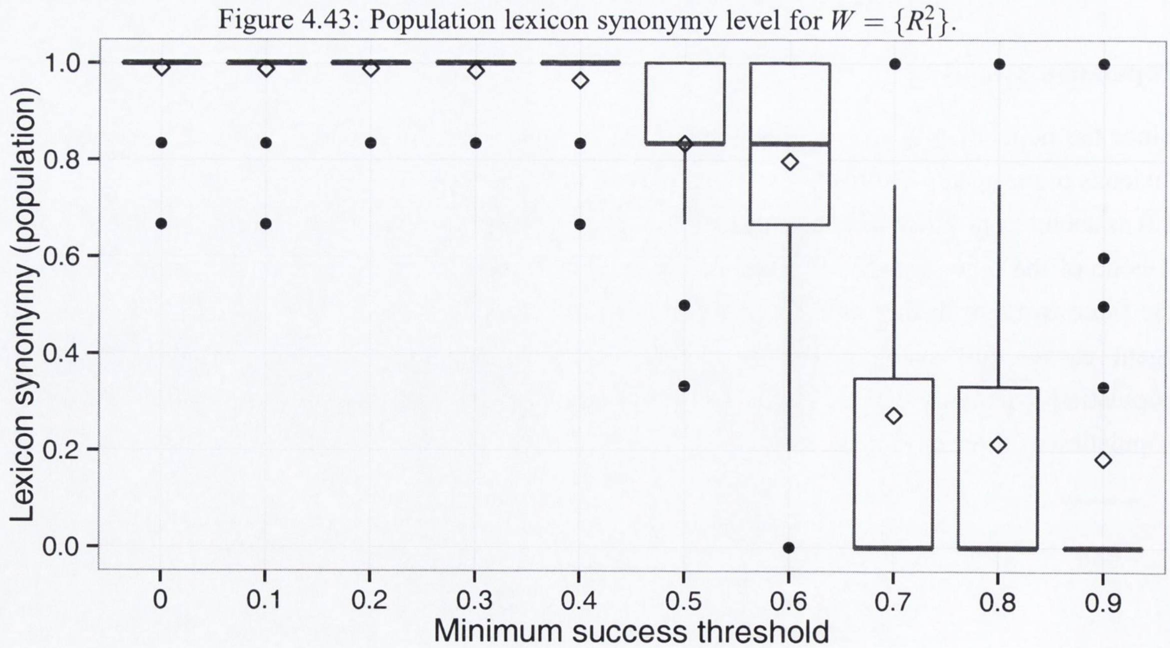


Figure 4.42 shows that, as discussed and anticipated in section 4.2.2, population lexicon synonymy is on average at least as high as agent lexicon synonymy. As can be seen, for $s_{min} \leq 0.4$, the majority of the populations end up developing a lexicon that has multiple forms assigned to each of the four available meanings across the agent lexicons. However, there are lexicons that exhibit no synonymy at all, even for $s_{min} = 0$, which is remarkable, given that every agent in such simulations was observed to have acquired at least three distinct forms (cf. figure 4.24), meaning that all three agents agreed on (or in the case where not all agents had a form associated with every meaning - did not disagree on) the mapping for at least three of the four potential meanings.

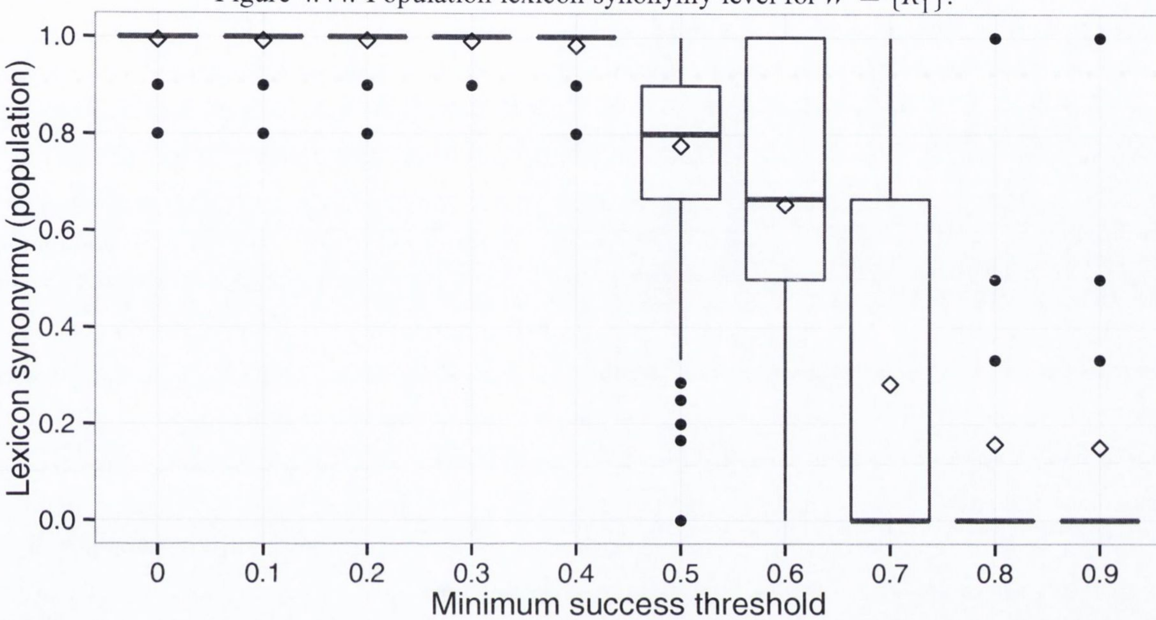
Figure 4.43: Population lexicon synonymy level for $W = \{R_1^2\}$.



In simulations with an R^2 or R^3 event, as depicted by figures 4.43 and 4.44, respectively, the population lexicon for simulations with $s_{min} \leq 0.4$ clearly exhibits an overwhelming excess of forms across the agent lexicons. As can be seen, the population synonymy level is predominantly fixed at

1.0, with only a handful of exception cases where the synonymy level is not that high. Admittedly though, even in such cases, the level of population synonymy is always above $\frac{2}{3}$. For higher levels of minimum success threshold, the trend reverses, in the sense that most population lexicons actually do not have any synonyms in them. However, one has to consider that the number of distinct forms in the agent and population lexicons in the corresponding simulations often did not exceed a single form, as shown by figures 4.26 and 4.28.

Figure 4.44: Population lexicon synonymy level for $W = \{R_1^3\}$.



To summarize the discussion of agent and population lexicon synonymy levels, it is important to note that low synonymy by itself is not necessarily an indicator of a positive lexicon acquisition outcome. Before coming to that conclusion, the actual number of forms learned by the agents needs to be considered. Only if this number is sufficiently close to the number of distinct meanings in the world, can one claim that agents have successfully acquired a useful and reliable lexicon. Unfortunately, in the above simulations, this claim can only be made for a handful of simulations with an R^1 event. In terms of comparing the levels of synonymy observed here to values in real languages, it appears that the nearest average match is for $s_{min} = 0.7$, going slightly up in terms of the minimum success requirement when compared to simulations with two agents.

Agent Homonymy

The fact that agents are not able to assign additional meanings to an existing form in their lexicon, either during production or interpretation, is in no way connected to the number of agents involved in the simulation itself. It remains hence impossible for agents to introduce homonyms to their lexicons, meanings that agent homonymy levels remain fixed at 0 throughout all experiments presented in this work.

Population Homonymy

In a population of two agents, the set of forms in the lexicons of the agents was always equivalent, meanings that determining population homonymy levels was as straightforward as calculating the maximum number of forms from non-shared mappings that would by definition have different meanings assigned to them between the two agents. In simulations conducted with three or more agents, agents no longer share the set of distinct forms amongst themselves. In fact, it has been shown in section 4.3.1 that it is possible for an agent to accumulate a theoretically unlimited set of forms, without the other two agents acquiring any forms at all. In such a scenario, the population homonymy level would be actually equal to 0 as only one agents would have any forms at all in his lexicon and, since agent homonymy is impossible, none of the forms would by homonymous on either the agent or population level.

However, as mentioned above, this section focuses on limits of lexicon synonymy and homonymy in a state where every agent has at least one form associated with every possible meaning in the world. Furthermore, it is important to note that population homonymy is only increased if a form is present in the lexicon of two or more agents and associated with at least two different meanings at the same time. If, on the contrary, one agent learns a large number of forms without any other agent remembering any of these, none of these forms can have multiple meanings assigned to them and thus the population homonymy will be actually falling.

As can be seen from the above, in order to estimate the maximum level of population homonymy, it is necessary to calculate the maximum number of forms that can be learned by multiple agents and at the same time not assigned to the same meanings between all of these. In addition to that, minimizing the number of total unique forms in the population lexicon also increases the level of lexicon homonymy. Following the discussion on lexicon size in section 4.3.1, maximum population homonymy is achieved when all agents involved acquire the same set of shared mappings, thus minimizing the amount of non-homonymous forms in the population lexicon. In this case, for every remaining meaning that is not a part of the shared set, every agent has the possibility of assigning a form to it that is different from the one selected by all (or some) of the other agents, thus making it a homonymous form and increasing the number of forms with non-shared meaning referents.

Figure 4.45: Population lexicon homonymy level for $W = \{R_1^1\}$.

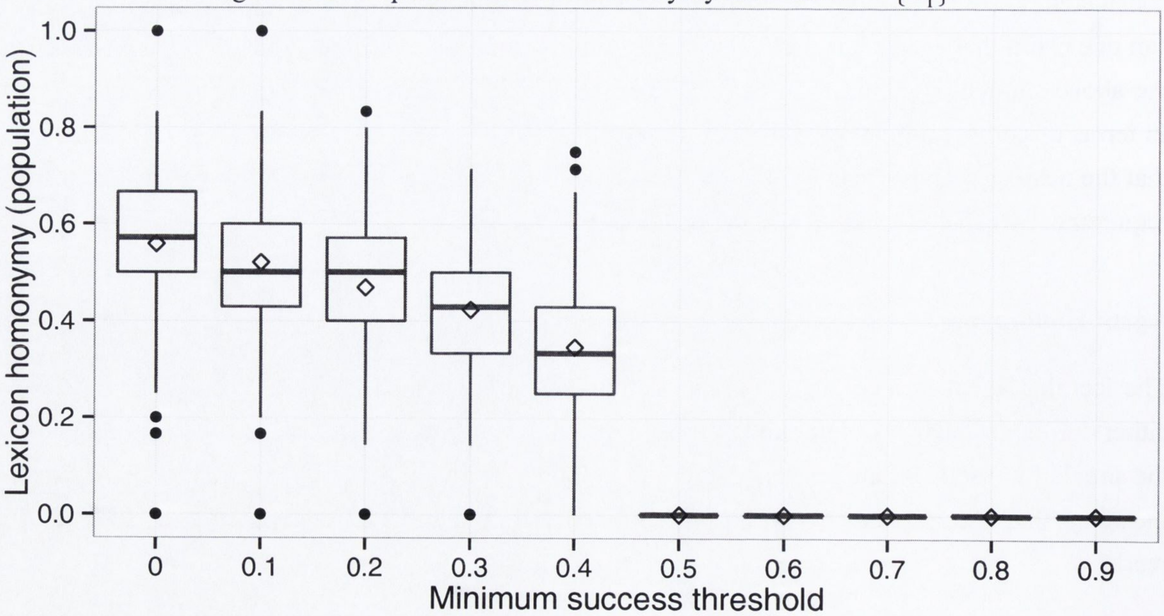
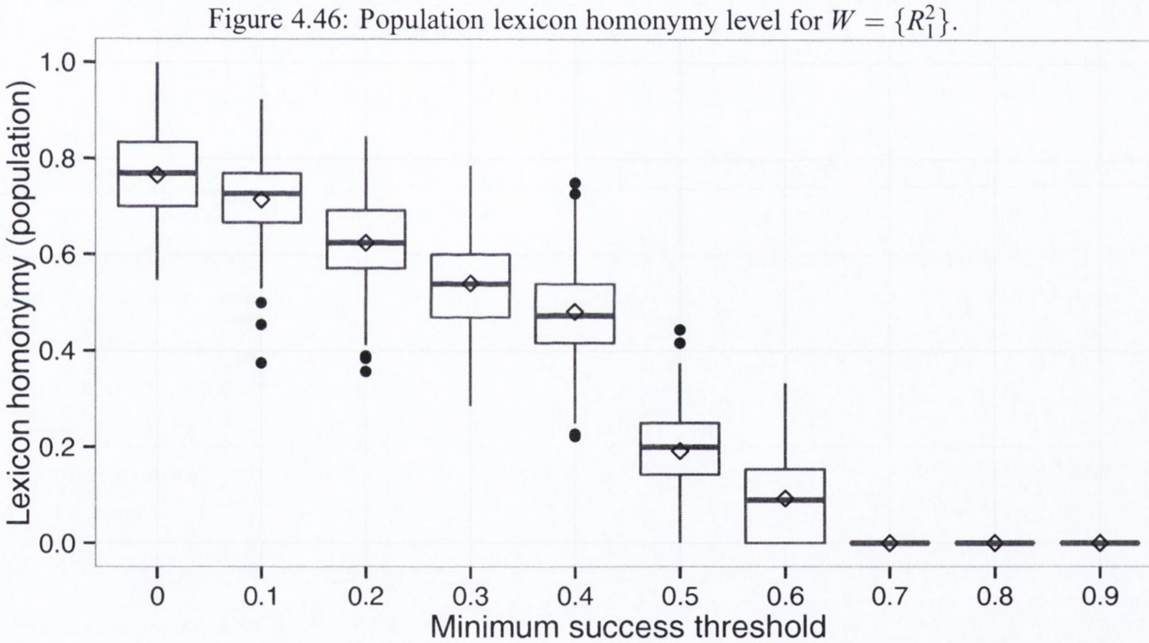


Figure 4.45 shows that in simulations involving three agents and one event of the type R^1, population homonymy hovers between 0.3 and 0.6 for $s_{min} \leq 0.4$ and is completely obsolete for $s_{min} \geq 0.5$. For the lower levels of minimum success required, it is interesting to note that simulations with no

resulting population homonymy are also observed, even these are rather few and far between. What this implies is that even given a very low success requirement, agents can end up learning the same lexical mappings across the population. On the other hand, agents with very low success requirement sometimes end up with population homonymy level reaching 1.0. This is particularly likely for the lowest level of s_{min} since here the punishments are either non-existent (for $s_{min} = 0$) or extremely low (for $s_{min} = 0.1$), meaning that mappings are never or almost never eliminated from the lexicons of the agents involved and thus that every form, even if incorrectly interpreted, is likely to be present in the lexicon of at least the two agents who utilized it once, thus resulting in maximum homonymy.

The lack of homonymy for higher levels of minimum success can be clearly explained by the higher success requirement, which implies that, in the case of simulations with one R^1 event, the agents can only obtain a payoff with $s_{min} \geq 0.5$ if they agree on all of the meanings involved in an interaction. Consequently, only matching mappings will be ever stored in the lexicons of the agents involved, thus making homonymy impossible.[12]

Figure 4.46: Population lexicon homonymy level for $W = \{R^2_1\}$.

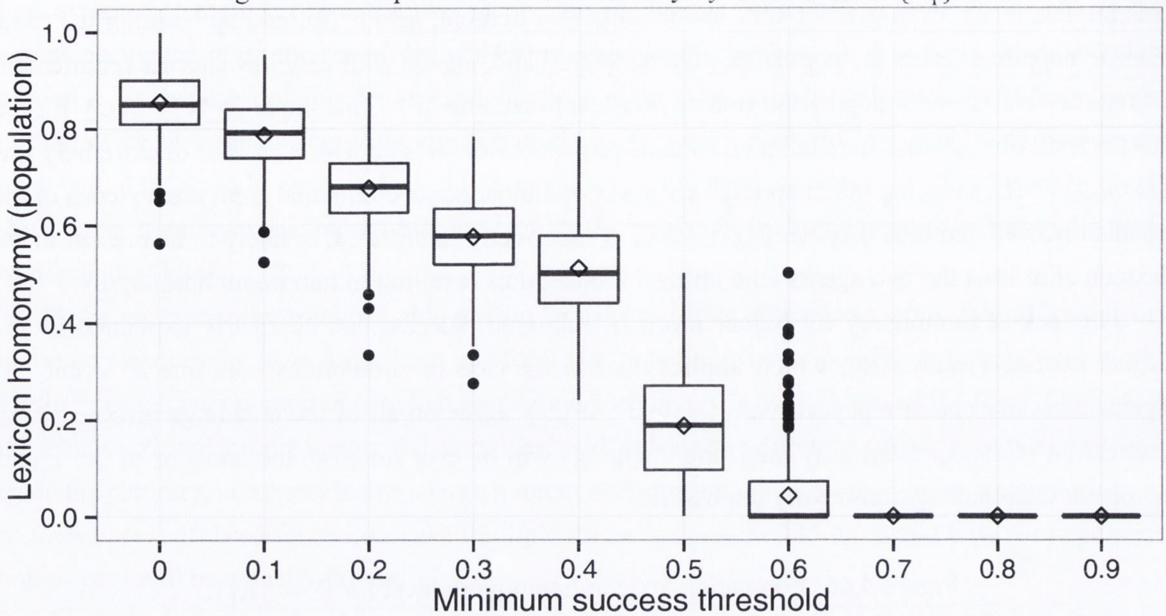


Figures 4.46 and 4.47 depict similar overall trends of population homonymy levels for simulations with an R^2 and an R^3 event, respectively. Population homonymy is at its highest for very low levels of minimum success, reaching 1.0 in some cases where $s_{min} = 0$. The homonymy level gradually increases as s_{min} is increased, falling to a fixed 0 for $s_{min} \geq 0.7$, i.e. where the agents would need to agree on all mappings of an utterance in order to add these to their lexicons.[13] Contrary to simulations with an R^1 event, there is not a single simulation run with $s_{min} \leq 0.5$ where an agents would not develop any cross-populational homonymy.

[12]Note that in simulations with four or more agents, disjunctive pairs could evolve fully agreed on dialects of their own even under a high requirement for communicative success. In this case, population homonymy would always be possible, regardless of the level of s_{min} imposed on the agents.

[13]In fact, for simulations with an R^3 event, agents could still learn one incorrect meaning-form association for utterances with the event partitioned into four parts and where $s_{min} = 0.7$, as decoding three of four forms correctly in such a scenario would actually imply a payoff of 0.05. However, it appears that this is either unlikely to happen, or is quickly corrected in future utterances, as the population homonymy level in corresponding experiments is fixed at 0 after 1000 interactions.

Figure 4.47: Population lexicon homonymy level for $W = \{R_1^3\}$.



As could be expected, population homonymy levels are reasonably high in simulations with multiple agents, where every subsequent pair of agents can have its own set of shared mappings and a large number of additional, non-shared mappings, which end up automatically generating additional homonyms when all the agent lexicons are combined. Naturally, for higher levels of minimum success, population homonymy becomes impossible in simulations with three agents. However, in these simulations runs, agents also only acquire a very limited amount of forms, which is usually not sufficient to express even the reasonably small amounts of meanings that are presented to them in the simulations outlined above. The implications of all of the above considerations on actual communicative success will be discussed in the following section. Finally, in terms of comparability to natural languages, population homonymy is just slightly off the 0.17 mark for $s_{min} = 0.5$ here, just as it was in simulations with two agents.

4.3.3 Communicative Success

In the previous sections, it has been observed that, in a population of three agents, lexicons tend to have significantly more forms than meanings for lower levels of the minimum success threshold, and just a few forms for higher levels of the parameter. Furthermore, the synonymy levels in both the individual and population lexicons appear to be fluctuating between the minimum and the maximum possible levels across the individual simulations performed for each of the ten different levels of minimum success that were experimented with. However, given the high levels of population homonymy for the lower levels of success threshold, it is expected that lexicon precision will suffer slightly in these simulations. On the other hand, the low average number of lexical mappings for the higher levels of s_{min} will certainly impact lexicon use, even if lexicon precision in these simulation can be expected be quite high. The following graphs depicting the correlation between lexicon use and lexicon precision for the three different event types in simulations performed with three agents show how the different properties of the lexicon discussed above actually influence the overall performance of the agents in the simulated interactions.

Figure 4.48: Lexicon-based communicative success for $W = \{R^1_1\}$.

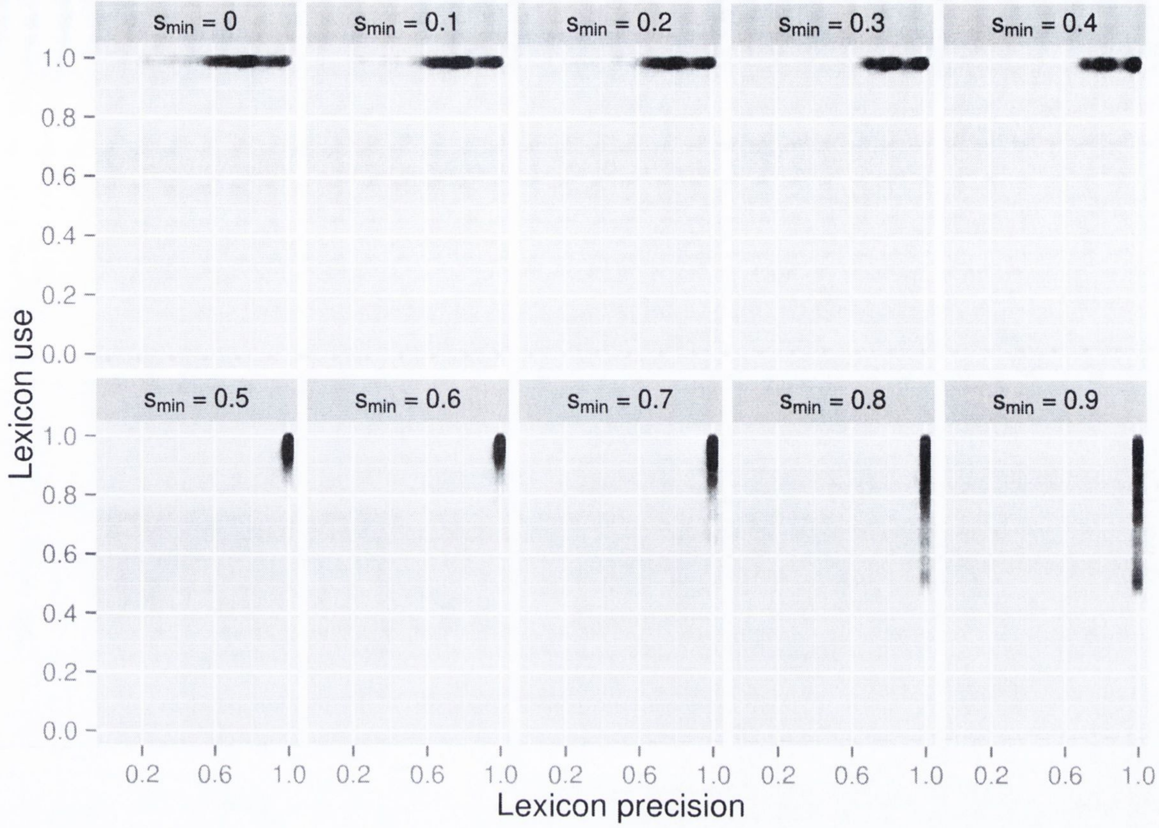
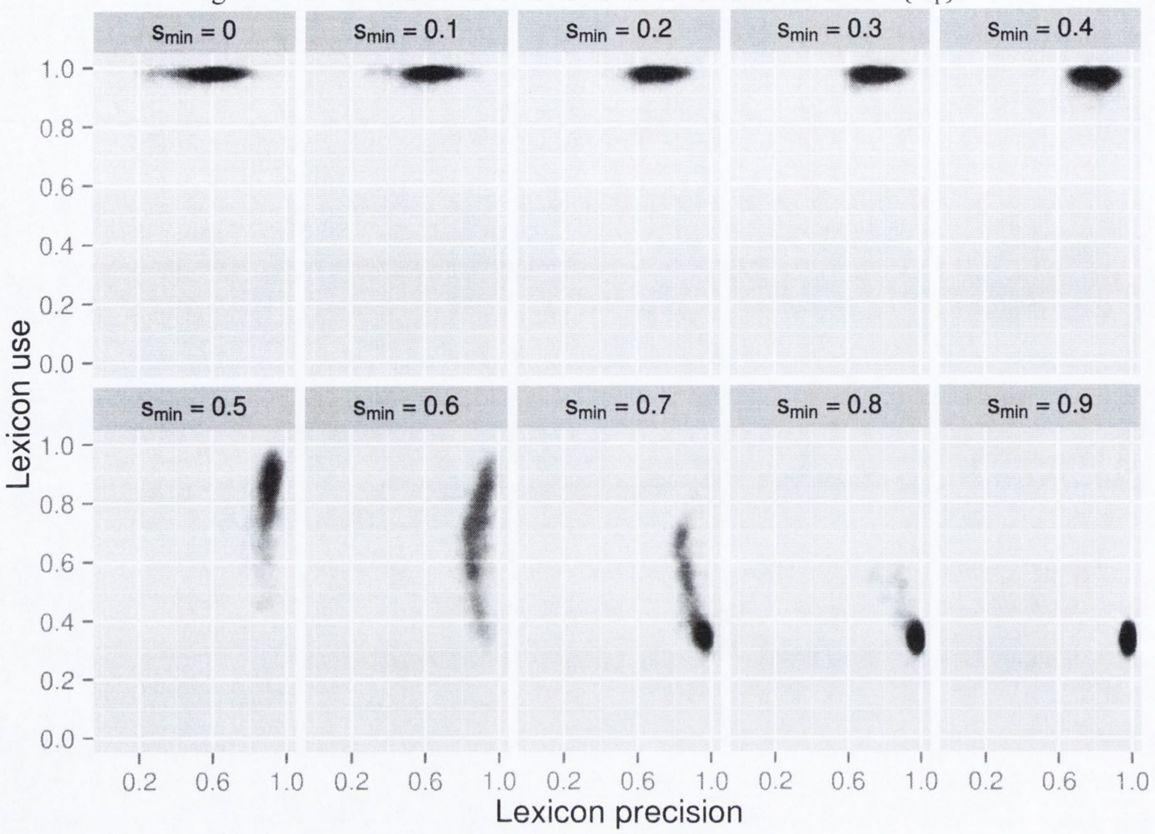


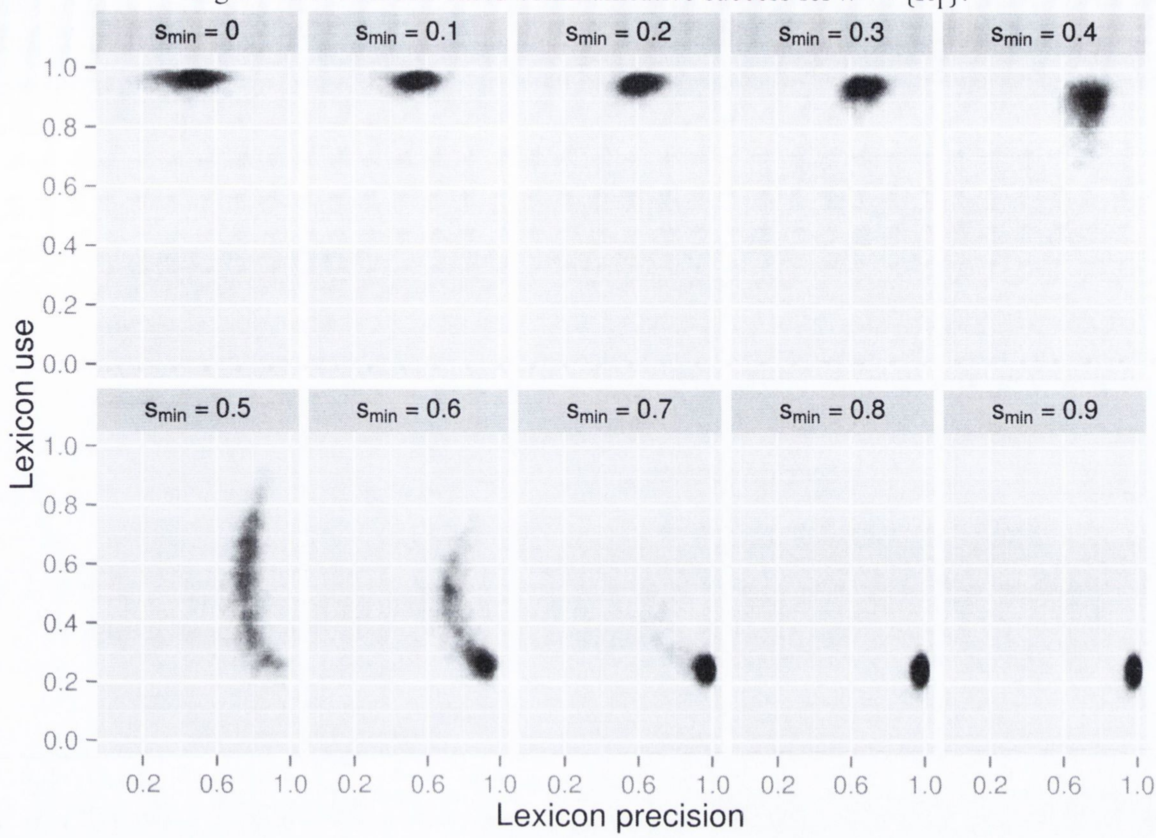
Figure 4.48 indicates a clear qualitative difference between simulations performed with $s_{min} \leq 0.4$ and those with $s_{min} \geq 0.5$. In the former case, agents are basically always able to find a lexical item associated with a heard form in their lexicon, even if the corresponding interpretation is not nearly guaranteed to be the correct one. Having said that, a very sizeable portion of the agents actually reside in the top right corner of their corresponding subgraph, i.e. are always able to use their lexicon in an interaction and also do so with maximum precision. For the higher levels of minimum success threshold, agents clearly take longer to add new lexical mappings to their lexicons, resulting in a gradual decrease of lexicon use. At the same time, the precision of their lexicons is fixed at 100% (which is logical, given that only perfect interpretation in an interaction can lead to a payoff in simulations with an R^1 event and $s_{min} \geq 0.5$).

Figure 4.49: Lexicon-based communicative success for $W = \{R_1^2\}$.



In simulations with the two bigger events, the results of which are depicted by figure 4.49 for R^2 and figure 4.50 for R^3, the interplay between lexicon use and lexicon precision exhibits slightly different effects. For the five lower levels of minimum success, agents in both these experiments seem to be still quickly acquiring a lexicon at the beginning of the simulations that covers all of the forms used by all three agents involved, resulting in near maximum lexicon use. However, contrary to simulations with an R^1 event, these agents never achieve perfect understanding, though admittedly hovering at very respectable lexicon precision levels between averages of 0.47 and 0.80.

Figure 4.50: Lexicon-based communicative success for $W = \{R_1^3\}$.



The major difference between the experiments conducted with one R^1 event and those with the two bigger events lies in the transition from full lexicon use to a more moderate use level, but with almost perfect precision. In the latter two experiments, this transition happens gradually, from $s_{min} = 0.5$ to $s_{min} = 0.8$ and $s_{min} = 0.7$ for R^2 and R^3, respectively. During the transition phase, lexicon use takes up values in the range between 1.0 and the final settling value, which for experiments with an R^2 event is at an average of 0.35 and for those with an R^3 event – at 0.25. Having decreased to this level of lexicon use though, lexicon precision becomes almost fixed at 100%.

Looking at the above results, the question that instantly poses itself is which is better: being able to always use one's lexicon, even if not always reliably so; or having the confidence that, if one interprets a form with the help of one's lexicon, then the interpretation will almost certainly be the correct one. It would seem that the former option would be more desirable in a scenario where interlocutors would be easily able to adjust the nuances of the meanings associated with different forms, without needing to resort too much to guessing after an initial learning period. The reliable lexicon option, on the other hand, seems perhaps even more attractive when considering that an interlocutor should be able to tell if he is using his lexicon or not and, in the case that he is, he can be quite confident of having understood the speaker correctly.

In a way, only having lexical items that one is quite sure of in one's lexicon implies that there should be no need to make any adjustments to the lexicon in the future. There is doubt, however, if that is the way humans actually learn a language, as studies from child language acquisition clearly show that many lexical items that children acquire at a very early stage in their language development are only partially correct, if at all, and get adjusted to the accepted lexical convention later on.

4.4 Interacting in Small Groups

The analysis of the LEW's fundamentals is concluded in this section with an evaluation of experiments performed with groups of ten agents. While the difference between two and three agents was predominantly a qualitative one, in the sense that the agents' involvement in interactions was significantly altered, the change between three and ten agents is of purely quantitative nature. The question here is, given the same average number of interactions per agent per simulation, if the agents will still be able to perform at the level of three-agent groups, or if having a larger number of involved agents will interfere with the lexicon acquisition process even further.

4.4.1 Lexicon

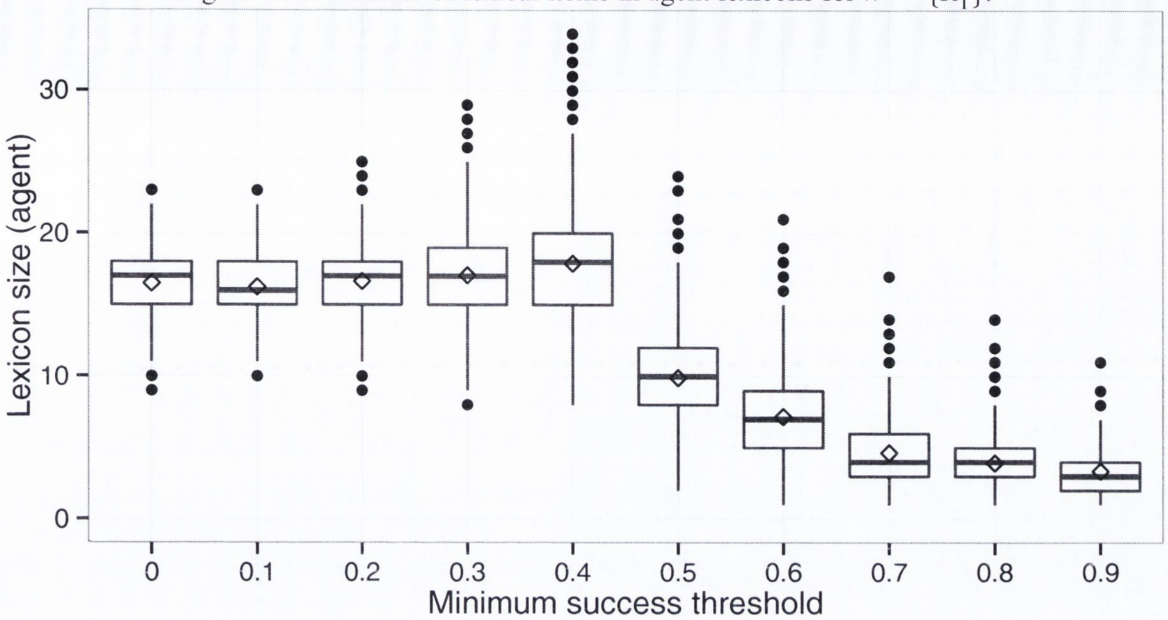
The same general principles that were discussed for the triads also apply to the small groups of ten agents, i.e. there are no limits on the maximum lexicon size of agents or the joint lexicon of the population (other than those defined by the number of interactions in the conducted simulations). Accordingly, this section will proceed directly to the evaluation of the experiments conducted in worlds consisting of one event of the same three types as before and populated by ten agents.

Lexical Mappings

As has been outlined in section 4.3.1, populations of three or more agents theoretically have no limits imposed on the size of the lexicon for at least one of the agents, regardless of the number or the types of available events or the level of minimum success imposed on the agents. What can happen in such groups is that, having just successfully learned a lexical mapping with one interlocutor, an agent can experience zero success when employing the same mapping in a conversation with another partner, which might potentially result in the agent in eliminating the mapping from his lexicon. In this case, the hearer from the first interaction will end up with a quasi-'orphaned' lexical mapping that only he has in his lexicon.

Given that the above scenario is only made likelier by the introduction of additional agents, in particular due to the increased chances of any one agent talking to different partners in successive interactions (50% in a population of three agents vs. 89% in a population of ten), it is expected that the average lexicon size of individual agents, as well as that of the overall population, will be significantly higher in experiments conducted with ten agents, even if the number of actual meanings to express remains the same.

Figure 4.51: Number of lexical items in agent lexicons for $W = \{R_1^1\}$.



Looking at figure 4.51 which depicts the distribution of the number of lexical items in agent lexicons, the predicted observation is that agents here indeed end up with roughly three times the number of words when compared to agents from a three-member group. In addition to that, and as opposed to simulations with triads, for $s_{min} \leq 0.4$ no agent ends up with less than eight mappings in his lexicon, nearly three times the amount of meanings in the agents' world. On the other hand, for $s_{min} \geq 0.5$, several agents do not actually have as much as three lexical mappings in their lexicon, which stands in contrast to simulations with triads where every agent ended up knowing at least three for $s_{min} \leq 0.7$.

Figure 4.52: Number of lexical items in the population lexicon for $W = \{R_1^1\}$.

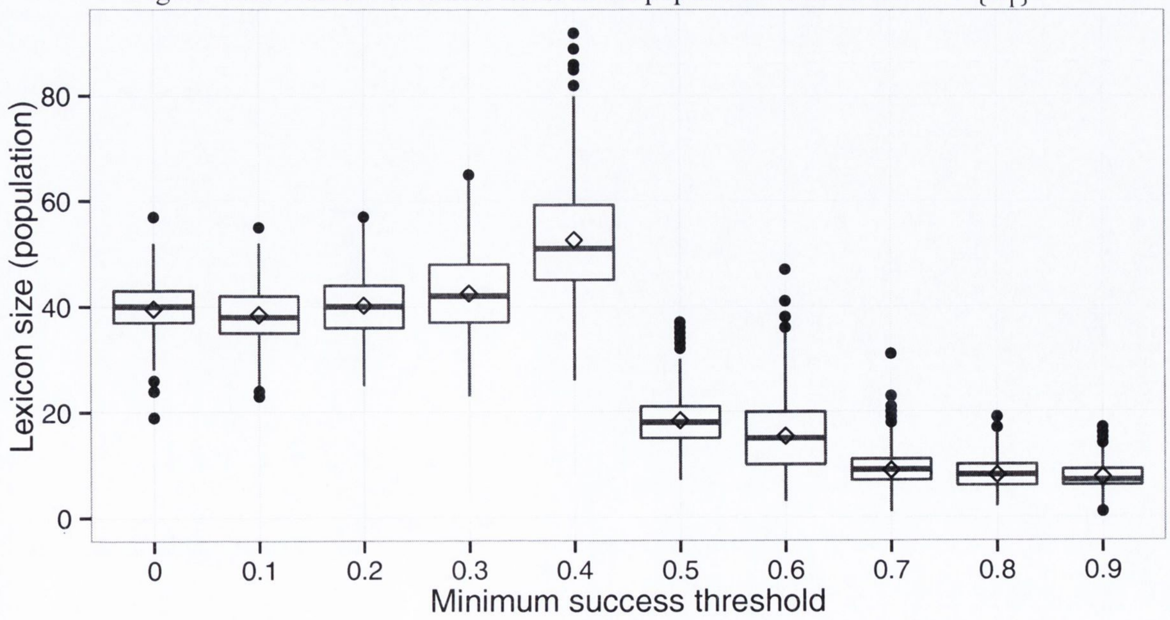
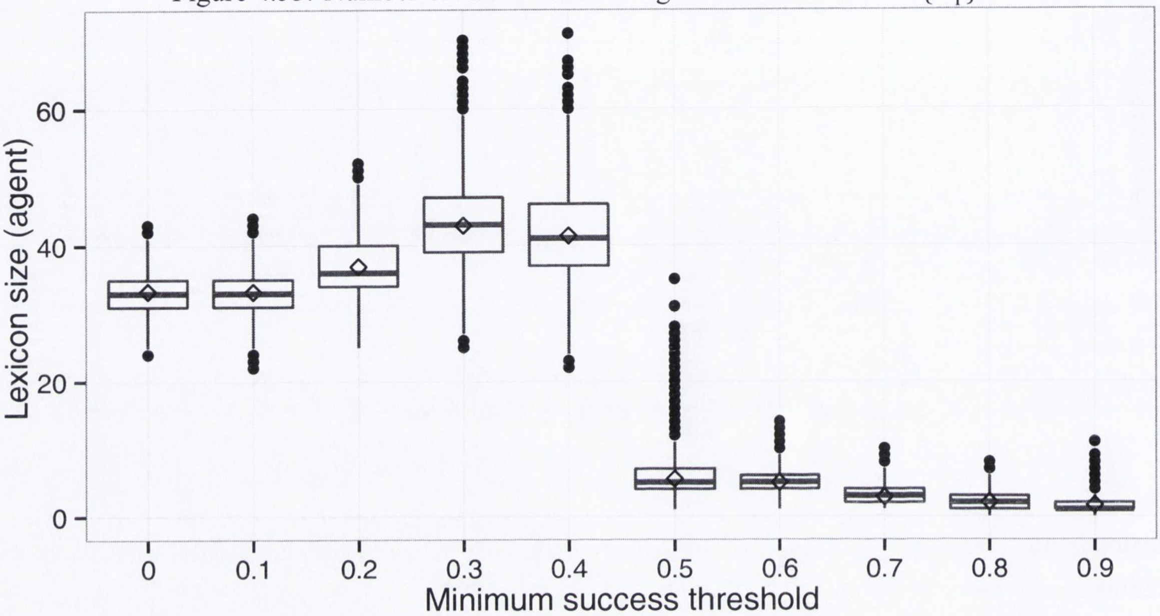


Figure 4.52 confirms the trend on the population level, as it can be clearly seen that, compared to figure 4.25, the population lexicon ends up being three to four times larger for all levels of minimum success, after the same number of interactions per agent (though after 5000 interactions in total, as

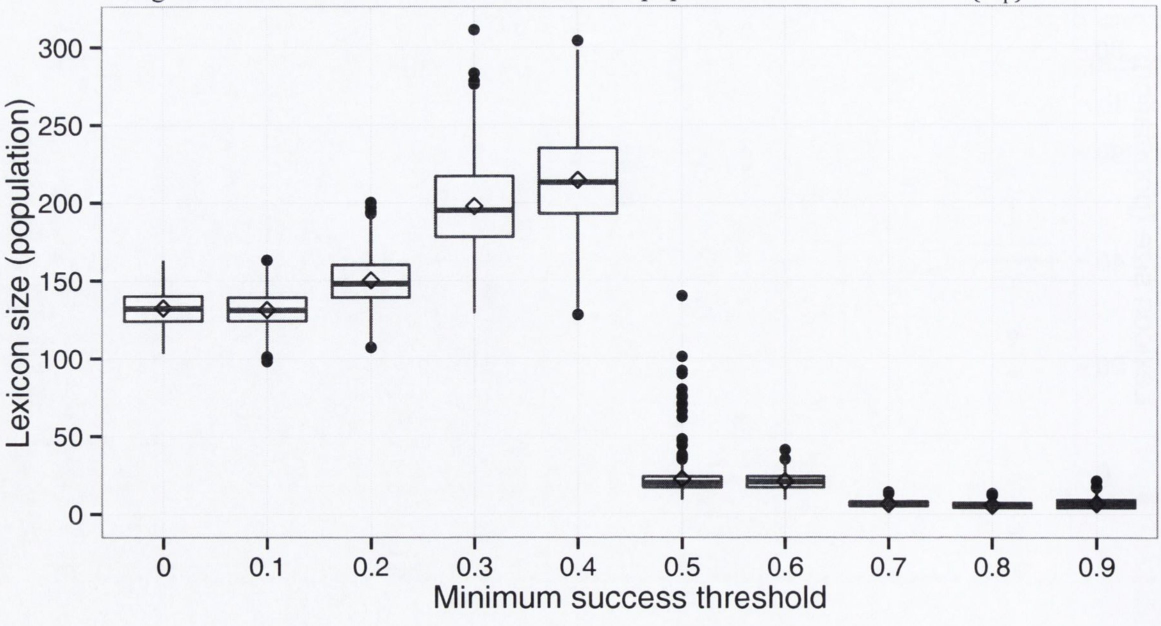
opposed to only 1500 interactions conducted in three-agent experiments). Notably, for $s_{min} \leq 0.4$, the population lexicon includes over 20 mappings at the end of nearly every experiment, peaking at over 90 for $s_{min} = 0.4$, i.e. 30 times the number of mappings that would actually be needed to refer to all of the world's meanings.

Figure 4.53: Number of lexical items in agent lexicons for $W = \{R_1^2\}$.



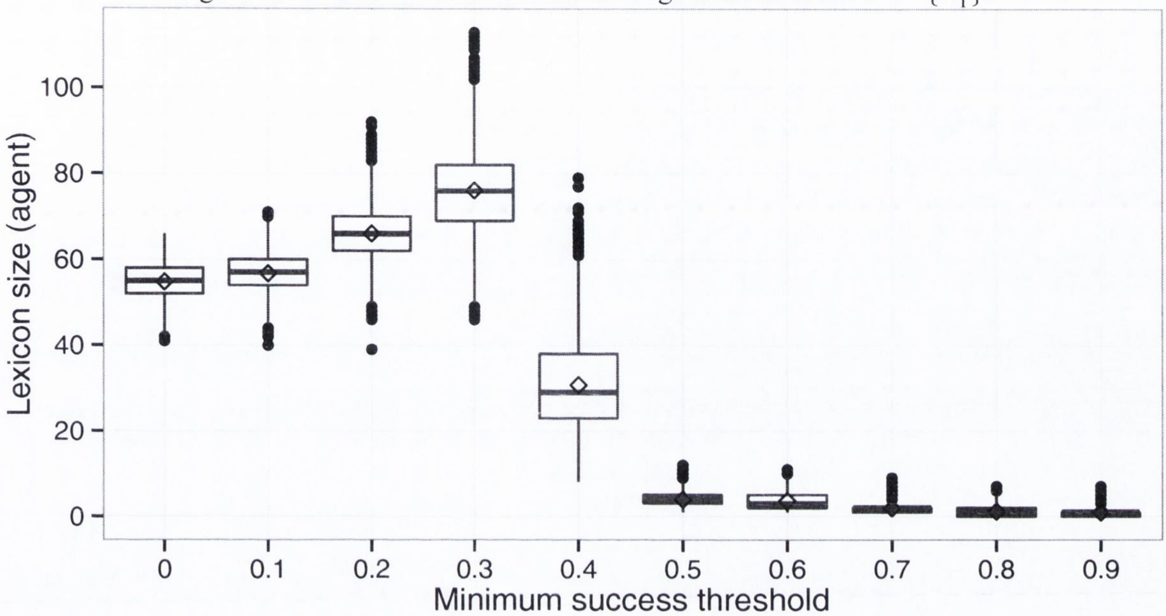
In experiments with an R^2 event, the size of agent lexicons has clearly grown even further in small groups by up to a factor of three, as can be seen from figure 4.53. While the increase itself was expected, an interesting feature in the small group experiments is that the lexicon size drops off significantly towards a more reasonable (requirement-wise, in world with six unique meanings) level at $s_{min} = 0.5$ already, as opposed to $s_{min} = 0.7$, observed in simulations with three agents.

Figure 4.54: Number of lexical items in the population lexicon for $W = \{R_1^2\}$.



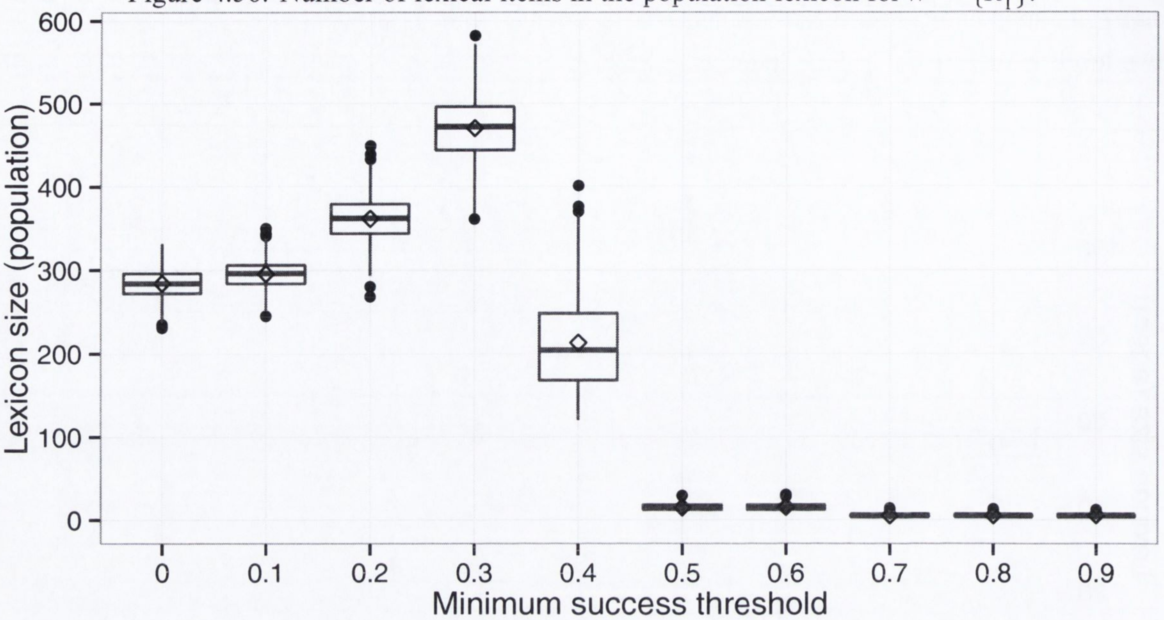
On the population level, the overall trend remains in simulations with both an R^2 and an R^3 event, as depicted by figures 4.54 and 4.56, respectively. In simulations with $s_{min} \leq 0.4$, populations of ten acquire on average four to five times the amount of mappings learned by agents from a triad. For $s_{min} \geq 0.5$, population lexicons of small groups actually exhibit a much smaller degree of bloating up, two levels earlier than observed in three-agent groups.

Figure 4.55: Number of lexical items in agent lexicons for $W = \{R_1^3\}$.



Finally, the trend continues in simulations with an R^3 event, as depicted by figure 4.55, with the lexicon sizes of agents doubling from the levels observed in simulations with three agents. For $s_{min} = 0.3$, agent lexicons including over 100 mappings are observed for the first time and this in a world where only ten different meanings need to be referred to. Just as was the case in small group simulations with an R^2 event, agent lexicons fall off to a more appropriate size from $s_{min} = 0.5$ onwards, two levels earlier than observed in triads. Notably, what did not occur in simulations with three-agent groups at all was for agents not to end up knowing at least a single mapping after the 1000 interactions. This appears to be possible (and in fact quite probable for $s_{min} \geq 0.8$) in simulations with small groups and either an R^2 or an R^3 event.

Figure 4.56: Number of lexical items in the population lexicon for $W = \{R^3_1\}$.



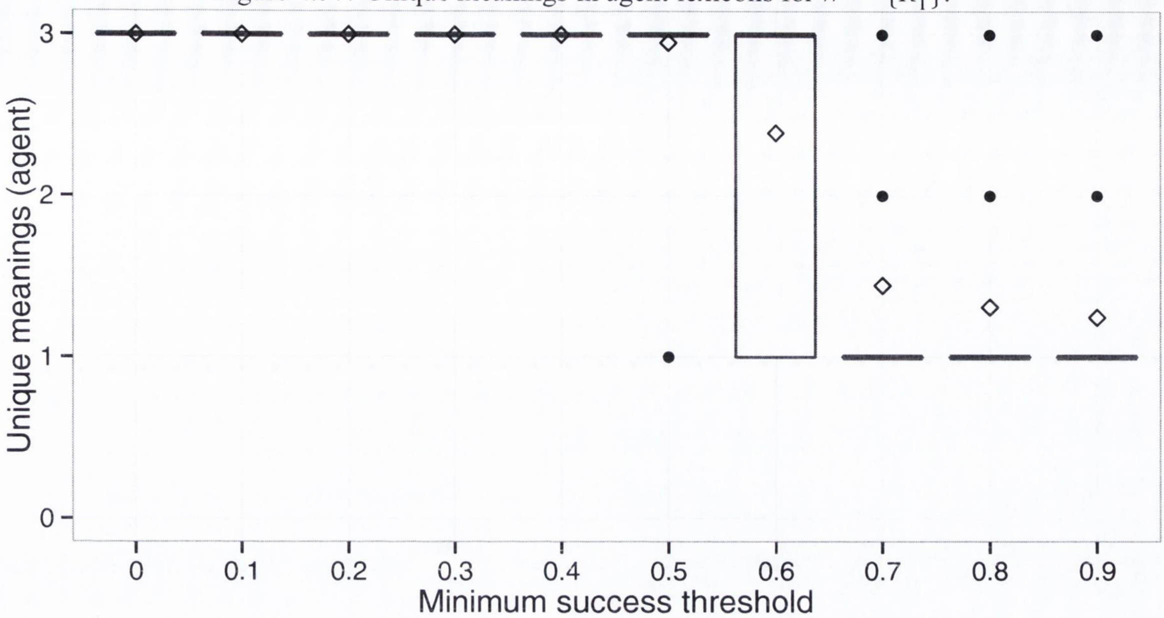
An interesting tendency that can be observed from the above figures is for the population lexicon size to gradually increase from $s_{min} = 0$ up to $s_{min} = 0.4$ for R^1 and R^2 and up to $s_{min} = 0.3$ for R^3. This might seem counter-intuitive at first, given that the reward for partially successful interactions is actually diminishing in the same space. Furthermore, considering the trend observed for $s_{min} \geq 0.5$, one would perhaps expect a similar, if not as pronounced trend to be observed for the successive levels of minimum success in the lower range as well. What is observed here, however, is exactly the orphaning effect, as outlined in more detail below.

The underlying difference between $s_{min} = 0$ and $s_{min} = 0.3$ e.g. is that for $s_{min} = 0.3$, the punishment for a completely failed interaction ($= -0.3$) actually outweighs the reinforcement for getting one out of three ($= 0.03$) or one out of two/two out of four ($= 0.2$) mappings right, which are some of the most likeliest scenarios. As a consequence, partially successful mappings experienced between Agent1 and Agent2 in an earlier interaction can get instantly removed from the lexicon of one of the agents, when Agent1 is involved in a completely failed interaction with another agent, thus not allowing Agent1 and Agent2 to consolidate the mappings in further interactions and forcing at least one of them to invent new ones, thus bloating up the lexicon of the other agent and, consequently, the overall lexicon of the population.

Unique Meanings

Looking at figure 4.57, depicting the number of distinct meanings in agent lexicons accumulated in simulations with an R^1 event in small groups, one can see a very significant difference to what was observed in triads, namely that the overwhelming majority of the agents knows a form for every mapping in the world only as far as $s_{min} \leq 0.5$. In fact, for $s_{min} \geq 0.7$, the majority of the agents only has a mapping for one of the world's meanings – the partition combining all of the components of the single event. This stands in stark contrast to simulations in triads, where all agents bar a few select exceptions managed to associated a form for every meaning even for the highest levels of minimum success threshold.

Figure 4.57: Unique meanings in agent lexicons for $W = \{R_1^1\}$.



When the lexicons of all agents in any given population are put together, the picture changes quite significantly, as shown by figure 4.58. In particular, the clear majority of populations have a lexical mapping associated with every available meaning in the population lexicon, even for $s_{min} \geq 0.7$. This mainly suggests that while many of the agents have failed to agree on lexical mappings with others, there are still some agents that experienced more success in their interactions and have thus contributed to a more complete looking population lexicon, when it comes to the number of expressible meanings.

Figure 4.58: Unique meanings in the population lexicon for $W = \{R_1^1\}$.

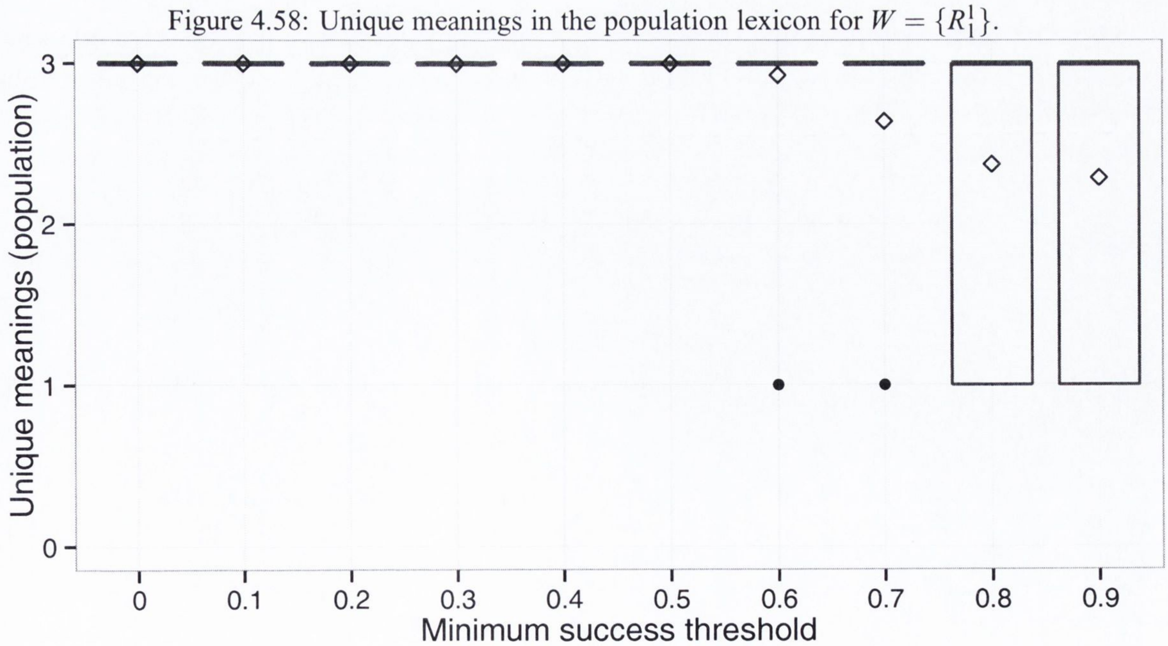
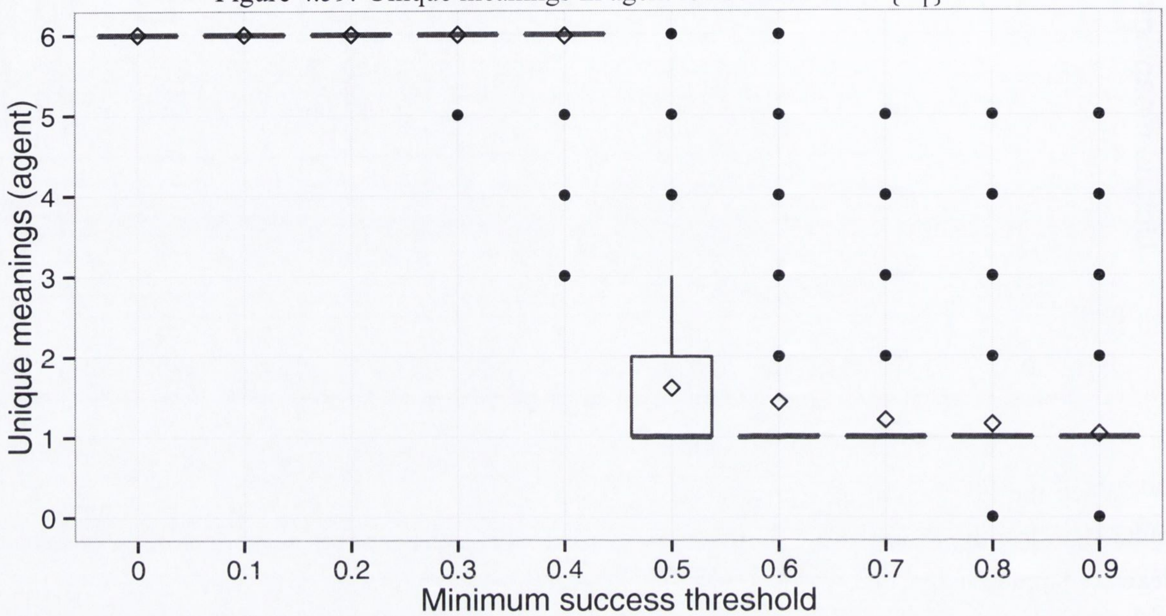


Figure 4.59 shows that in simulations with an R^2 event, a number of agents emerge that after around 1000 interactions do not have a single mapping in their lexicon and thus naturally do not have a single meaning represented in it. If one looks at the corresponding figures for simulations with three agents, one would see that every agent in every simulation ended up knowing how to express at least

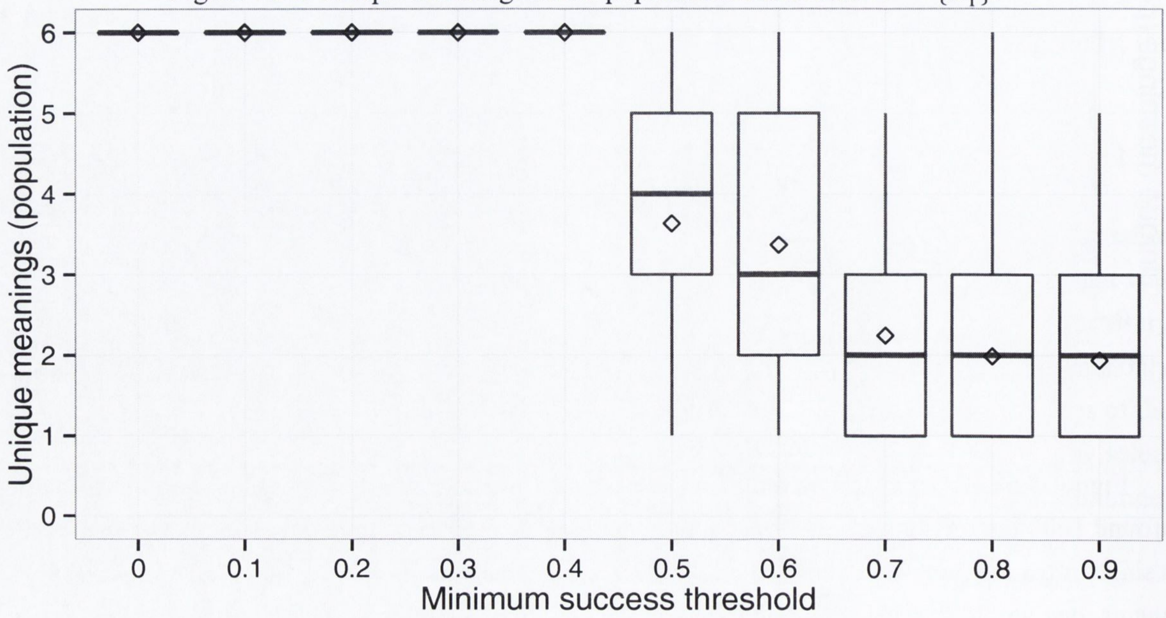
one meaning in the world, for all types of events and all levels of minimum success. Additionally, for $s_{min} \geq 0.5$, the majority of the agents only ends up having one meaning represented in their lexicons, further underlying the complexity of agreeing on lexical conventions through dialogue in a relatively large group of interactors.

Figure 4.59: Unique meanings in agent lexicons for $W = \{R_1^2\}$.



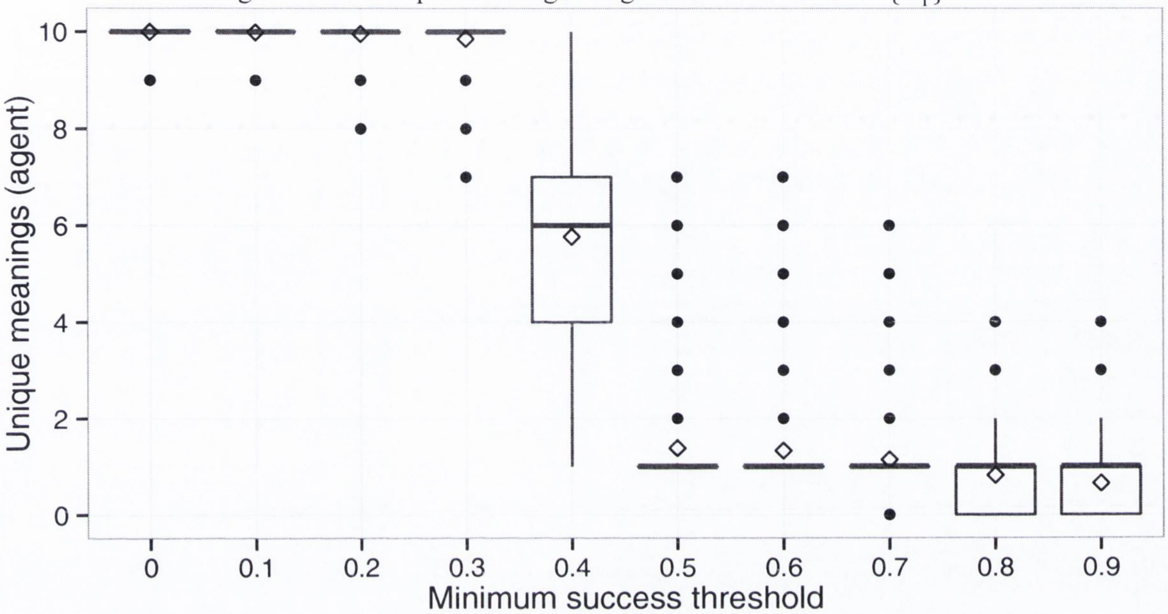
In contrast to the agent lexicons, the number of distinct meanings in the combined lexicon of each population for $s_{min} \geq 0.5$ is significantly higher, as shown by figure 4.60, ranging, on average, between two and four meanings, though reaching all six on a much more numerous number of occasions. Given that the majority of the agents only have one meaning available in their lexicons in the corresponding simulations, the only explanation for this trend is that different pairs of agents occasionally happen to agree on lexical mappings for different meanings, without being able to subsequently spread these around the group.

Figure 4.60: Unique meanings in the population lexicon for $W = \{R_1^2\}$.



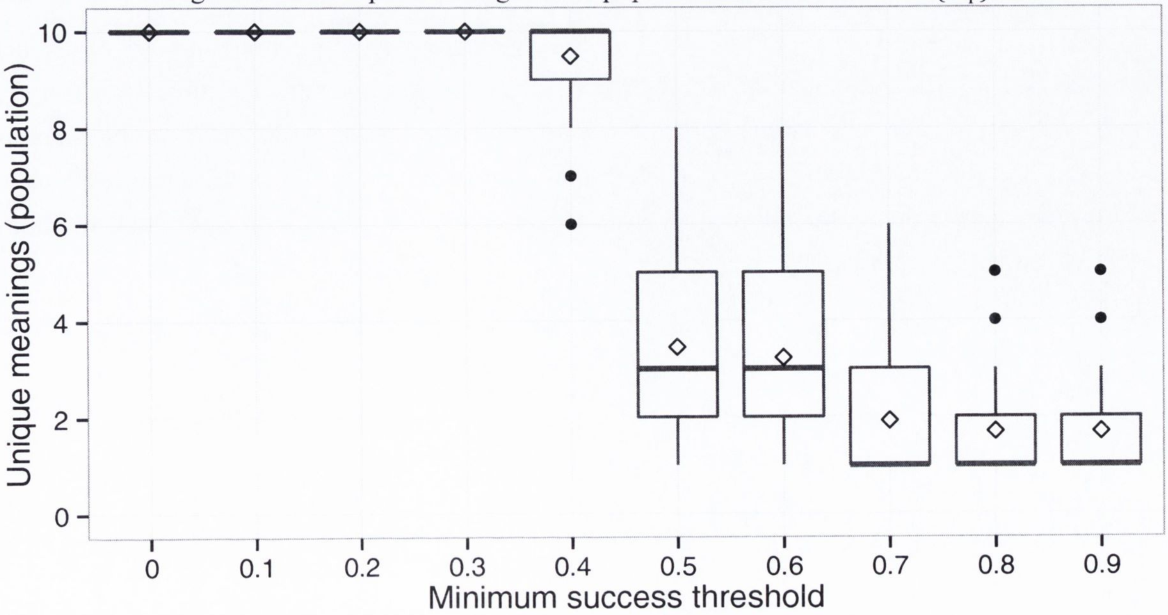
Finally, in simulations with the largest event, figure 4.61 shows that only for $s_{min} \leq 0.3$ does the majority of agents have all of their meanings represented in their lexicons. On the other side of the success threshold, i.e. for $s_{min} \geq 0.8$, the portion of agents without a single represented meaning becomes increasingly more sizeable as agents struggle with the increased number of potential meanings in every interaction.

Figure 4.61: Unique meanings in agent lexicons for $W = \{R_1^3\}$.



The trends on the population lexicon level, depicted by figure 4.62, are quite similar to those observed in simulations with an R^2 event. However, the combined lexicon effect becomes much less pronounced, with population lexicons averaging at just one meaning for $s_{min} \geq 0.7$, quite clearly as a consequence of many agents in the corresponding groups not being able to learn a single lexical mapping.

Figure 4.62: Unique meanings in the population lexicon for $W = \{R_1^3\}$.

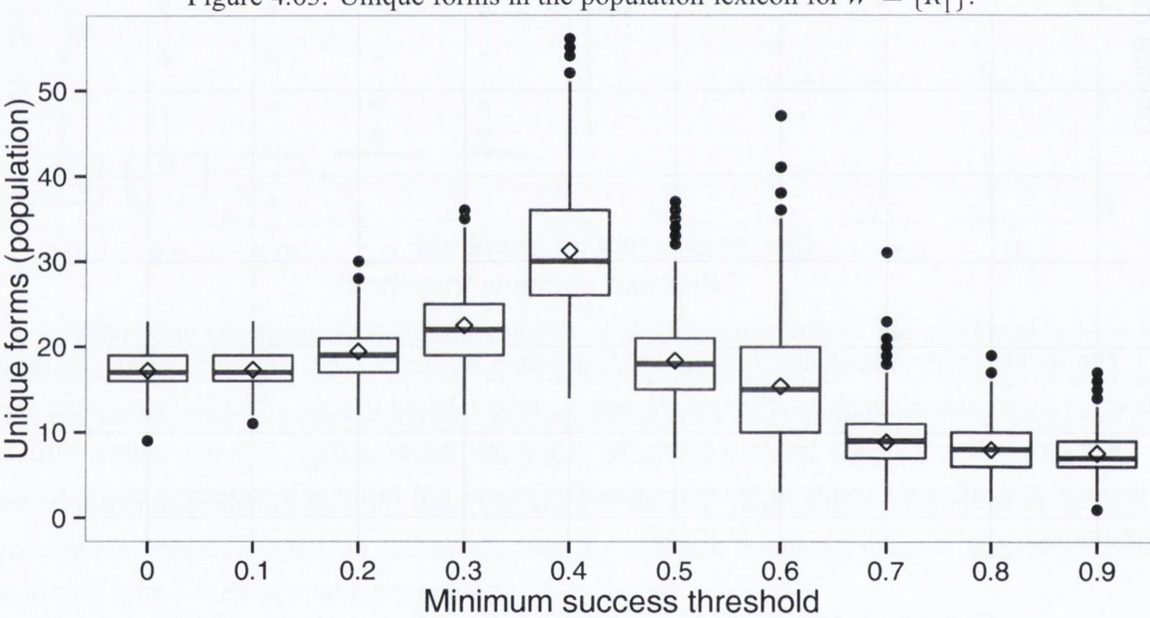


The reason for most of the above observations are the significantly reduced chances of speaking to the same agent in several consequent interactions, due to an extended pool of agents that any particular speaker has to pick a hearer from. Since the hearer selection is always made at random and does not depend on past interactions in any way, chances are high that the speaker will end up talking to different agents and all the good work (or luck, if one prefers) from past interactions with one of the other agents may be quickly undone by a single failed interaction with another partner.

Unique Forms

In experiments with agent triads, it has been observed that agents tend to learn a number of multiples more forms that is required to be able to express every meaning in their world. On top of that, for the two bigger event types, the overall lexicon of the population was shown to be significantly larger than that of an average agent, indicating a large amount of mappings with non-shared forms in individual agent lexicons that had multiple synonyms across the population.

Figure 4.63: Unique forms in the population lexicon for $W = \{R_1^1\}$.



Similarly to simulations with three-agent groups, the difference between the number of distinct forms learned by any individual agent (cf. figure 4.51) and the number of distinct forms in the combined lexicon of the population, depicted by figure 4.63, in a world with an R^1 event is not as striking as it is for larger events, even considering the increased size of individual agent lexicons in small group experiments. The only exception is provided by simulations with $s_{min} = 0.4$, which as discussed earlier, provide the best conditions for the accumulation of 'orphaned' forms and thus the corresponding bloating up of the distinct forms represented in the population lexicon.

Figure 4.64: Unique forms in the population lexicon for $W = \{R_1^2\}$.

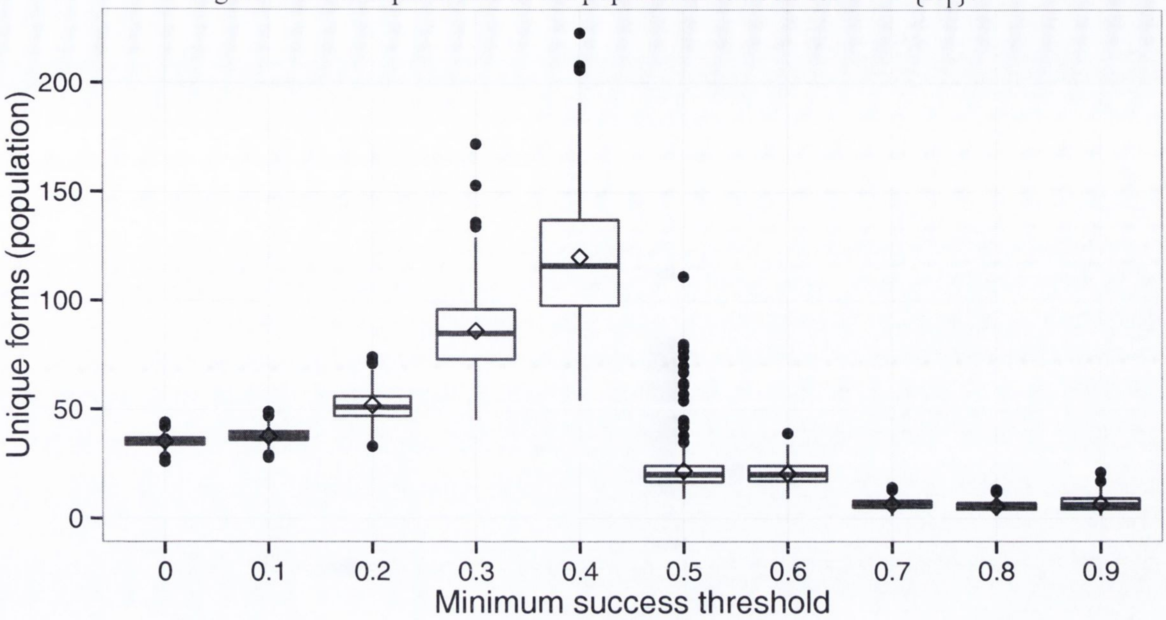
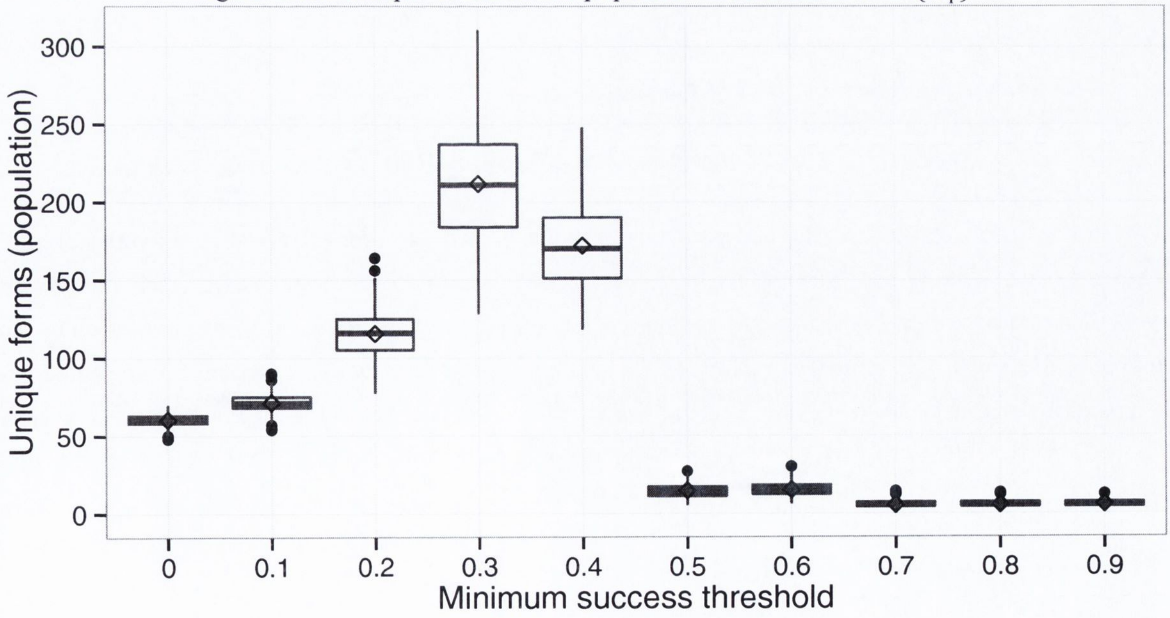


Figure 4.64 shows that the more distinct difference between the number of distinct forms on the agent and population level reappears in experiments with an R^2 event in small groups, just as it did in triads. While the factors of the increase are comparable, it is quite stunning to note that the number of distinct forms in the population lexicon has reached over 200 items in some cases for $s_{min} = 0.4$, again underlying the role that the level of minimum success plays on the emergent lexicon of the population even in a simple world consisting of six different meanings.

Figure 4.65: Unique forms in the population lexicon for $W = \{R_1^3\}$.



Finally, figure 4.65 suggests that, in experiments with an R^3 event, the overall number of distinct forms learned across all agents of a population is largest for $s_{min} = 0.3$, with the most 'lexicon-hungry' populations learning over 300 forms. In summary, looking at the number of distinct mappings, expressible meanings and used forms per level of minimum success, it would appear that having a success requirement of 50% or slightly more yields the most promising results, with any level just below 50% actually generating the seemingly most unmanageable lexical conditions for the agents.

4.4.2 Synonymy & Homonymy

As was the case with the basic lexicon properties, the analytic limits of synonymy and homonymy in agent and population lexicons do not change between triads and small groups. Accordingly, the following sections will largely focus on the evaluation of the experimental outcomes, referring back to the analytical section in the preceding sections where necessary.

Agent Synonymy

In a scenario where agents are not limited in the number of distinct forms that they can learn for only a fixed number of distinct meanings, it is clear that agent lexicon synonymy always has a chance of reaching the maximum level, i.e. having multiple forms assigned to each of the world meanings.

Figure 4.66: Agent lexicon synonymy level for $W = \{R_1^1\}$.

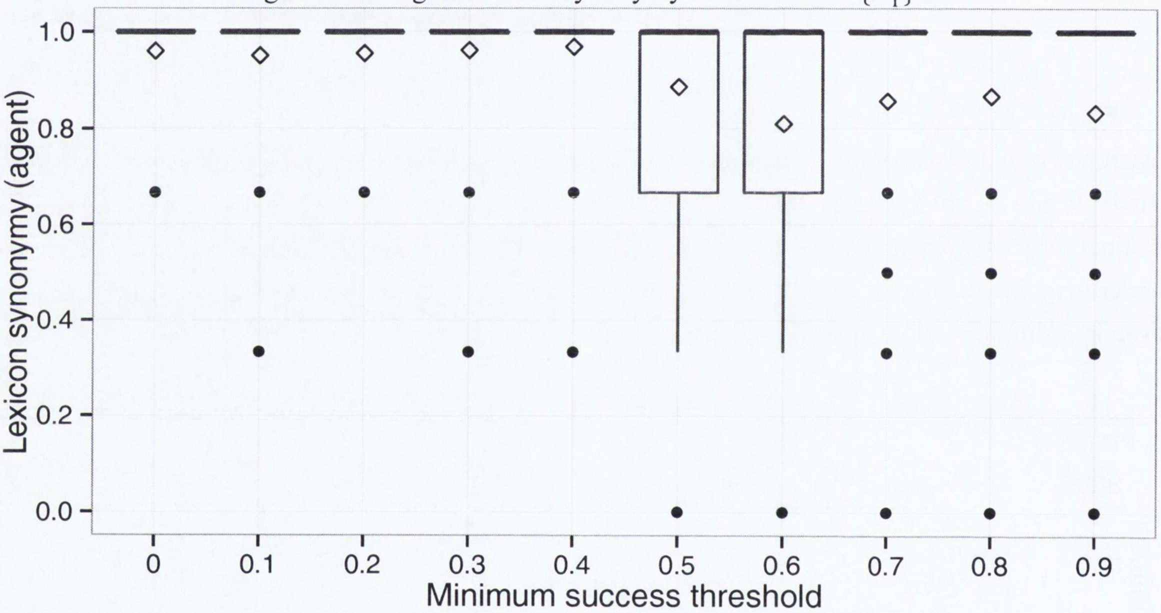
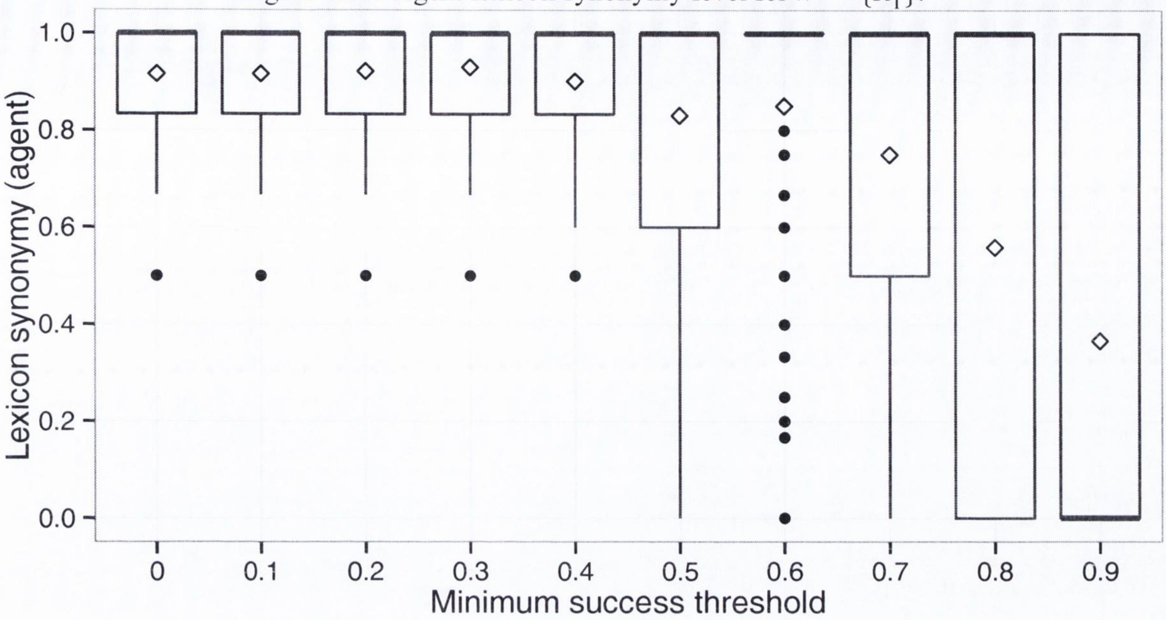


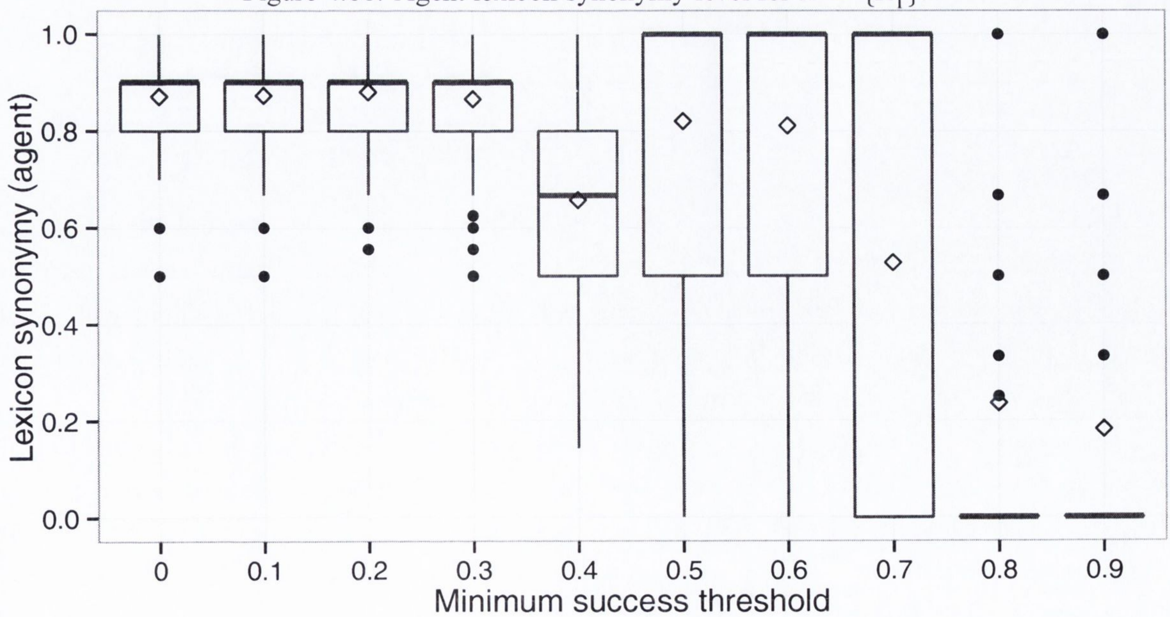
Figure 4.66 confirms this by showing that, in simulations with an R^1 event, the overwhelming majority of the agents have multiple form options associated with each of the three meanings, for every level of minimum success experimented with. This is a clear increase from simulations with three-agent groups where only a fraction of the agents observed full lexicon synonymy in their lexicons and a sizeable amount of the agents actually had no synonyms in their lexicons. As can be seen from the above figure, in small groups, only agents in simulations with $s_{min} \geq 0.5$ occasionally end up without a single synonym.

Figure 4.67: Agent lexicon synonymy level for $W = \{R_1^2\}$.



The trend continues in experiments with an R^2 event, with the majority of the agents again having multiple forms available for each of the six meanings for all levels of minimum success bar the highest, where the dominating outcome shifts from full synonymy to no synonymy. Notably, in simulations with three agents, having no synonymy was the prevalent result for three levels of $s_{min} \geq 0.7$. Admittedly, agents interacting in triads about this event never ended up without synonyms in their lexicon for $s_{min} \leq 0.4$, however, the lowest level that they were able to reach was having one in six meanings associated with multiple forms, compared to a much higher 50% synonymy observed for the same levels of minimum success in small groups.

Figure 4.68: Agent lexicon synonymy level for $W = \{R_1^3\}$.

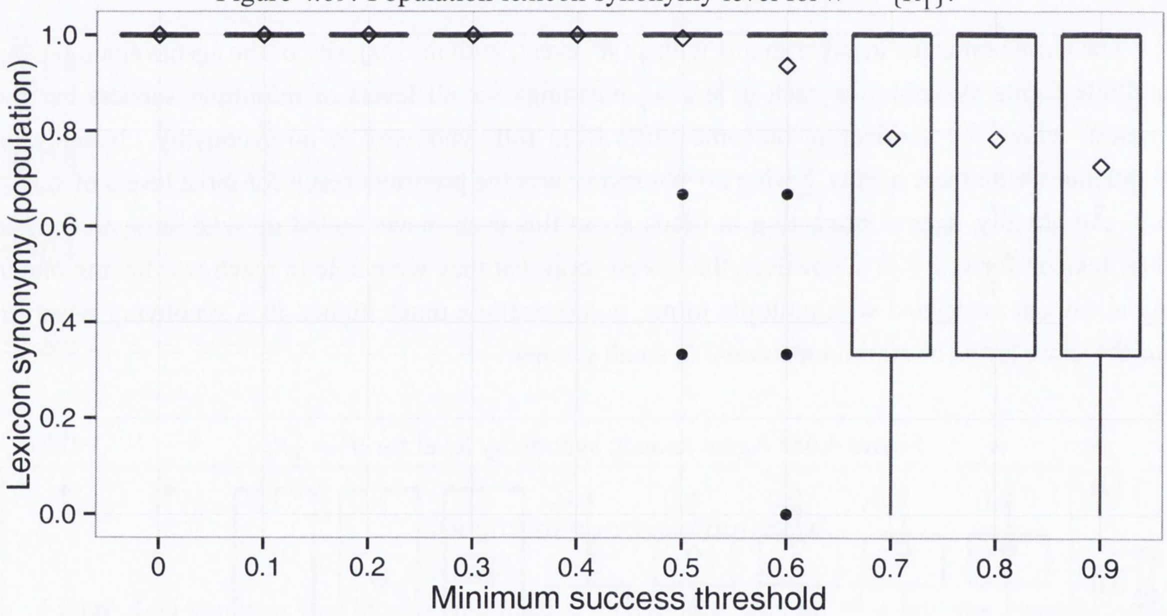


In simulations with an R^3 event, most agents observing a lower level of minimum success $s_{min} \leq 0.4$ do not actually reach full synonymy, though they are markedly closer to it than the corresponding agents in three-member groups. In fact, average lexicon synonymy goes slightly down for $s_{min} = 0.4$, but then goes up to the maximum level again for $0.5 \leq s_{min} \leq 0.7$. Finally, for $s_{min} \geq 0.8$, most agents do not have any synonyms in their lexicon, mirroring precisely the situation in triads. However, this is the only parameter combination where small groups end up having similar results to those of agents from a group of three. In general, the tendency is clearly for agents in small groups to have a significantly higher amount of synonymous forms in their lexicons.

Population Synonymy

As a direct consequence of increased lexicon sizes on both the agent and the population level, the levels of synonymy in the population lexicon have increased accordingly. Figure 4.69 shows that, in simulations with an R^1 event, synonymy at the population level has reached its maximum in the majority of conducted simulations for each level of minimum success. In fact, population synonymy is fixed at 1.0 for $s_{min} \leq 0.4$, owing to the very loose success requirements in these simulations which allow for every other form to be incorrectly interpreted and assigned to different meanings between agents and still obtain a positive reinforcement.

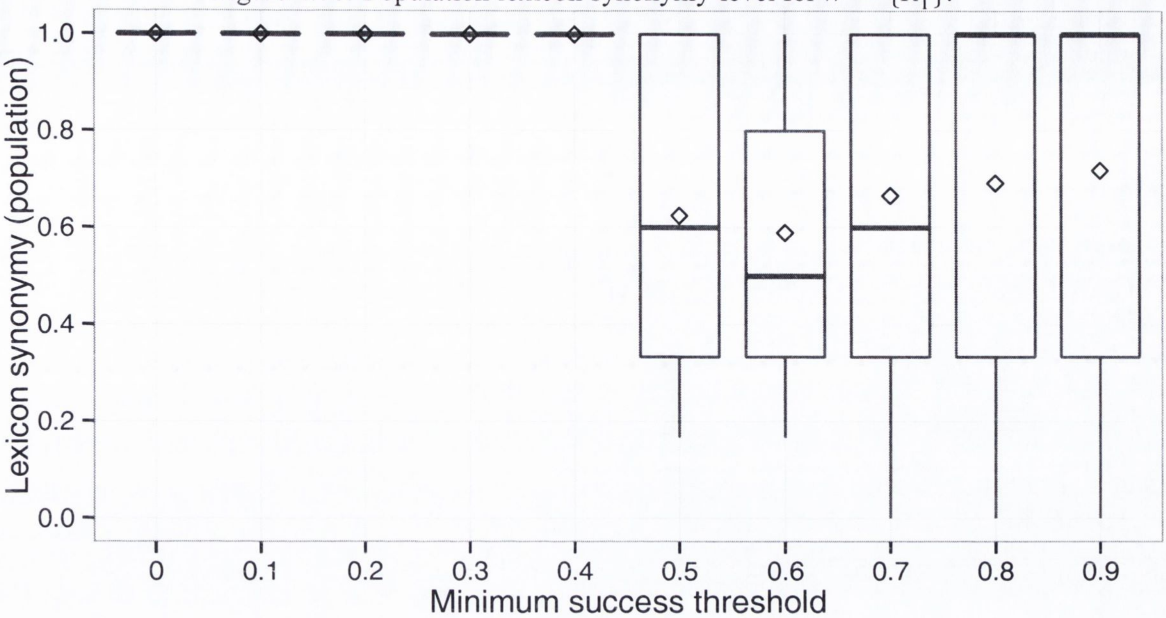
Figure 4.69: Population lexicon synonymy level for $W = \{R_1^1\}$.



In simulations with an R^2 event, the results of which are visualized by figure 4.70, population synonymy for the lower range of minimum success actually does not differ from what was observed in triads, as it had reached 100% in those experiments already. However, in simulations with small groups, population synonymy has actually dropped off for $0.5 \leq s_{min} \leq 0.7$. A similar tendency is observed in simulations with an R^3 event for $0.5 \leq s_{min} \leq 0.6$, as shown by figure 4.71.

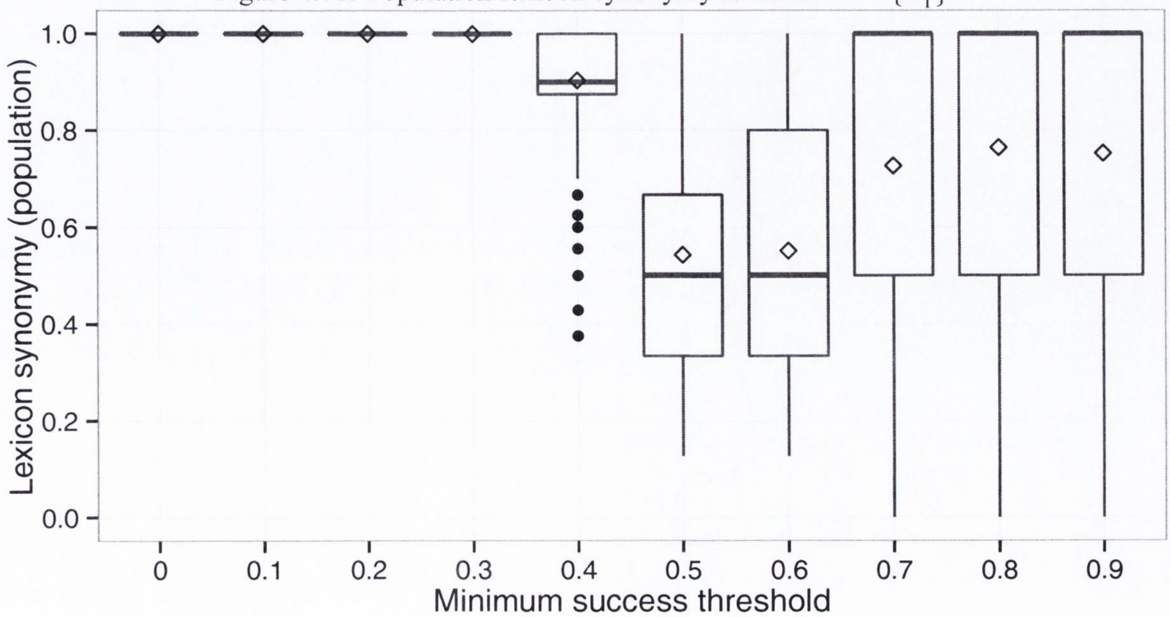
The reason for the lower synonymy levels in these simulations is that even though agents tended to end up knowing a form for one meaning only, different agents ended up knowing how to express different meanings, in many cases sharing the mapping with one or more other agents. Although at least half the meanings still ended up having multiple forms associated with them by different agents, this relatively high levels of minimum success ensured that some meanings actually need to belong to shared mappings, while at the same time not eliminating too many mappings, as is the case for the highest levels of minimum success, as described below.

Figure 4.70: Population lexicon synonymy level for $W = \{R_1^2\}$.



For the higher levels of minimum success, the population synonymy has gone from mostly zero observed in three-agent groups, to mostly full in small groups, in simulations with both an R^2 and an R^3 event. Looking back at the figures in section 4.4.1, one can see that in these simulations, only one meaning is known by the agents most of the times, while the population lexicon has a much larger number of forms. What happens here is that pairs of agents are sometimes able to understand each other's utterances, but are unable to agree on the corresponding mappings across the population to the extremely high success requirements. Consequently, individual pairs of agents end up sharing a mapping for one of the meanings, but since these mappings are different across agent pairs, the population lexicon exhibits a high level of synonymy.

Figure 4.71: Population lexicon synonymy level for $W = \{R_1^3\}$.



In summary, it can be seen that changes in the number of unique meanings and forms the lexicons can have very diverse effects on the corresponding synonymy levels, with a higher number of forms not necessarily implying higher synonymy levels and vice versa. However, nearly all of the agent and population lexicons in simulations with ten agents have synonymy levels that are significantly higher than those of human languages, indicating a deteriorating trend in terms of the comparability of the emerging system to human languages.

Agent Homonymy

Agent homonymy levels remain at zero in experiments with ten agents as the principles underlying homonymy generation (or lack thereof) have not changed with the increase in population size.

Population Homonymy

On the population level, homonymy acts quite similarly to what was observed in simulations with triads. However, there is a slight deviation in terms of the highest levels of population homonymy reached for the lower levels of minimum success and in the rate of its decline with the successive increase of s_{min}, as described below.

Figure 4.72: Population lexicon homonymy level for $W = \{R_1^1\}$.

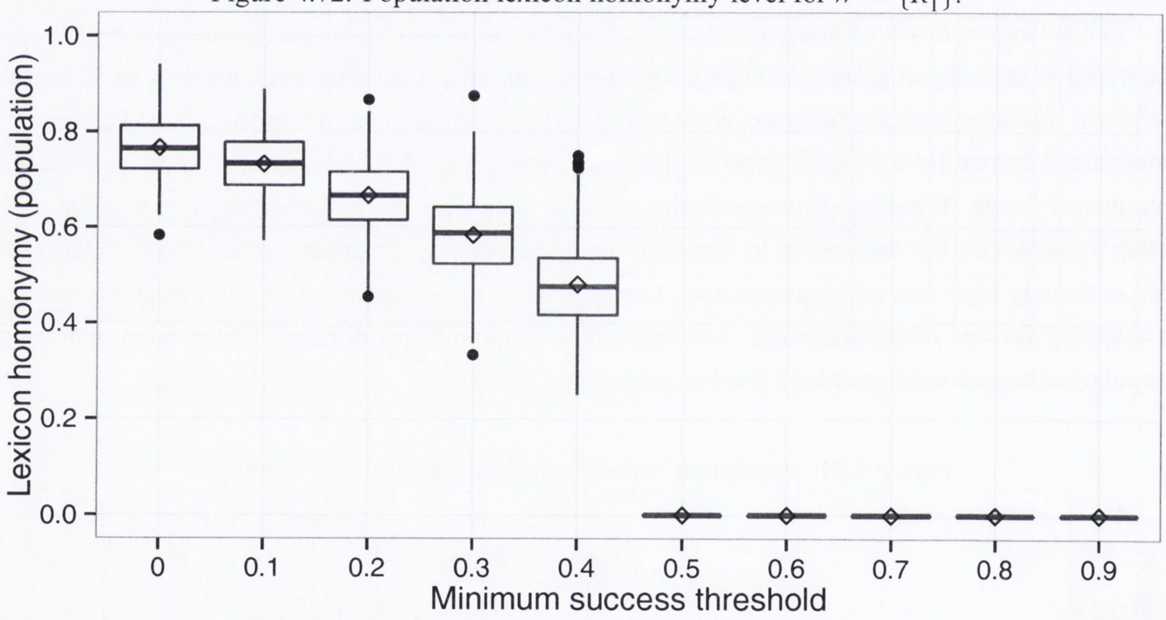
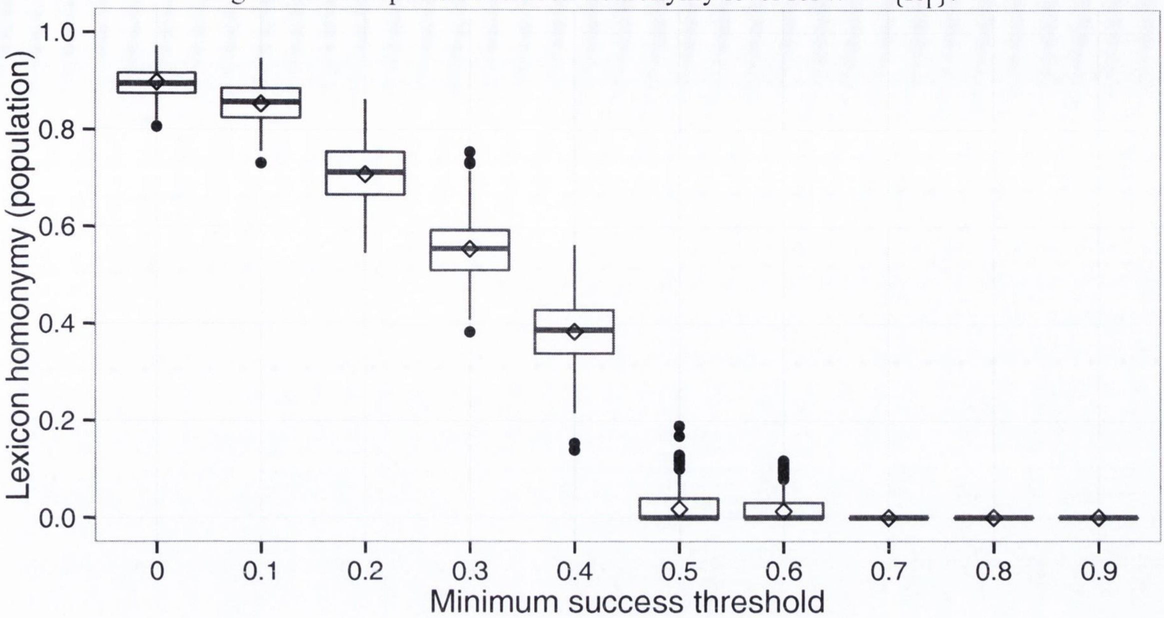


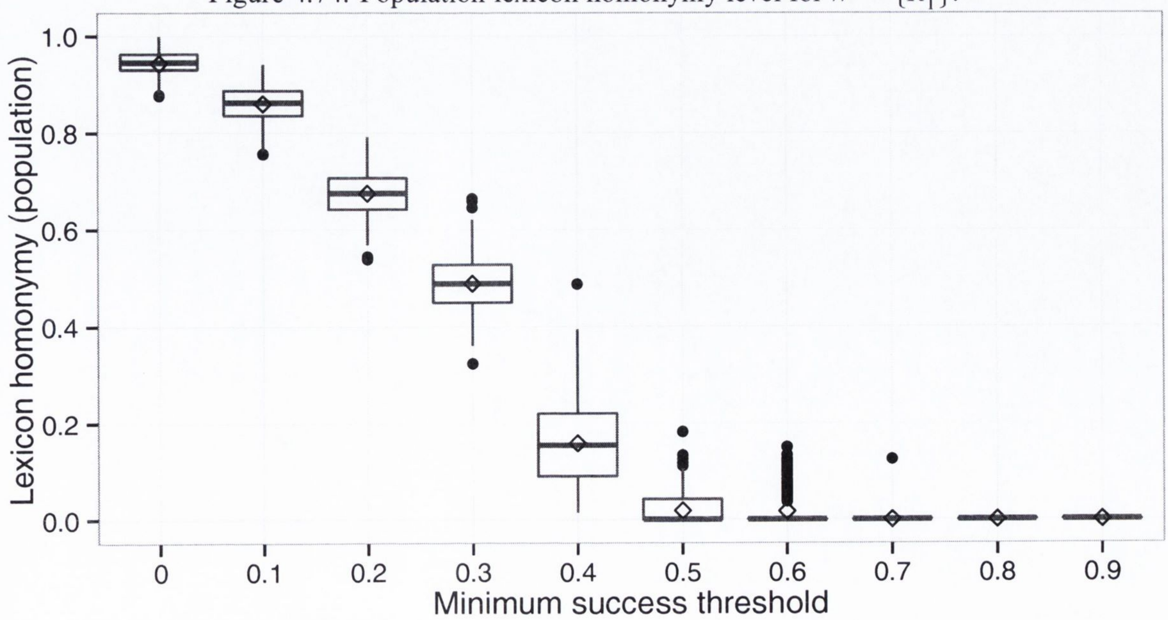
Figure 4.72 shows that, in simulations with an R^1 event and ten agents, population homonymy is on average up to 0.2 decimal points higher than in comparable simulations with three agents for $s_{min} \leq 0.4$. Furthermore, for $s_{min} \geq 0.5$, population homonymy remains at zero, just as was the case in three-agent groups.

Figure 4.73: Population lexicon homonymy level for $W = \{R_1^2\}$.



Figures 4.73 and 4.74, depicting population homonymy at the end of experiments with an R^2 and an R^3 event, respectively, illustrate that, for $s_{min} \leq 0.2$, population homonymy in small groups is again slightly higher than it was in triads. However, for the following levels of minimum success, population homonymy is actually lower in small groups, reaching zero in most simulations with $s_{min} \geq 0.5$ already (as opposed to 0.6 or 0.7 observed in triad experiments). The reduced population homonymy levels for $0.3 \leq s_{min} \leq 0.4$ is a direct consequence of the increased number of distinct forms in the population lexicon in these experiments, as described in section 4.4.1 above. On the other hand, the population lexicon of agents is drastically reduced for $0.5 \leq s_{min} \leq 0.6$, basically only containing correctly interpreted and agreed on forms and thus a very low, if any, level of population homonymy.

Figure 4.74: Population lexicon homonymy level for $W = \{R_1^3\}$.

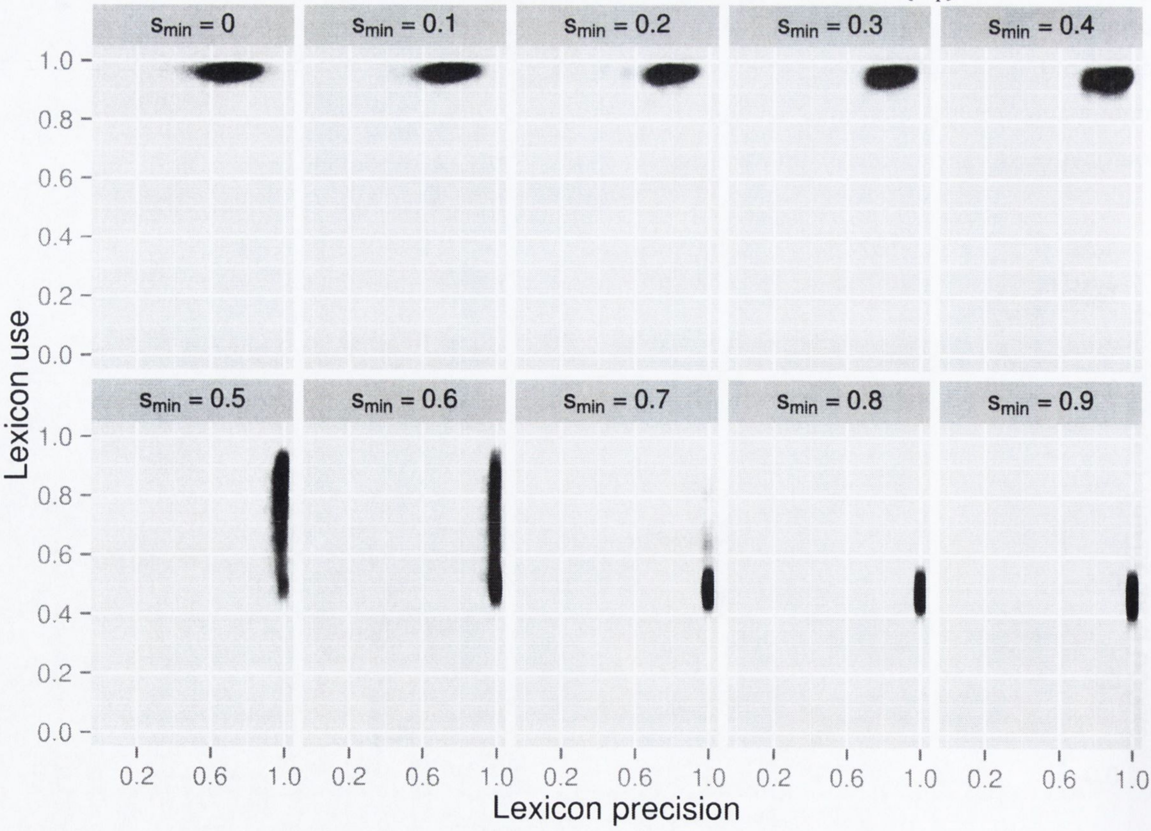


While the trends observed in the latter two of the three figures above might indicate that population homonymy is actually improving for $s_{min} \geq 0.3$ in a population of ten agents, the figures are a great example of the importance of considering a number of different lexical properties before making any qualitative conclusions. In particular, while the relative level of population homonymy was indeed on decline, it was not due to an improved ability of agents to acquire a lexicon. Instead, for $0.3 \leq s_{min} \leq 0.4$, the reason was a significant explosion in the number of distinct forms acquired by the agents, while for $0.5 \leq s_{min} \leq 0.6$, it was exactly the opposite. As can be concluded, having a low level of population homonymy alone does not imply that agents have managed to learn a less ambiguous lexicon, but could actually just as well mean that they have learned either too few or too many lexical items.

4.4.3 Communicative Success

Having performed an analysis of the types of lexicons that emerge in groups of ten agents, the question that remains is how the agents equipped with these lexicons actually end up performing in their interactions. Some of the effects can be easily predicted by e.g. the average counts of distinct mappings, meanings and forms in agent lexicons, where not having enough of lexical items will clearly prevent the agents from reaching a high level of lexicon use. However, the effects of other properties such as lexicon synonymy and homonymy are perhaps slightly less predictable, it remaining an open question if e.g. a high synonymy level could actually be beneficial for an agent when trying to understand a number of different agents. Hopefully, the following analysis of lexicon use and lexicon precision should shed some light on these issues.

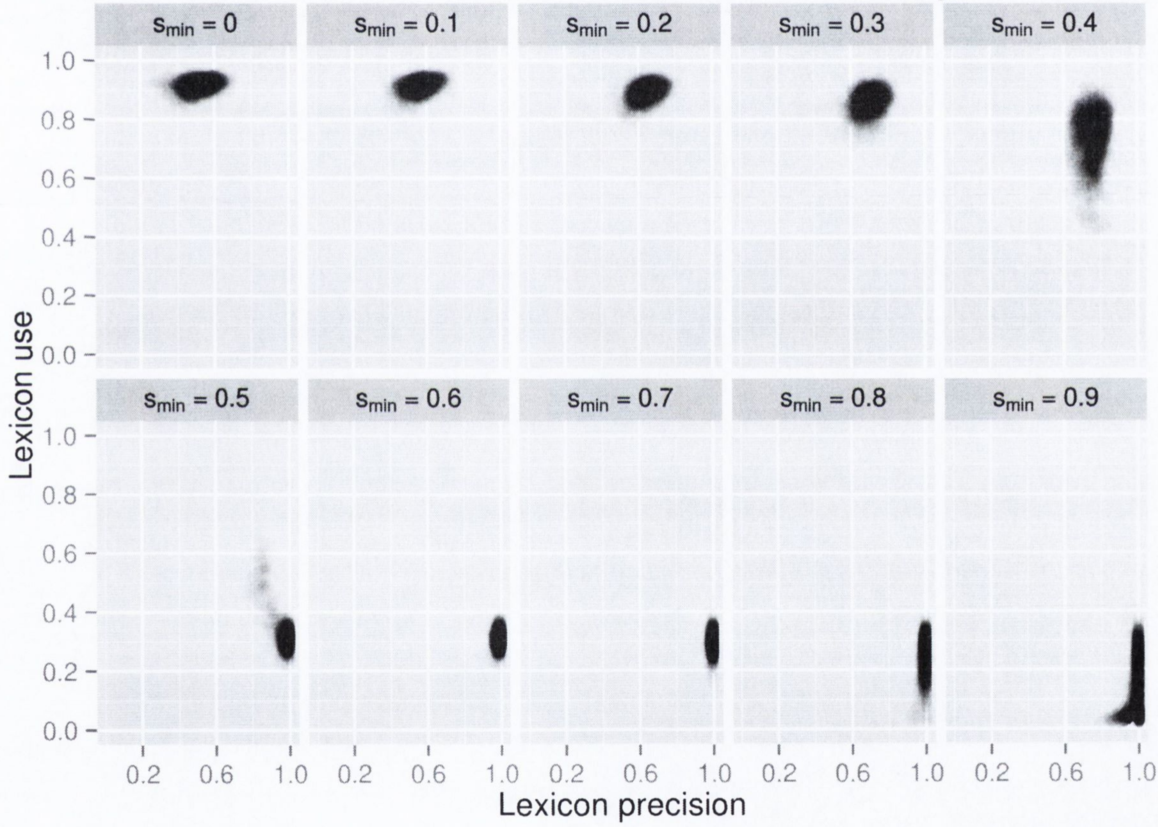
Figure 4.75: Lexicon-based communicative success for $W = \{R^1_1\}$.



In simulations with ten agents and an R^1 event, the interaction between lexicon use and lexicon precision, depicted by figure 4.75, is largely comparable to that observed in three-agent groups, although with some modifications. For $s_{min} \leq 0.4$, the overall trend is very similar, with lexicon use nearly reaching 100% and lexicon precision improving towards the higher levels of minimum success. The only thing to note here is that lexicon use is not actually fixed at 100% as it was in triads. This is mostly due to the larger number of agents naturally taking more time to experience every lexical item from every other agent, from which point on lexicon use would be guaranteed.

Another difference between experiments with triads and small groups is in the degree of the drop in lexicon use observed for $s_{min} \geq 0.5$. While in groups of three, lexicon use gradually reduced between these levels of minimum success, still sometimes reaching 100% even for $s_{min} = 0.9$, in small groups, lexicon use drops to around 50% by $s_{min} = 0.7$ and remains at that level for all subsequent values of s_{min}. The reason for this are the reduced chances of all ten agents actually agreeing on a set of reliable lexical items under the strict success conditions, something that clearly persists throughout experiments with all three event types.

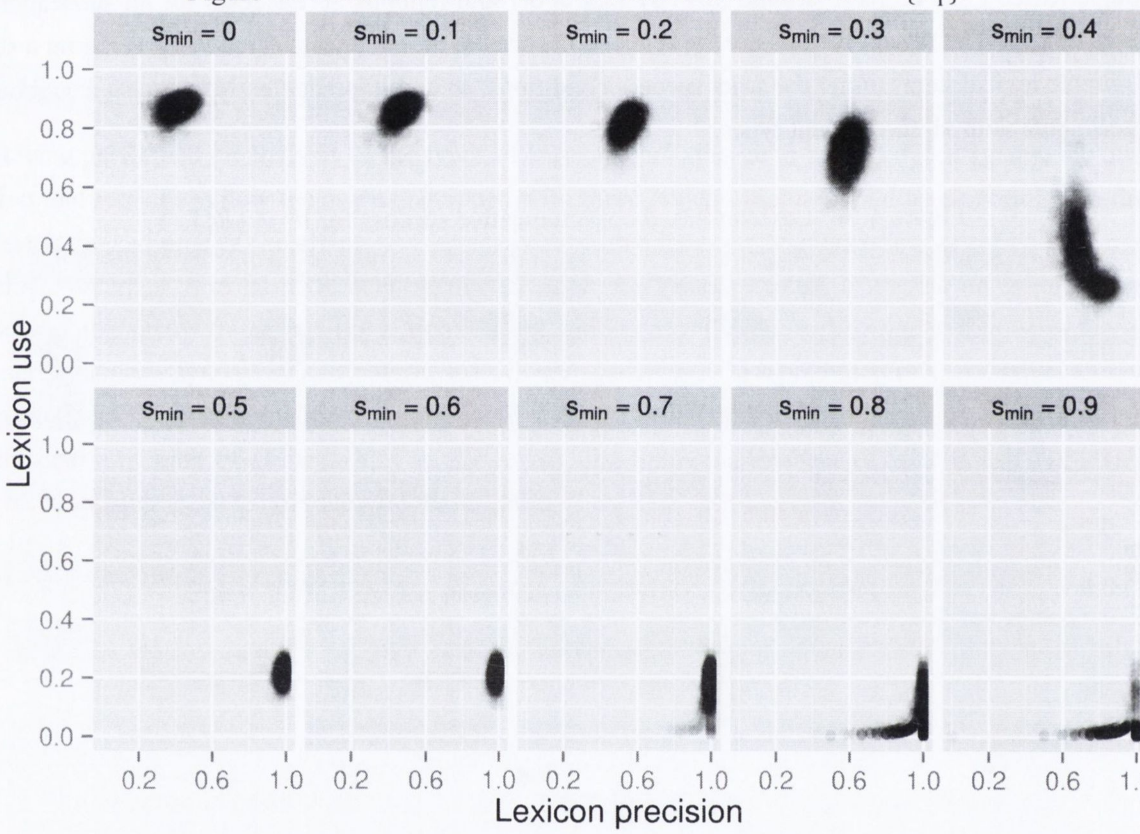
Figure 4.76: Lexicon-based communicative success for $W = \{R_1^2\}$.



Similarly to simulations with an R^1 event, in experiments with an R^2 or an R^3 event, lexicon use is highest for the five lower levels of s_{min}, though not quite reaching 100% in any of the experiments, as depicted by figures 4.76 and 4.77, respectively. Again, the transition from nearly constant lexicon use with varying precision to a much more occasional lexicon use, but a very high precision occurs slightly earlier in small groups, with the population size making its effect in simulations with a higher success requirement.

One effect that has not been present in graphs depicting the interaction of lexicon use and precision in simulations with two or three agents, that is quite evident in simulations with ten agents and either an R^2 or an R^3 event, is a further decline in lexicon use for the highest levels of s_{min}, also accompanied with a slight decrease in lexicon precision, which is rather uncommon. As can be seen from figure 4.77, for $s_{min} = 0.9$, lexicon use drops to just above zero in most experiments, with lexicon precision varying between a very high 1.0, which is more common for such low levels of lexicon use, and a very mediocre 0.5. The drop is a direct consequence of agents not being able to develop any sort of a reasonably-sized lexicon, as was discussed in section 4.4.1, with lexical items only emerging sporadically between individual pairs of agents and then quickly disappearing as subsequent interactions with other agents do not yield sufficient success.

Figure 4.77: Lexicon-based communicative success for $W = \{R_1^3\}$.



In summary of the simulations with ten agents, it can be stated that the optimal combination of lexicon use and precision results from a slightly lower success requirement than in a population of just two or three agents. In particular, for the largest event type, the level of minimum success at which the agents seem to perform best is somewhere between 0.2 and 0.3. This is despite the population lexicons in these experiments reaching over 500 items at times, with agent synonymy averaging at 90%, population synonymy fixed at 100% and population homonymy hovering at around 50%. What appears to be the critical aspect here is that for $s_{min} = 0.4$, the levels of synonymy and homonymy are significantly lower, yet the average agent only has a word for six of the ten meanings in his lexicon, which is clearly having a very negative impact on the communicative performance. It is undeniable that learning an average of 75 different lexical forms for any given meaning from the world would have been an unnecessary and unacceptable burden on the storage capacities of early hominids, however, from the communicative point of view, agents appear to be able to deal quite efficiently with the extra forms during their interactions even with a very simple frequency-based word selection model.

4.5 Summary of Fundamentals

The experiments presented in this chapter deal with a very basic configuration of the LEW that included only one event consisting of between one and three arguments and a small population of up to ten agents per simulation run. The main goal of conducting these experiments was to determine if there are certain analytical limits imposed on the different properties of the emergent agent and population lexicons and how they are influenced by the varying success requirements during interactions.

As was presented in section 4.2 that discussed a version of the model with a population of just two agents, the amount of communicative success imposed on the agents as a requirement for obtaining a payoff and thus learning a lexical item does indeed have a significant effect on almost all of the lexical properties. In particular, it sets a limit on the size of the lexicon, the number of distinct meanings and forms represented by it and, consequently, the maximum level of agent and population synonymy and homonymy present in these lexicons. In some cases, the analytical limits did not actually impose any restrictions on the system, as is e.g. the case for the level of population synonymy for $s_{min} \leq 0.4$, which, regardless of the event type, could theoretically take up any value between 0 and 1 in the conducted experiments.

While the analytical discussions in section 4.2 were meant to determine what is in fact possible in the simulations conducted with the help of the LEW, the experiments themselves provided an overview of the actual probability distributions of the values taken up by the observed lexical properties. The results of these experiments indicate that a level of minimum success around 0.5 results in the most optimal combination of lexicon use and precision, with all of the meanings being represented in the agents' lexicons and relatively moderate levels of synonymy and homonymy on both the agent and the population levels.

Having gone through the simplest possible version of the model, the following sections turned to simulations with an increased number of agents, namely three in section 4.3 and ten in section 4.4. Right at the beginning of section 4.3 it has been established that the quantitative limit imposed on the lexicons of agents in a population of two disappears with the addition of at least one further agent. In effect, agents could now keep on infinitely accumulating lexical items, given a certain alteration between speaker-hearer pairs as well as successes and failures.

The above was directly reflected by the simulations conducted with the different groups of agents, with the population lexicons growing up to 100 items in some of the three-agent groups and up to 600 items in ten-agent groups, after around 1000 interactions per agent. The increased lexicon size had a naturally strong effect on the lexicon synonymy levels, considering that the amount of meanings was fixed at three, six and ten, respectively, depending on the selected event type. Additionally, increasing the number of agents also clearly affected the chances of agents agreeing on a set of shared lexical items, with too few mappings learned in experiments with a high level of minimum success and too many learned if the success requirements were loosened too much.

In particular, the optimal level of minimum success shifted down to 0.4 in simulations with three agents and to around 0.25 in groups of ten. This decline was predicted by Bachwerk (2011), where it was shown that, in a population of ten agents and conversing about a large number of events, the optimal level of minimum success is determined by the average guessing rate of agents, which for R^3 is around 0.25. The fact that, in simulations with two and three agents, the optimal level of minimum success was much higher, can be explained by the low variety between agent lexicons, resulting in quick agreement on lexical items after just a few successful interactions.

Finally, when comparing the lexicons resulting from the conducted simulations to those of real languages, the synonymy and homonymy levels appeared to be quite similar for certain levels of the minimum success threshold in simulations with two or three agents, mainly around $s_{min} = 0.5$. However, in experiments with ten agents, both properties of the emerging lexicons reached significantly higher levels than their real-language counterparts for almost all levels of s_{min}.

A fundamental analysis of any computational model is essential in building up the understanding of its main effects and interactions. In effect, it allows one to correctly interpret the results of the more complex simulations performed with the model, which is normally the real goal of building the model in the first place. The analysis performed in this chapter can be thus seen as an evaluation of the model itself, but less so of the language acquisition process in a realistically complex setting. The following chapter will draw on the conclusions presented in this chapter to analyse three advanced social configurations and their potential role in the history of language evolution.

Chapter 5

Advanced Configurations

5.1 Introduction

Computational models of social behaviour are normally created with the purpose of investigating complex interactions between a multitude of parameters in a fully controlled simulated environment. In addition to that, a computer model also allows one to abstract over some of the features of the real world that are either not central to the research issue at hand, or are simply too complex to be represented in full detail in the model. The unfortunate consequence of introducing abstractions into a model is that it quickly becomes subject to criticism from scientists working hands-on in the field that the model is situated, e.g. linguists in the current case. Therefore, it is important to make it perfectly clear at this point that a mathematical or a computer model is not meant to create a perfect representation of the real world (at least for now), but rather to aid the field researchers by providing them with often hard (or simply impossible) to obtain empirical evidence, thus providing them with some guidance and in no way trying to replace their invaluable work.

Having said that, it is also important that computer scientists not only look at the very basic, and largely unrealistic, configurations of their models, but also try to extend the parameters to values that would result in a closer representation of the real world. In the previous chapter, the general fundamentals of the LEW model and the corresponding Lexicon Acquisition Game have been examined with the goal of estimating some analytical limits and probabilities of the various properties of the emergent simulated lexicons. This exercise was important in order to understand what the model is capable of in theory; however, the next important step is to apply the model to some of the more realistic scenarios of language evolution.

First of all, the number of different entities and events that can be perceived by the agents in their worlds is increased from one event and three entities to 50 instances of 10 events populated with 25 different entities that are admissible as arguments for any of the events. While this increase is still far away from the literally endless number of distinct atoms surrounding us in the real world, it could be argued that the number of information units that were truly critical to the early hominids, and thus were relevant to them in obtaining some sort of payoff in terms of finding a food source or avoiding a predator, was actually quite limited. Furthermore, it is hard to imagine that the early inventors of language would start talking about literally everything that surrounded them from the very beginning, but rather that they built a language up from interacting about the essential things to everything else later on.[1]

[1] Just like a child who starts with 'mummy', 'daddy' and then moves on to the less important things in life.

Despite the increase in the size of the agents' world, the number of interactions per agent was kept at 1000 on average. Admittedly, giving the agents more time to agree on the best lexical conventions for expressing all the meanings from their surroundings would certainly have increased their chances of communicative success. On the other hand, it seems rather implausible that early humans would spend too much time trying to communicate with the help of a new approach, such as the human language, if it did not bring at least some rewards pretty early on. Having said that, no claim is being made here that early humans would not be satisfied with a fairly moderate payoff from their early interaction attempts (after all, anything is better than nothing). Accordingly, for the experiments presented in this chapter, the level of minimum success that serves as a threshold for obtaining a payoff was set at just over a third of an utterance, i.e. $s_{min} = 0.35$.

The main focus of this chapter though is placed on different social configurations of agent populations, with three of these being presented in the following sections. In section 5.2, ten agents will be divided into five subgroups of two, with different interaction constraints imposed on the agents within these pairs. In the first part of the first experiment, a varying intra-group communication rate is set, which defines how often the agents will talk within their pair and how often they will talk to other members of the population. The motivation behind this setting is that agents have been observed to be much more efficient at evolving a language in pairs and less so in slightly bigger groups. The question is then if a combination of the two would solve some of the problems of the homogeneous groups. The second batch of simulations addresses the same question from a slightly different perspective, namely by letting agents interact in pairs for a specified period of time before switching the pairings, repeating the process until every agent has been paired up with every other agent. The results of both these configurations will be also compared to the usual groups of both two and ten agents. Afterwards, section 5.3 will look at how so called 'friendships' could be established naturally in a population of ten agents by dynamically adjusting one's interaction rates with other agents based on the success in previous interactions. Section 5.4 will then summarize the three experiments and discuss the effects that different social configurations have on the lexicon acquisition process.

5.2 Population Subgroups

The notion that social groups of one type or another play a central role within the evolution of the hominid species as such as well as the emergence of a communicative system like a proto-language in particular is apparent from a variety of evolutionary theories and modelling approaches. From the anthropological point of view, it has been repeatedly suggested that the emergence of language is strongly connected with the increase of hominid group sizes and the arguably directly related neocortex ratio increase between 500,000 and 250,000 years ago (see Aiello & Dunbar, 1993). Being unparalleled in any other species, this evolutionary change has become the focal point of several theories on the emergence of language.

Admittedly, an increase in group size alone does not in any way guarantee the emergence of language, nor does it really support it in any way to begin with, given the complexities of establishing linguistic conventions in groups as small as ten agents that were observed in the previous chapter. When faced with this issue, Dessalles (2007) suggests that language emerged together with the evolution of *Machiavellian Intelligence* (see Byrne & Whiten, 1988) for the purpose of detecting free-riders. While the above is an interesting theory, the argument that will be adopted here is that larger groups required better communication in order to make hunting and scavenging activities more efficient, which is extremely important in a big group, as noted by Bickerton and Szathmáry (2011). Notably, living in a group does not imply that all activities are performed collectively, in fact, most of the time, members of the group will be divided into small subgroups, particularly when going on scavenging runs. The question then is how communicatively integrated these small unions need to be with the whole group in order for them to be able to inform others of their findings and, subsequently, for a group language to be established in the population.

5.2.1 Experiment Design

Based on the above discussion, the first of the more advanced experiments explores the effects of different interaction constraints imposed on agent groups on the developed lexicons in the corresponding populations. The first setting that is considered here, and which has been explored previously by Bachwerk and Vogel (2010) in a slightly different overall configuration, has ten agents paired up in five fixed partnerships, with a new parameter of the model – the *intra-group communicate rate* p_{intra} – defining the proportion of the time that the agents would interact within their own group, with the remaining interactions bouts being performed with agents from other pairings. This configuration was motivated by the findings from the previous chapter, which suggested that agents have much higher chances of evolving a reliable and unambiguous lexicon in pairs, rather than in groups consisting of ten or even just three agents. However, if one assumes that language actually did evolve in groups of ten to twenty agents, the question is then how would the early hominids go about most efficiently in agreeing on lexical conventions within the whole group. In order to model the different levels of subgroup closeness, three levels of p_{intra} have been experimented with: 0.75, 0.5 and 0.25.

The second branch of the experiment is based on a communication pragmatics experiment conducted by Garrod and Doherty (1994). In this, it was suggested that while isolated pairs of interlocutors will achieve certain levels of understanding quicker, a group consisting of constantly interchanging partners will slowly but surely evolve a better communication system thanks to a large number of potential linguistic variants available for selection. This setup has been realized in the LEW by Bachwerk and Vogel (2011) with the help of a new parameter which allows to have speaker-hearer pairs rotated after a certain number of interactions, with every agent interacting about the same amount with every other agent throughout the simulation as a whole.[2] As with the preferred partner setting above, the intuition behind this configuration was that giving the agents a prolonged period of time to interact with one particular partner should give them a good opportunity to agree on the lexical conventions with this partner, before negotiating their use with other members of the group.

Finally, a set of simulations has been performed with homogeneous populations of two and ten agents each, enabling a comparison of the results from the previous chapter to the ones obtained here. Notably, a population of two agents is largely equivalent to having ten agents split into five pairs of two and $p_{intra} = 1$, i.e. allowing pair-local interactions only. On the other hand, having equally distributed interaction chances in a group of ten agents is the same as pairing them up with $p_{intra} = \frac{1}{9}$, effectively setting the probability of interacting with the paired agent to be the same as that of selecting any of the other eight agents as a partner.

The experimental question that is explored here is if agents in the above configurations would be able to establish some parts of the lexicon with a selected interaction partner first and then spread

[2] As can be clearly seen, this setting is equivalent to the pair-free group in terms of the amount of interaction between any particular pair of agents. The big difference is, of course, in the sequential nature of these interactions in the alternating pair configuration.

such lexical items throughout the community. The simulations presented in this section will examine if such an approach could in fact be beneficial and if so, then how do different pairings affect the evolution of a communal lexicon in the corresponding populations. The main conflict that was expected in this set of simulations was that while agents communicating within a fixed pair should achieve higher levels of understanding, these agents are likely to evolve their own sub-dialects that would be quite distinct from those of other pairings, thus making them unable to actually properly cooperate with most other members of the population.

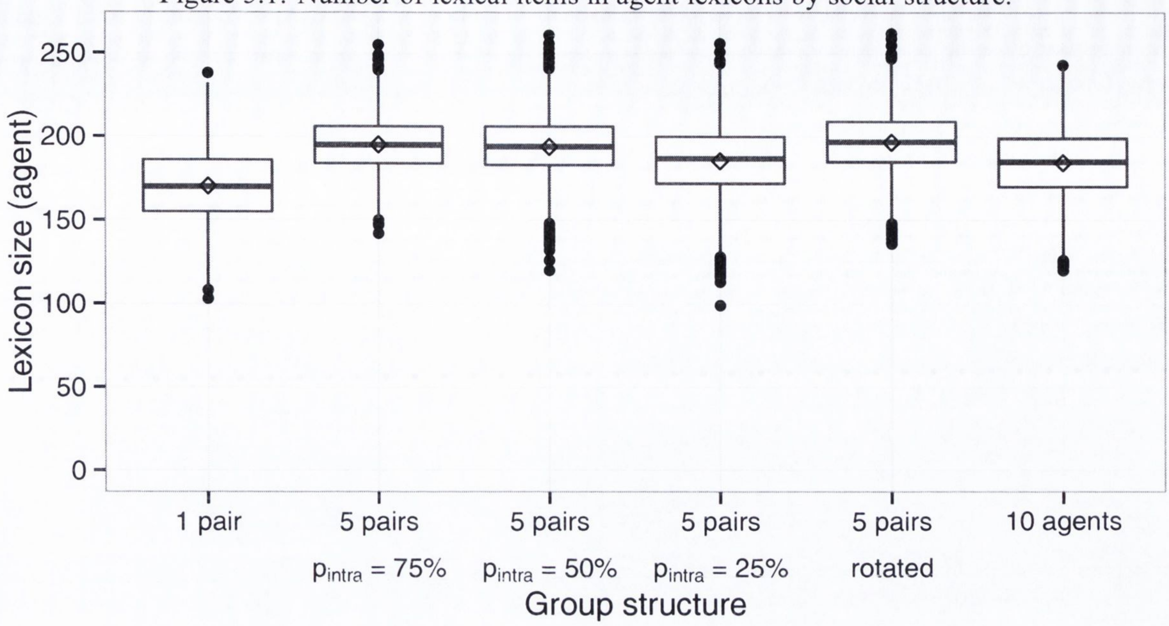
5.2.2 Results

In order to determine if having a preferred partner at the early stages of lexicon acquisition is in fact beneficial for the establishment of a joint population lexicon, a thorough analysis of the main lexical and communicative features of the emergent agent and population lexicons has been performed, based on the metrics presented in section 3.4 and following largely the same script that was followed throughout chapter 4. The experiments presented in this section were also based on 500 randomly re-seeded simulation runs for each of the conditions, with an average of 1000 interactions (either as speaker or as hearer) being performed by each agent. As has been mentioned above, keeping the number of interactions per agent per simulation fixed restricts the agents in terms of the time that they can spend on negotiating the use of different words. However, this has been done quite intentionally as it could not be expected of early humans that they would spend an awfully long time trying to come up with a new communication tool if they do not obtain any rewards from it early on. In addition to that, keeping the length of the simulations constant makes it possible to easily compare the results obtained with the more complex social configurations to the fundamental experiments presented in the previous chapter. In the remainder of this section, the outcomes of the conducted simulations will be analysed in detail, followed up with a conclusion of the results in section 5.2.3.

Lexicon

In a world with 50 instances of 10 events, distributed randomly (though with equal probability) between the three types presented in section 3.2.2, and 25 entities, the number of distinct meanings is at its highest if all ten events are selected to be of type R^3 and all of the entities are utilized in the 50 instances, with no entity sequence coming up more than once. If this is the case, then agents can potentially observe $10+25=35$ single-component meanings, $50+50+50=150$ two-component meanings, $50+50=100$ three-component meanings and 50 four-component meanings, resulting in a total of 335 different meanings for the agents to talk about. However, if the event instances are evenly distributed between the three types, then the agents would experience $35+\frac{(3+2+1)*50}{3}+\frac{(2+1)*50}{3}+\frac{50}{3}=201.67$, i.e. just over 200 distinct meanings on average, again assuming that no entity sequence comes up twice across the 50 event instances.

Figure 5.1: Number of lexical items in agent lexicons by social structure.



Having performed 1000 interactions, agents in all of the presented social configurations end up knowing between 150 and 200 different lexical mappings on average, figure 5.1 shows. As could be expected, agents from the single isolated pair tend to learn a slightly smaller number of lexical items as a result of both these agents taking part in all of the interactions of their population and thus eliminating the possibility of multiple dialects emerging. However, agents from the small groups only learn around 10% more lexical items, which would suggest that if there are any dialects emerging, the mappings from these are not being spread around throughout the whole population. Finally, it is interesting to note that no significant difference in agent lexicon size can be observed between the five different small group configurations, with the alternating pair configuration resulting in the biggest lexicons on average. Before trying to explain this observation, it is worth taking a look at the combined lexicons of the corresponding populations.

Figure 5.2: Number of lexical items in the population lexicon by social structure.

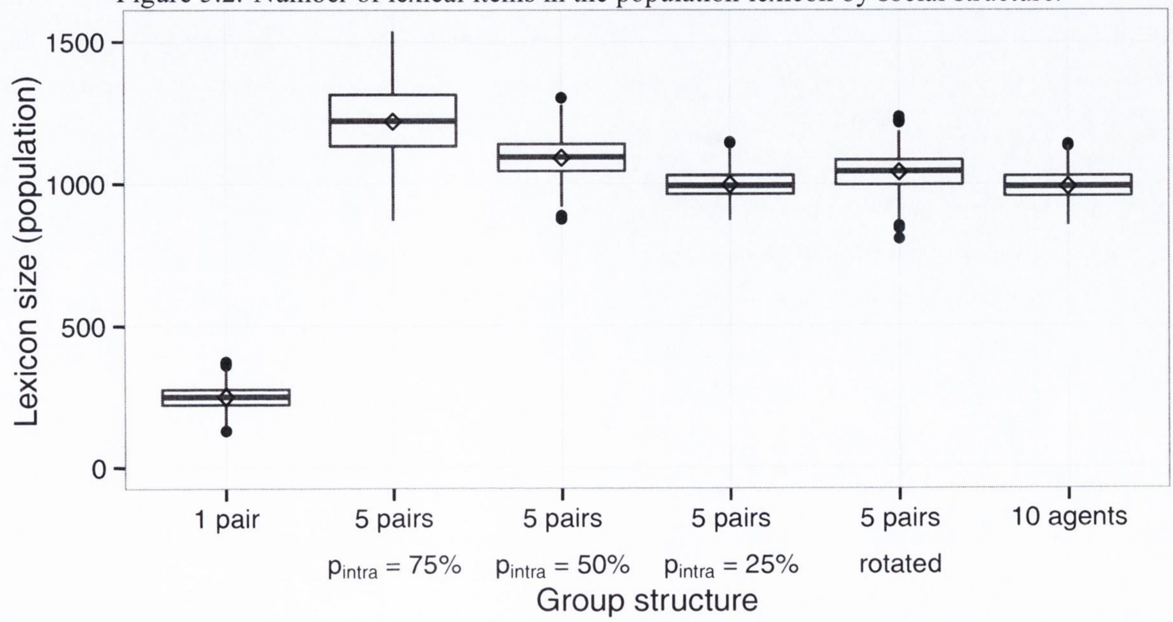
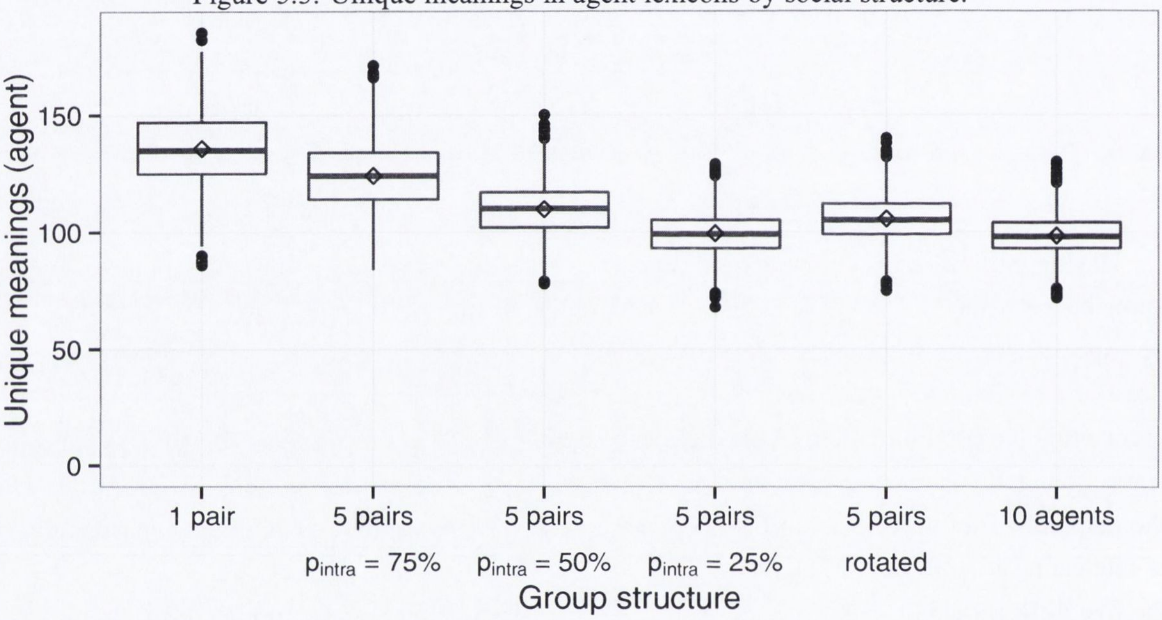


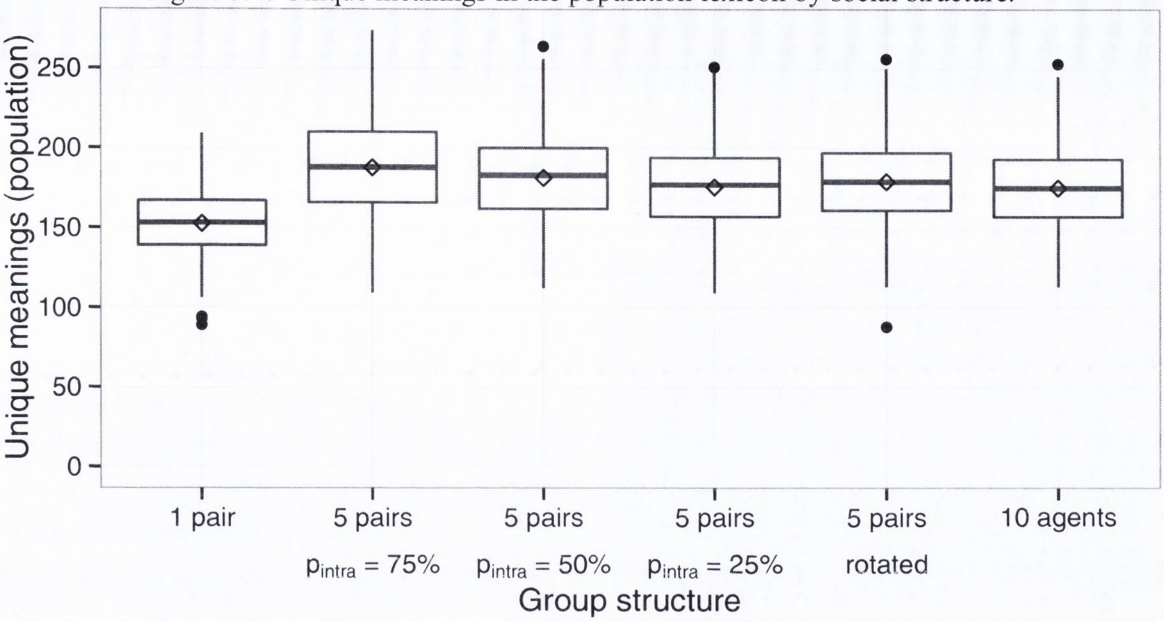
Figure 5.2, depicting the distribution of population lexicon sizes between the six simulated conditions certainly fills in some of the gaps that were left open in the evaluation of individual agent lexicons above. In particular, it is clear that pair-specific dialects are certainly emerging in the presented experiments, especially for the higher two levels of intra-group communication rate. Notably, these dialects are clearly being combined much better in the alternating pair condition, with the observed population lexicon size being the lowest of the five group conditions, and the agent lexicon size – the highest of the five. Going back to the agent lexicons, the relatively similar levels over the five group conditions can be confidently attested to the interaction between the amount of cross-group interactions and the number of dialect-specific lexical items, with the two levelling each other out in agent lexicons.

Figure 5.3: Unique meanings in agent lexicons by social structure.



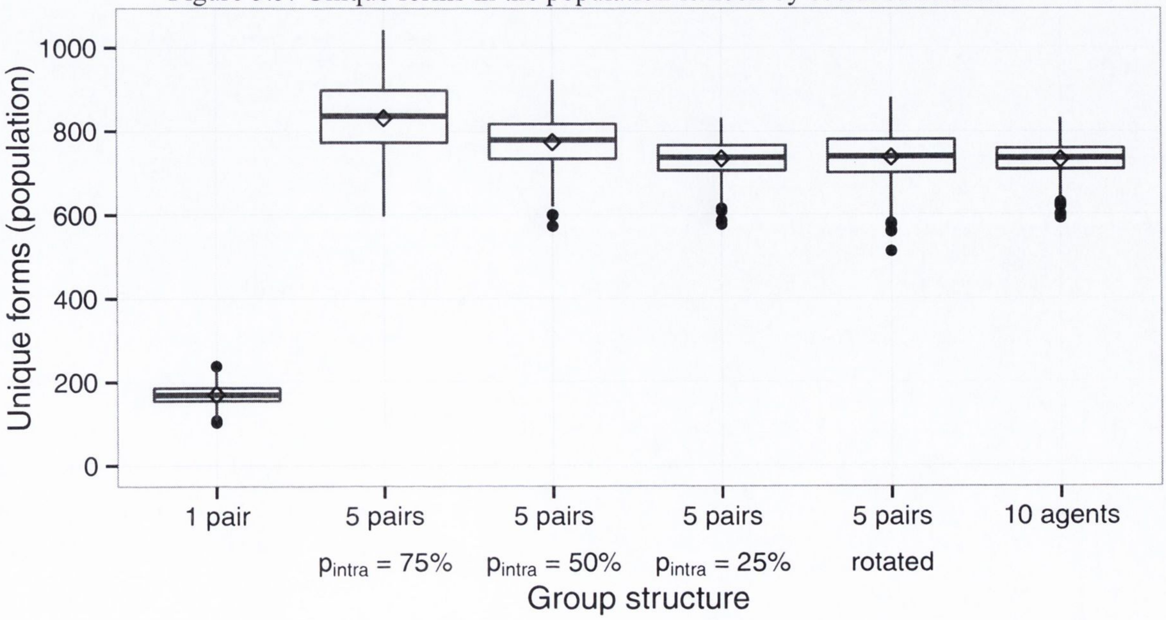
Looking at the number of unique meanings in agent lexicons, represented by figure 5.3, it is interesting to note that the agents from isolated pairs, who ended up having the smallest individual lexicons of the five simulated configurations, actually manage to learn to express a significantly larger amount of different meanings. While this might seem illogical at first, the observation starts to make sense if one considers that these agents can focus on developing their local lexicon throughout the whole simulations and not just 75% of the time or less. In effect, agents who also have to interact with others outside their group are likely to experience more failures due to distinct dialects or simply as an effect of a larger group size, resulting in meaning-form associations actually being eliminated from their lexicons in cross-group interactions, rather than new ones being learned. Finally, it appears that agents from alternating pairs manage to not only learn a higher number of lexical items, but in doing so, also learn to express a significantly larger amount of distinct meanings than their pair-free counterparts, suggesting that the configuration allows the agents to organize their lexicons in a slightly better fashion.

Figure 5.4: Unique meanings in the population lexicon by social structure.



On the population level, however, figure 5.4 suggests that population lexicons of groups with more agents, and in particular those with a higher proportion of pair-local interactions, actually end up having a higher number of distinct meanings represented in them. While this is in fact the case, it would be misleading to conclude that agents in these groups are actually able to communicate a larger number of meanings. In fact, as has been observed from figure 5.3 above, quite the opposite is the case. What is causing the effects observed on the population level is that pair-local dialects will likely happen to include different meanings, simply based on chance. Consequently, if all of the meanings from the different dialects are combined, the overall population lexicon will appear to include a significantly higher number of meanings than the one observed in isolated pairs, even if, individually, every dialect's lexicon had a much lower number of distinct meanings in it.

Figure 5.5: Unique forms in the population lexicon by social structure.

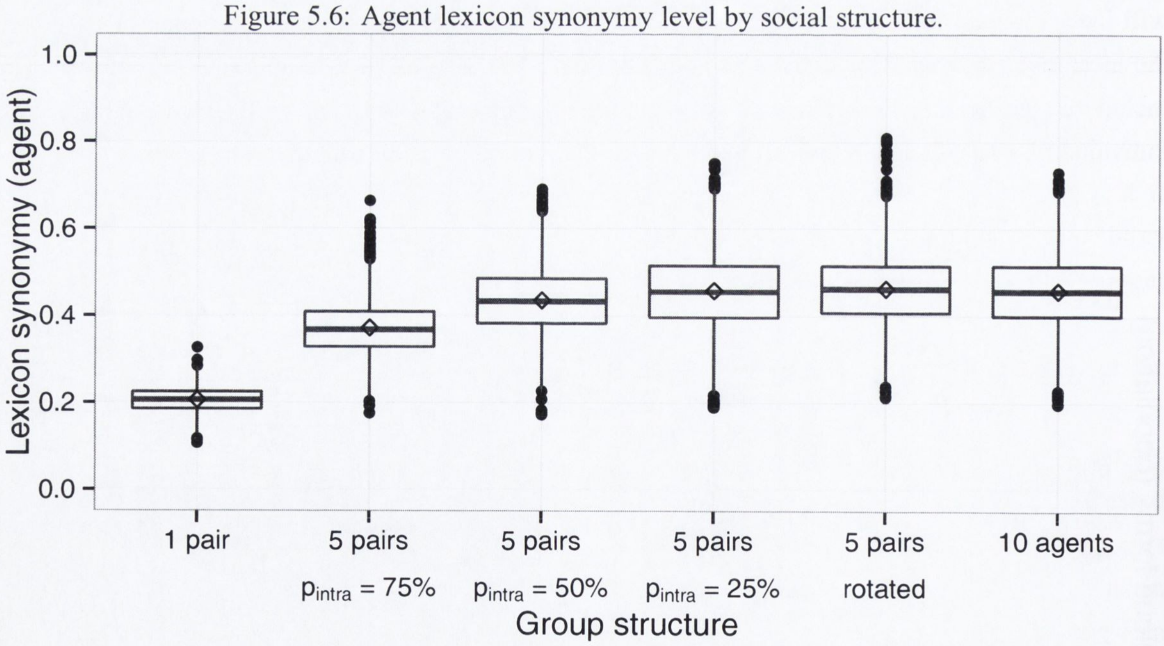


The number of distinct forms learned throughout the population, depicted by figure 5.5, is a clear reflection of the corresponding populations' lexicon size, with a similar trend of groups with more pair-centric interactions ending up with more forms in the overall population lexicon as opposed to those with more homogeneous interactions. As was suggested before, the reason for this trend is that the five agent pairings tend to develop their own dialects, with different forms proving to be successful in different pairs, resulting in a higher overall number of distinct forms being utilised in the group language.

Synonymy & Homonymy

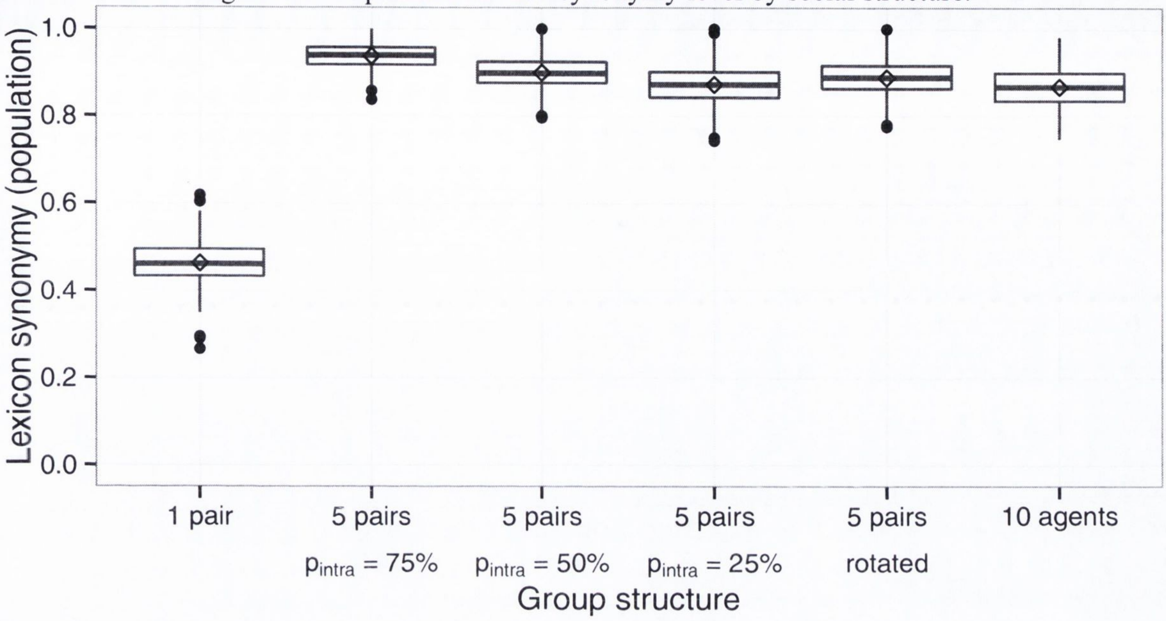
Having evaluated the resulting lexicons from the conducted experiments from a quantitative point of view, the next step is to consider the amount of ambiguity and redundancy in these lexicons. The first measure that gives an indication of the quality of agent lexicons on an individual level is the amount of synonymy, which is summarized by figure 5.6. What this figure clearly shows is that the level of agent synonymy in small groups is nearly twice as high as in isolated pairs, which implies that double the proportion of meanings represented in the lexicons of agents interacting in a group have multiple forms assigned with them. In principal, this outcome is not unexpected, given that, compared to an isolated pair, a population of ten agents has basically five times as many sources for innovation and opportunities to interpret words differently across interactions. However, it is perhaps striking that the amount of agent-level lexical synonymy is as high as it is in a population of highly pair-centric agents. In effect, this shows that communicating with the rest of the group even as rarely as once in four interactions is sufficient for it to take a very significant effect on the structure of an agent's lexicon.

Figure 5.6: Agent lexicon synonymy level by social structure.



While synonymy levels on the individual level were still reasonably low, as well as comparable to the levels observed in human languages, it is perhaps not surprising that lexical synonymy reaches almost 100% in the combined lexicons of the small group populations, as shown by figure 5.7. As has been outlined before, the agents in these groups have a high chance of assigning quite different forms to the same meanings across the local dialects that emerge between pairs of agents. Accordingly, the highest levels of population-level lexical synonymy is observed in simulations with the ten agents divided into five pairs and the highest level of pair-internal communication rate $p_{intra} = 0.75$.[3]

Figure 5.7: Population lexicon synonymy level by social structure.

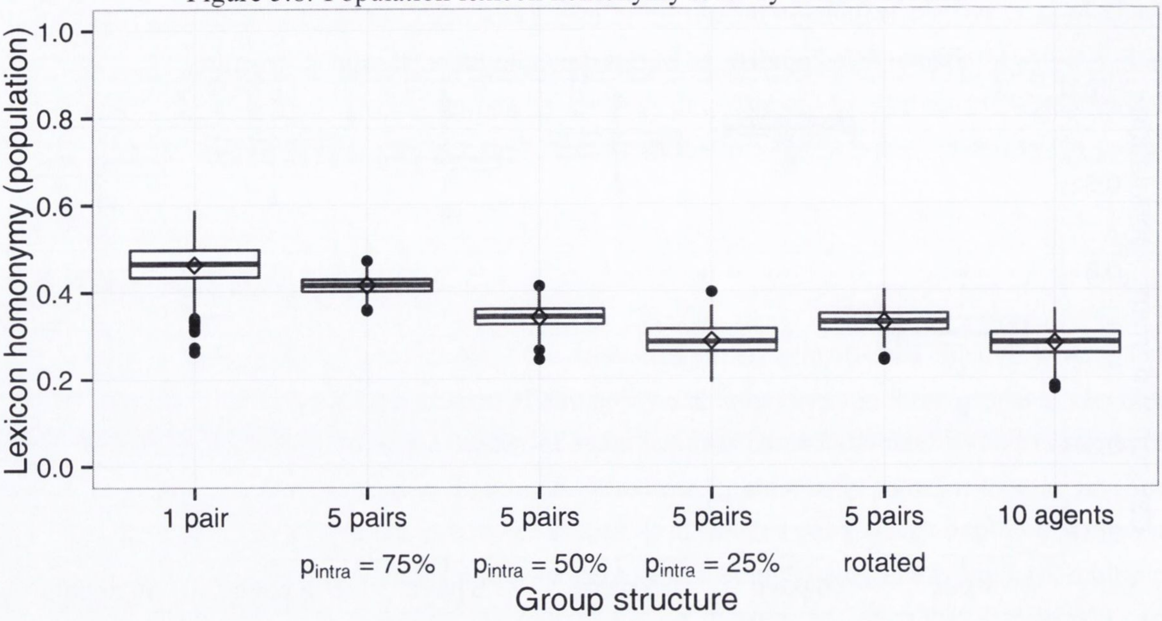


The distribution of population lexicon homonymy levels, depicted by figure 5.8, suggests that homonymy levels are significantly lower in groups of agents, with homonymy at its lowest (and nearest to the levels observed in human languages) in a homogeneous group of ten. In order to understand why this is the case, one needs to consider the different dynamics of pair-internal and pair-external interactions. In both these types of interactions, forms used by a speaker agent will be often misinterpreted by the corresponding hearer. Occasionally, the hearer will correctly interpret a sufficient number of other forms from the utterance in order to obtain a payoff, in which case a new population homonym is created. In subsequent interactions, these two agents are likely to use the same forms for the same events, resulting in repeated borderline success and consolidated homonymy.

However, when an agent will engage in pair-external interactions, the forms that were previously agreed on with his main partner will not be known to others, resulting in a high chance of these being misinterpreted and punished, along with the form that was wrongly interpreted in the first place, eventually eliminating the form from the agent's lexicon. When this happens, the form that was acting as a homonym on the population level will only remain in the lexicon of this agent's partner and thus no longer be a homonym, as it would only have one meaning associated with it in one agent's lexicon. In general, the more agents change their partners the more chances there are that forms wrongly interpreted by some two agents will be removed from the lexicon of at least one of these, leaving numerous low-weight forms across agent lexicons, but driving down population homonymy, as confirmed by the trend observed in the above figure. The only exception to the above rule appears to be the alternating pair setting, where mappings have time to accumulate higher weight values within a pairing and are thus not eliminated as quickly at a later stage.

[3] Arguably, a simulation conducted with five fully isolates pairs of agents would result in even higher levels of lexicon synonymy. However, it was decided against performing such a simulation as performing an analysis on its results would just as meaningful as studying the combined lexicon of French, English, German and two other languages.

Figure 5.8: Population lexicon homonymy level by social structure.



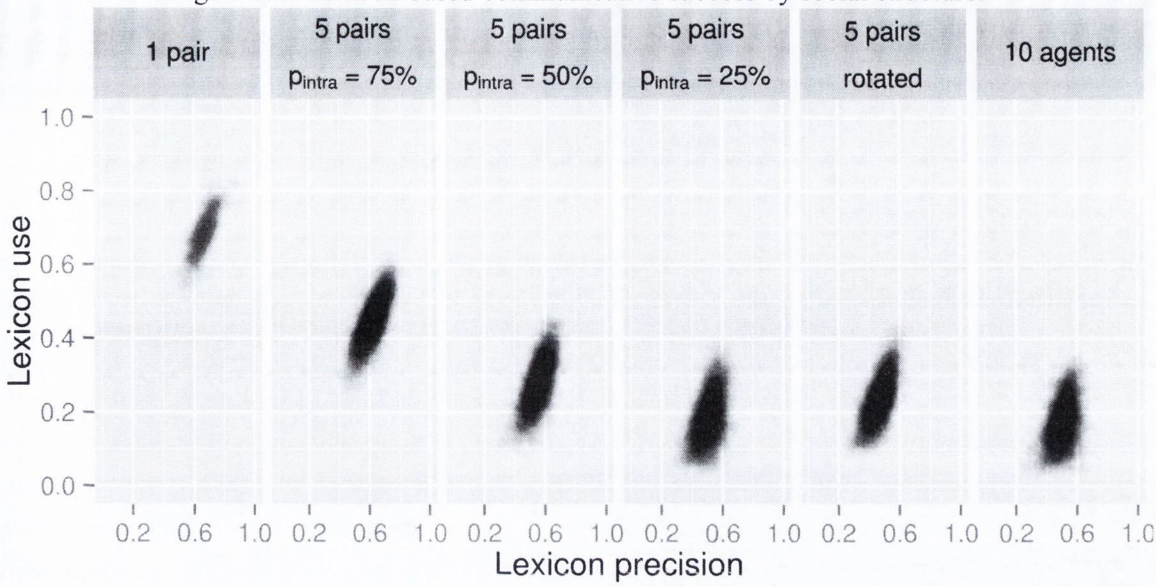
It has already been observed in previous chapter 4 that simulations with a higher number of agents result in less clearly defined lexicons, with the different agents all contributing with their own utterances and interpretations, both successful and less so, to the lexicon learning task, resulting in numerous variations of lexical mappings as well as mismatched meaning-form associations. The results of the current experiment have shown that placing agents in pairs and letting them communicate a dedicated amount of time within the pairing does not reduce the amount of ambiguity in the lexicon of the population, but in fact adds to it through the emergence of pair-specific dialects.

Communicative Success

Finally, the real test for the different social configurations lies in the communicative performance of agents from the corresponding populations. Figure 5.9 confirms that an isolated pair of agents still significantly outperforms any larger group, regardless of its configuration.[4] This result could be expected, given that no matter how the ten agents are connected, they will always struggle to reach the same level of population-wide lexical agreement as two agents after the same number of interactions per agent.

[4]The difference in the shade of grey between the isolated pair and small group settings is solely an effect of a smaller number of data points and can be completely ignored.

Figure 5.9: Lexicon-based communicative success by social structure.



Furthermore, one can observe a strong correlation between the intra-group communication rate and the level of lexicon use across all groups of agents except for the population with alternating pairs. In contrast to other groups, and in particular – to the pair-free setting with the same overall distribution of interactions between agents, being isolated with different interlocutors for fixed periods of time clearly helps the agents agree on the mappings with every successive partner significantly better.

5.2.3 Conclusions

In this section, three experiments have been presented, in which a group of ten agents has been divided into five pairs and assigned one of three rates of intra-pair communication, as well as one experiment with the pairs alternating between ten equal phases of the simulation. Alongside these, two additional simulations have been conducted with a basic population configuration – one with just a pair of agents and one with ten agents (with equal chances of any agent picking any other as an interaction partner). The results of these experiments show that there is clear evidence of dialects emerging within the individual pairs in the small group conditions, with different pairs of agents learning to express different sets of meanings with the help of different forms, resulting in particular in very high levels of agent and population lexicon synonymy.

In terms of communicative success, having a preferred interaction partner does not appear to be benefiting the agents in the long run. In particular, the more self-centred pairs are certainly able to achieve higher levels of communicative success when interactions are performed within the pair. At the same time, these agents obviously find it quite infeasible to also agree on the lexical items to use in different situations with the rest of the population, resulting in a very low number of successful cross-pair interactions, subsequently pulling the average levels of communicative success down below those of fully isolated pairs. However, if the agent pairs are fully isolated for a certain period of time, before being rotated, the positive effects of the two-agent condition are slowly carried over onto the population as a whole, as observed in the alternating pair setting.

5.3 Friendships

In most computational models of language evolution, agent relationships are normally defined at the beginning and kept fixed throughout the individual simulation runs. In such configurations, agents can be basically represented as vertices that are connected by predefined, and usually equally weighted edges. In some cases, these agents would be divided into two or more 'age' groups, whereby the vertices in some groups, e.g. the 'child' agents would not be connected among themselves (see Gong, 2010, for a review of such models). However, the common trait of nearly all such models is that the social ties between agents remain fixed throughout any simulation run, i.e. agents can neither become 'friendlier' with some agents, nor break off all contact with others. This lack of dynamic social networks has been identified by Gong and Wang (2005) as one of the major issues with the majority of computational models of language evolution to date. In order to address this, Gong and Wang (2005) and Bachwerk and Vogel (2010) have introduced varying inter- and group communication rates into their models of syntax and lexicon formation, respectively. In the corresponding experiments, dynamic interaction patterns have proven to affect most aspects of the emerging communication systems.

However, even the introduction of more sophisticated social structures into models of language evolution is not sufficient as the ties between (groups of) agents in such models remain fixed from the start of a simulation, whereas it has been observed by e.g. Panzarasa, Opsahl, and Carley (2009) that humans constantly re-evaluate their relationships with their interlocutors depending on the obtained feedback and the partner's reciprocity. In effect, apart from contributing to the question of the emergence of language *per se*, a computational model with dynamic ties between agents should also shed light on how agents' decisions during individual interactions influence the overall social structure of the population, i.e. how the very first language-induced friendships were made in human evolution. This section will present an experiment with dynamic social ties in the context of language evolution, with the setup of the experiment described in section 5.3.1, following with an evaluation of the executed simulations in section 5.3.2 and a discussion of results in section 5.3.3.

5.3.1 Experiment Design

As outlined above, the final experiment of this project has been largely inspired by a model of language evolution presented by Gong and Wang (2005), in which agents continuously updated the strengths of their relationships, i.e. the probabilities of future interactions with each other based on the success of preceding communication bouts. Furthermore, if the tie connecting a pair of agents exceeded a certain weight threshold, the agents were considered as 'friends' (the labelling had no additional implications in the model). The results of experiments performed with this model have exhibited friendships to be established between agents, as well as a certain amount of clustering occurring throughout the population. Interestingly, the model introduced no limit on how many 'friends' an agent may have and considered agents who were 'friends' with every other agent simply extremely popular.

The ability of an agent to become friends with every other agent in the model by Gong and Wang (2005) seems slightly questionable. In particular, it does not seem fitting that the definition of 'friendship' is local for every pair of agents and does not take into consideration the strengths of an agent's ties with other agents.[5] In order to address the above issue, it has been proposed by

[5] An agent who has a connection to every other agent with $r = 0.0001$ is as likely to communicate to any of these as

Bachwerk and Vogel (2012) to utilize a relative partner selection approach, similar to the lexical selection strategy presented in section 3.3.2. In particular, the idea is to perform updates of the strengths of agent relationships r for every agent $i, j \in \{1, .., N\}$ in a population of N agents based on the communicative success of their interactions, with the level of minimum success s_{min} kept at 0.35, as follows:[6]

$$r_{ij}(t+1) = \begin{cases} r_{ij}(t) + \alpha * \pi(t) & \text{if agents } i \text{ and } j \text{ were involved} \\ r_{ij}(t) & \text{otherwise} \end{cases} \tag{5.1}$$

When executing the above strategy, the same dynamic payoff definition from equation 3.5 was used, with the only adjustment being that the connection strengths between every pair of agents i, j were set to $r_{ij} = 1$ at the beginning, i.e. every agent started out with nine social ties with relative weights $w_{ij} = \frac{1}{9}$, i.e. $w_{ij} > 0.1$. Furthermore, when updating the connection strengths between agents, the payoff value was multiplied by a scaling parameter α that essentially defines the ratio between the speed of social and lexical learning (cf. Santos, Pacheco, & Lenaerts, 2006). In the presented experiments, the value of the social update rate α was varied between 1, 0.5 and 0.1, i.e. the agents were always at least as quick in updating their lexical mappings as they were in adjusting their social ties.

The dynamic connection strengths between agents also allow one to observe what kinds of social ties evolve in a population which is not bound by a predefined network structure, but where its members are capable of becoming more familiar with some, and more distant towards others. In order to be able to evaluate the results on a per-agent basis, three types of social connections are defined here, that an agent i might end up establishing with a second agent t based on the connection's relative weight $w_{it} = \frac{r_{it}}{\sum_{j=1}^{N} r_{ij}}$, i.e. the probability of i selecting t as a partner in an interaction: an acquaintance ($w_{it} \geq 0.1$), a friend ($w_{it} \geq 0.25$) and a best friend ($w_{it} \geq 0.5$).[7] In particular, it was observed if and how many connections of any given type were established by the agents, given any particular value of α, with the goal of determining how the first language-induced friendships could have been established in human evolution, along with the corresponding effects on the success of the lexicon acquisition task itself.

5.3.2 Results

The evaluation of the experiment results that is presented in this section follows the usual script and compares the three different dynamic group configurations with results from simulations with the standard populations of both two and ten agents. As has been the case for all of the presented experiments, 500 simulation runs have been executed for each of the three different α values, with 5000 interactions in each of the simulation runs. The focus of the analysis is to determine the effects of allowing agents to adjust their interaction ratios with each other within the group as well as to evaluate the emergent social structures. For the purpose of the latter task, an additional section has been added to the analysis that should provide some insights into the state of the relationships that have evolved between agents by the end of the simulation runs.

an equally interconnected agent, but with $r = 1.0$, making the latter agent's 'friendship' ties no different from the former agent's weak ties.

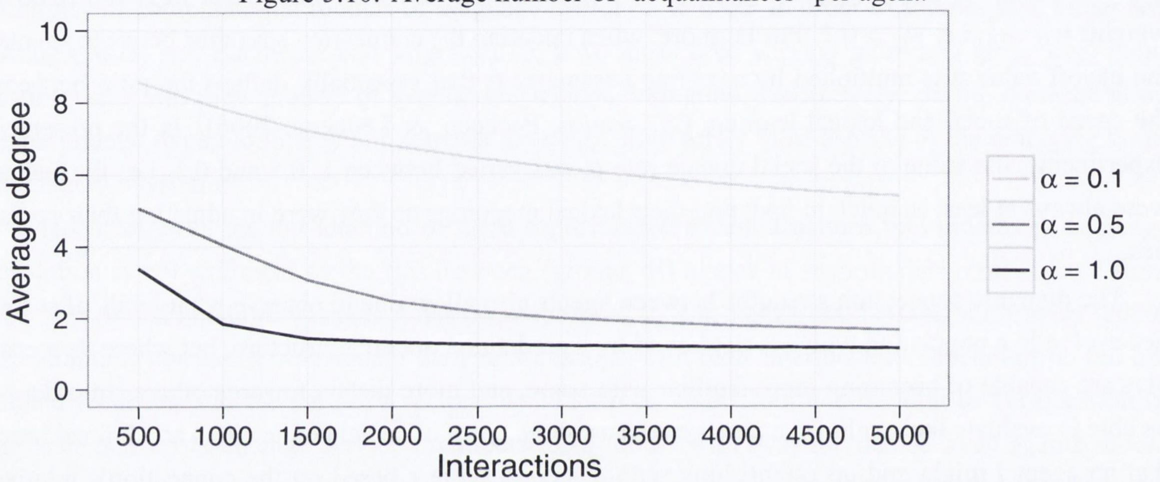
[6] If an agent's connection strength with another agent reaches 0, the former will avoid the latter in all future interactions.

[7] Given this definition, the analytical limit for any particular agent is to have 2 best friends, 4 friends and 10 acquaintances, population size permitting.

Agent Relationships

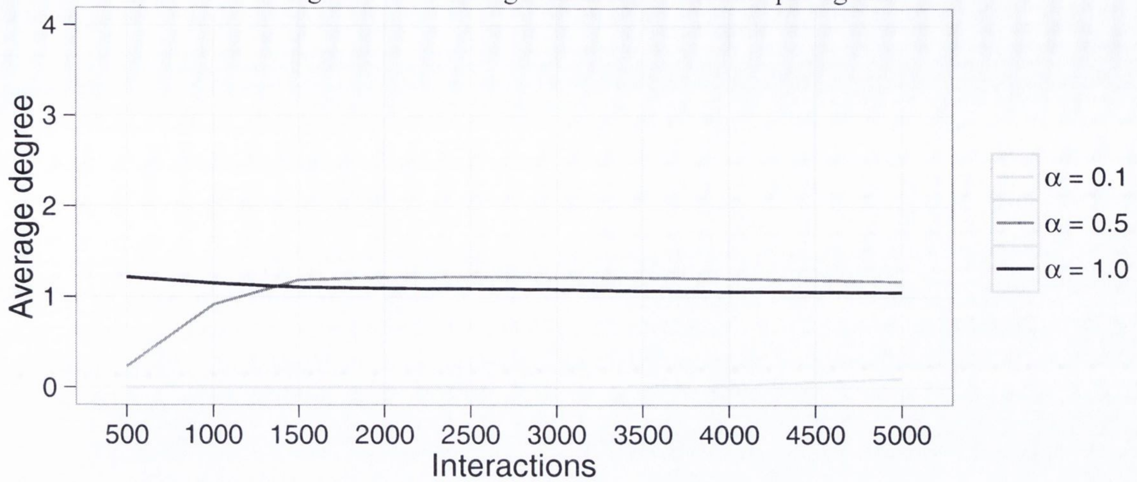
Before going into the usual analysis of the agent lexicons, the setup of this particular experiment warrants a thorough examination of the agent social networks. In particular, the main question here is if the social structure of the simulated groups is significantly evolving at all under the new conditions. Judging by figure 5.10, one can say that the initial relative connection weights of $w = \frac{1}{9}$ do not remain for long as the average number of connections per agent with $w \geq 0.1$ steadily decreases throughout the course of the simulations. In particular, for higher values of α, one can observe a radical drop in the number of 'acquaintances' from the initial nine to five for $\alpha = 0.5$ and almost three for $\alpha = 1$ by the end of the first 500 interactions already, i.e. after just 100 interactions per agent.

Figure 5.10: Average number of 'acquaintances' per agent.



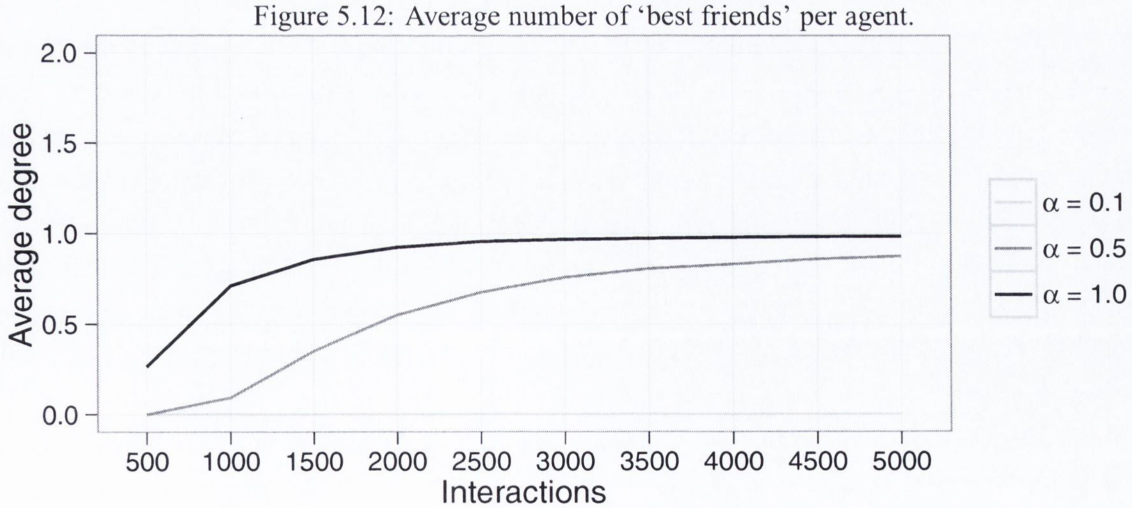
The shrinking number of 'acquaintances' observed for all values of α suggests that the interaction weights are being redistributed more densely between the more popular communication partners of an agent, which leads to the establishment of first 'friendships'. Figure 5.11 confirms that this is the case for agents with $\alpha \geq 0.5$ who on average make at least 1 'friend' by the end of the simulation runs. However, agents with $\alpha = 0.1$ do not achieve a connection weight $w \geq 0.25$ with any of their conspecifics until very late on in the simulation runs (and even then just a handful of the agents do so) due to their 'unwillingness' to give up on other partners too hastily. In effect, the different social update rates model different human characters, ranging from patient pragmatics that are always cautious and thus end up having very practical friendships, if any at all, to the more impulsive types who are ready to become friends with anyone quickly.

Figure 5.11: Average number of 'friends' per agent.



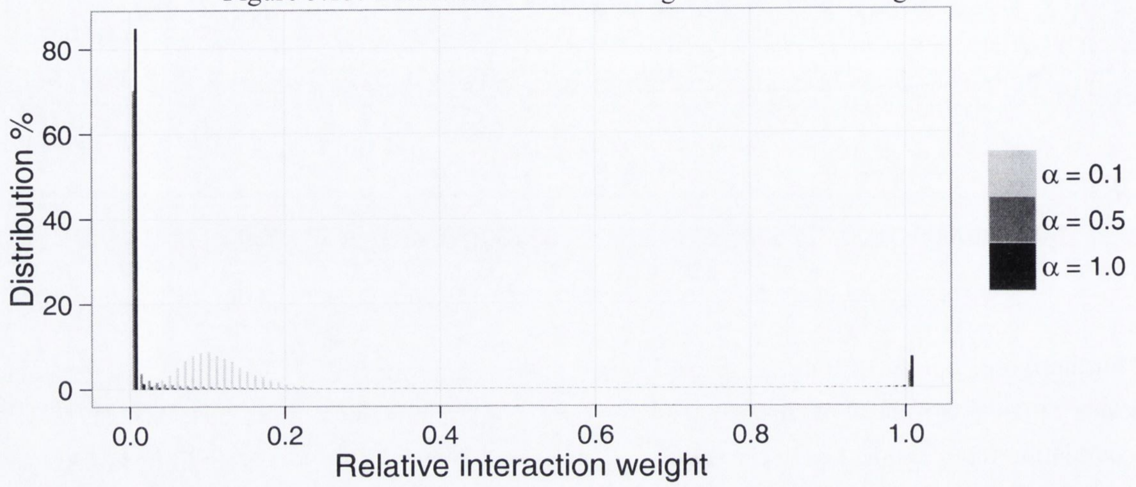
Furthermore, figure 5.12 shows that all agents in simulations with $\alpha = 1$ as well as the overwhelming majority of those in simulations with $\alpha = 0.5$ end up having a 'best friend' by the end of the simulation runs, in effect only ever communicating with one fixed partner from the population of ten. While these agents could be expected to perform significantly better than their highly interconnected conspecifics, based on the previous findings from simulations with individual pairs of agents, one has to remember that friendships as defined in the LEW are not necessarily reciprocal, i.e. one could end up with any social structure from five isolated pairs, two subgroups of five agents, to a cyclic arrangement of friendships where every agent has befriended the next one in line. Naturally, each of these configurations will have a different effect on the various lexical properties. However, in the following, the focus will be predominantly on trying to explain the more general effects of introducing dynamic social ties into the model.

Figure 5.12: Average number of 'best friends' per agent.



Finally, when comparing the social structures that have emerged in the current experiment to empirical studies of human networks, one quickly spots two striking similarities in their construction and organization. First of all, the two-stage process of establishing social relationships that was observed for agents with a higher social update rate (in particular for $\alpha = 1$) has been also reported by Panzarasa et al. (2009) among others, who were able to conclude from their studies that social relationships were established at the very onset of group formation and were then slowly reinforced afterwards. Additionally, it appears that for these agents, the frequency distribution of edge weights, i.e. the connection strengths between agents, depicted in figure 5.13, has a distinct scale-free character (with the exception of an increase in $w = 1$ ties at the tail), which Barabási (2009) considers to be typical of most dynamic networks.

Figure 5.13: Distribution of relative agent interaction weights.



Lexicon

Given that the number of events and entities in the agents' world has been kept fixed from the previous experiment at 50 instances of 10 events with 25 different entities, the number of total meanings that could potentially be observed by the agents has remained constant at 335 as well. While agents in previous experiments have been learning between 150 and 200 distinct mappings in the varying group configurations, figure 5.14 shows that especially for the two higher levels of the interaction ratio update rate, some agents can end up learning as many as 600 different items.

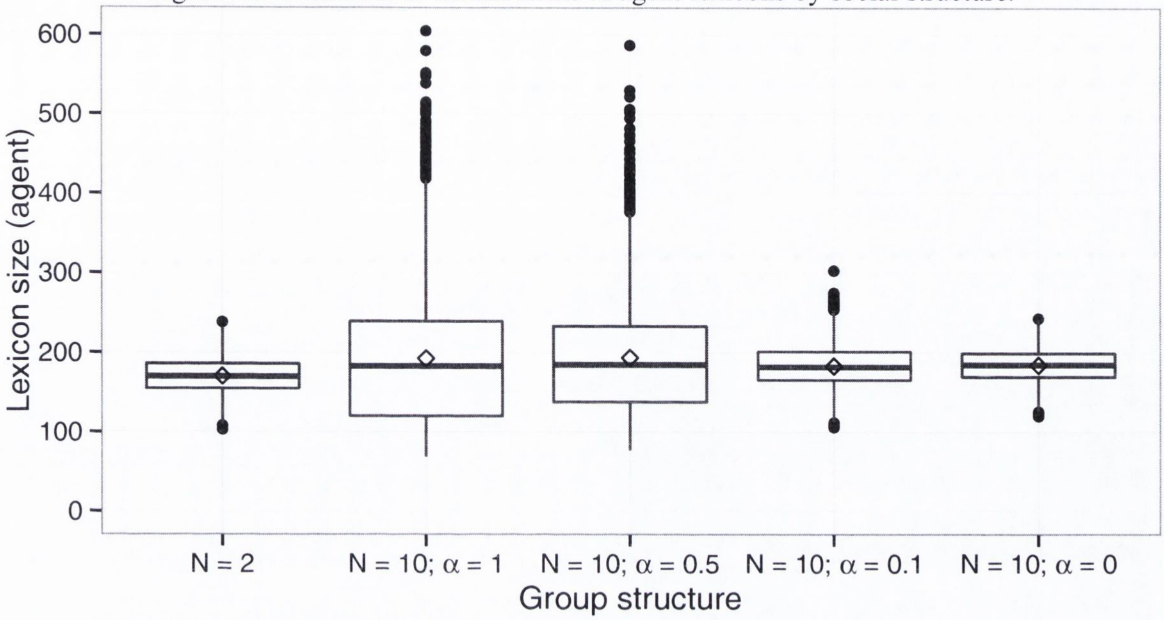
The reason for this lexical explosion lies in the way interaction rates are updated by the agents in the corresponding simulations. In particular, for high values of α, social ties with other members of a group are bound to be significantly weakened, especially during the first, mostly unsuccessful interactions. When this happens, experiencing success with an agent for just a few times can lead to a long-lasting 'friendship' relationship, effectively dominating all of the future communication attempts of an agent. While this per se would not result in a configuration that is significantly different from a group divided into fixed isolated pairs, one has to consider that when such dominant relationships are established, it is unlikely that each of the ten agents from the group will end up having a different 'friend'.

On the contrary, chances are that some agents will be 'friended' by several others, while some might not be anyone's friend, i.e. preferred communication partner.[8] The main consequence of this development is that interaction counts will no longer be equally distributed between agents, e.g. if one agent end ups being the only social connection for five others, he will not only participate in the usual average of 500 speaker-role and 500 hearer-role interactions, but also in a further 2000 hearer-role communication bouts. As a result, such an agent would be exposed to a significantly larger number of lexical variants, allowing (or forcing) him to develop a much bigger lexicon. On the other

[8]Remember that friendships in the current experiment are unidirectional relationships, representing the chances of an agent being picked as a hearer partner by an agent acting as a speaker in an interaction.

hand, agents who have not been selected by anyone as the preferred partner will only take part in 500 speaker-role interactions,[9] thus reducing their chances of establishing a large lexicon.

Figure 5.14: Number of lexical items in agent lexicons by social structure.



On the population level, a similar trend in the evolution of lexicon size is observed for the higher levels of α, as shown by figure 5.15. Admittedly, the overall number of interactions performed in any given simulation run remains exactly the same overall, regardless if every agent has ended up interacting with the same partner based on the performed social updates. However, what is observed here is the dialect effect, which is augmented by the fact that some of the dialects will be completely local and have no chance of merging into a communal language, especially if the social tie adjustments result in several isolated subgroups.

Figure 5.15: Number of lexical items in the population lexicon by social structure.

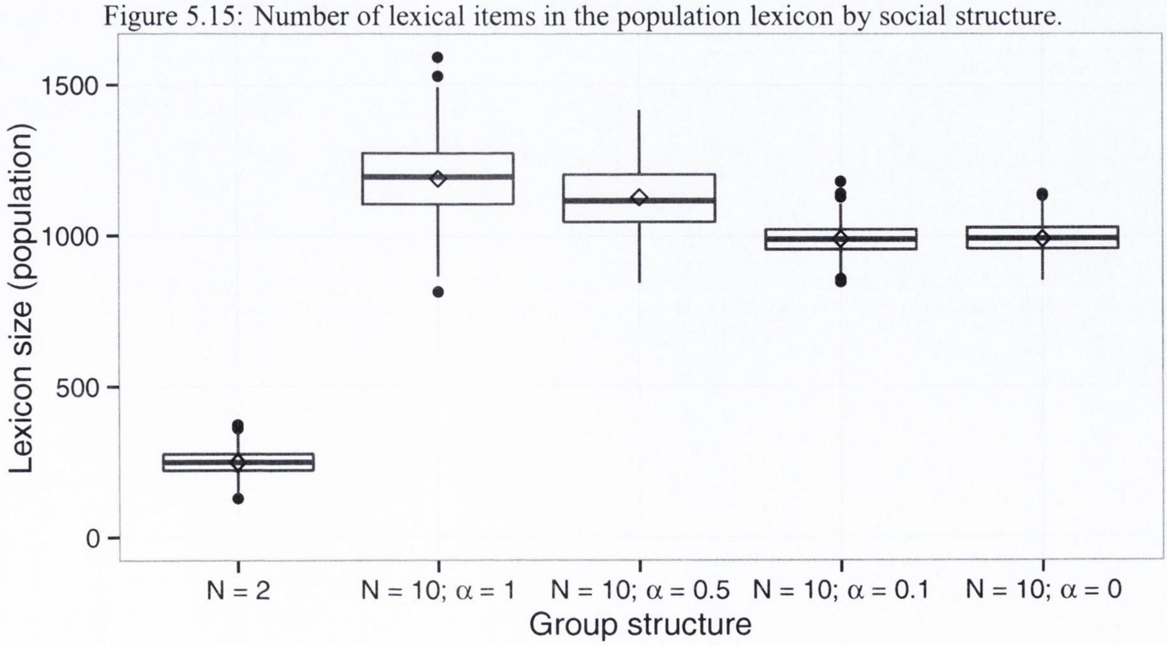
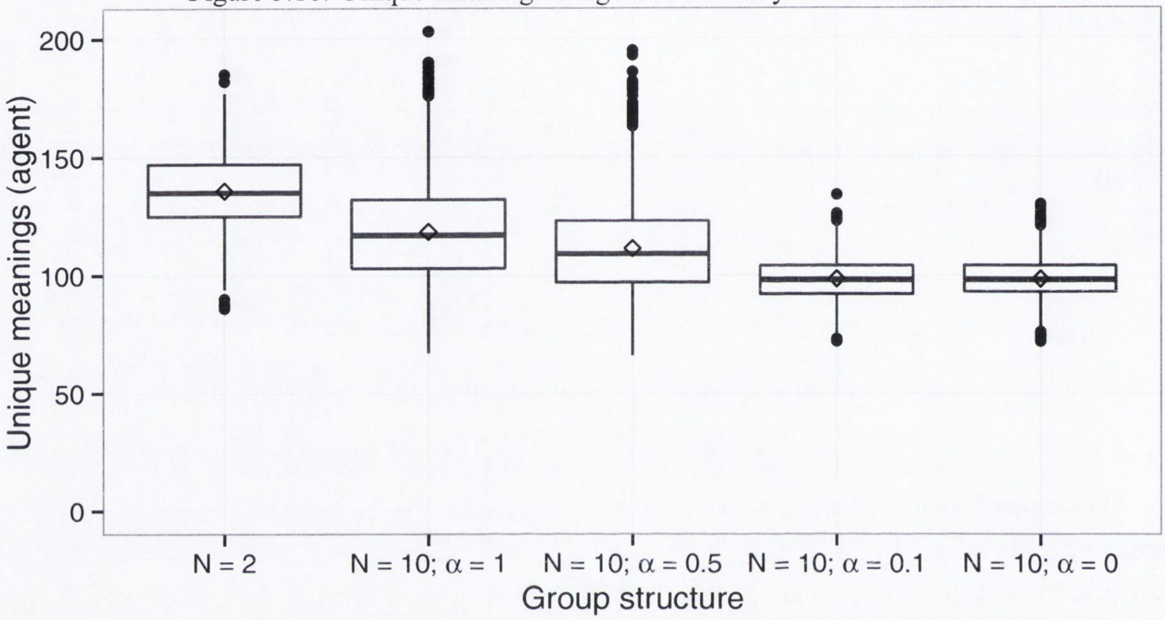


Figure 5.16 shows that agents who end up learning a considerably larger number of words do not

[9]Speaker selection remains fully random regardless of who is friends with whom.

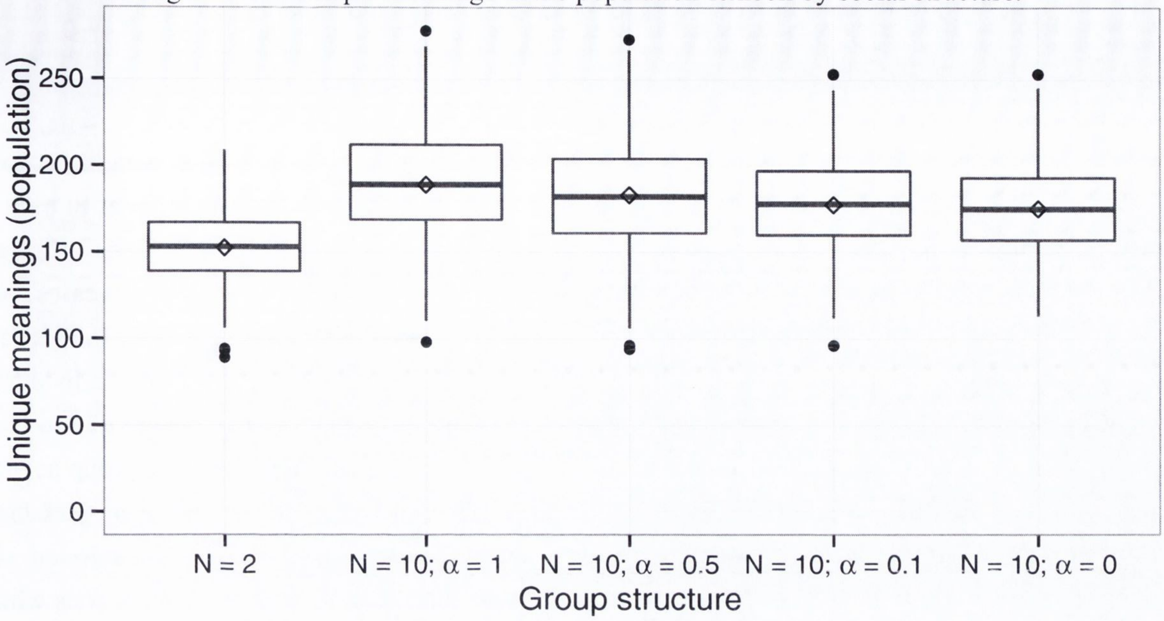
just acquire a variety of different forms for the same meanings, but actually learn to express up to twice as many meanings as agents who adjust their social ties very slowly, or those who do not adjust their interaction rates at all. This effect has been observed before and is mainly due to the fact that an agent exposed to a larger number of events will be able to partition these in a wider variety of ways and thus, assuming that his interaction partners had been referring to the same parts, learn to express a bigger range of different meanings.

Figure 5.16: Unique meanings in agent lexicons by social structure.



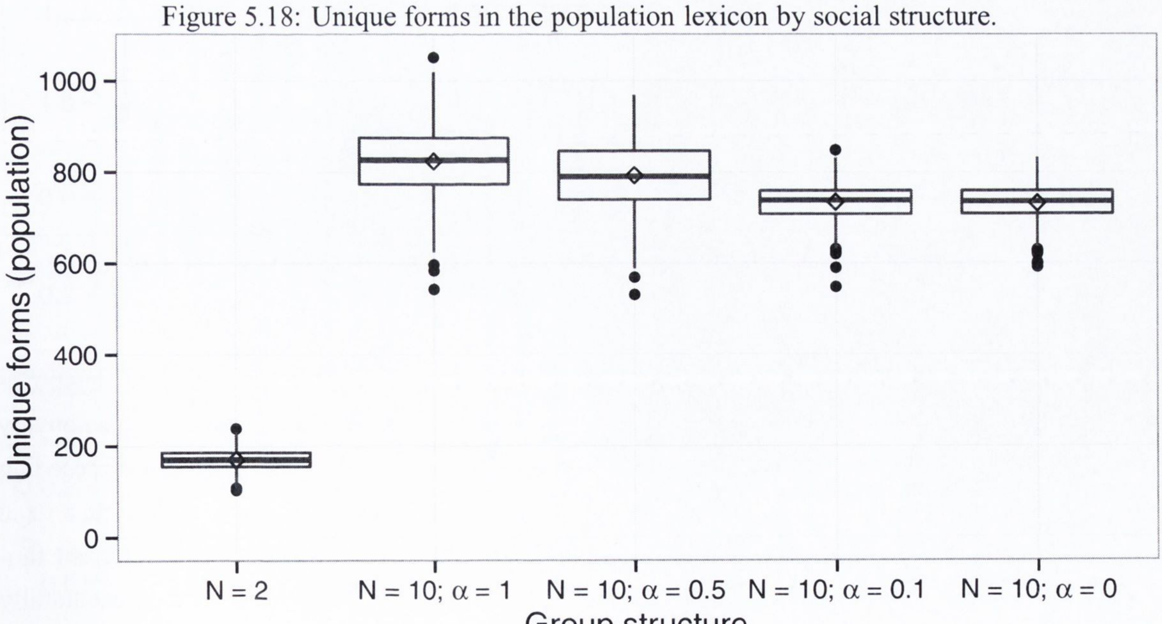
Despite the significantly higher number of distinct meanings represented in some of the agent lexicons in simulations with high levels of social update rate, the trend is only marginally carried over onto the overall lexicon of the population, as suggested by figure 5.17. The social configuration of the population plays a much smaller role here because no matter what it is, the number of simulated events and their perspectives will not be affected by that in a fixed-length simulation run. However, the fact that a slight improvement is actually present can be explained by the lower chances of failure within a reduced social clique, as previously agreed on mappings do not need to be guessed by new agents and are thus more rarely eliminated.

Figure 5.17: Unique meanings in the population lexicon by social structure.



As expected though, the number of distinct forms in the combined population lexicon grows significantly with an increased tendency to have a single interaction partner, figure 5.18 confirms. Just as was the case with the size of the population lexicon, having agents arrange themselves into small, sometimes fully isolated subgroups results in different dialects being developed that utilise different forms. Consequently, when all of the population dialects are combined into one communal language, it is not surprising that the resulting lexicon has much more variety than that of a more homogeneous population.

Figure 5.18: Unique forms in the population lexicon by social structure.

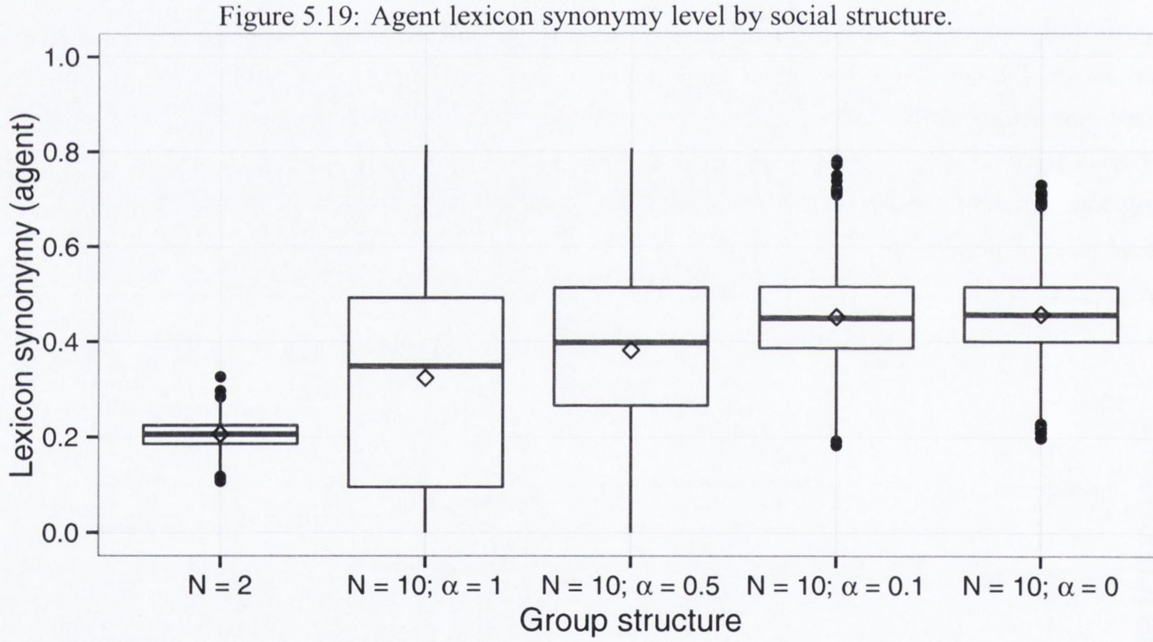


In summary of the lexical analysis, it can be said that being able to adjust one's social ties surely has an effect on the developing language of the population, with a significant increase observed in the number of distinct words, as well as individual forms and meanings represented in both agent and population lexicons. However, it remains questionable if such a development is necessarily of a positive nature, especially if it comes as a consequence of small world cliques being established in the community and local dialects being learned by the agents within these cliques.

Synonymy & Homonymy

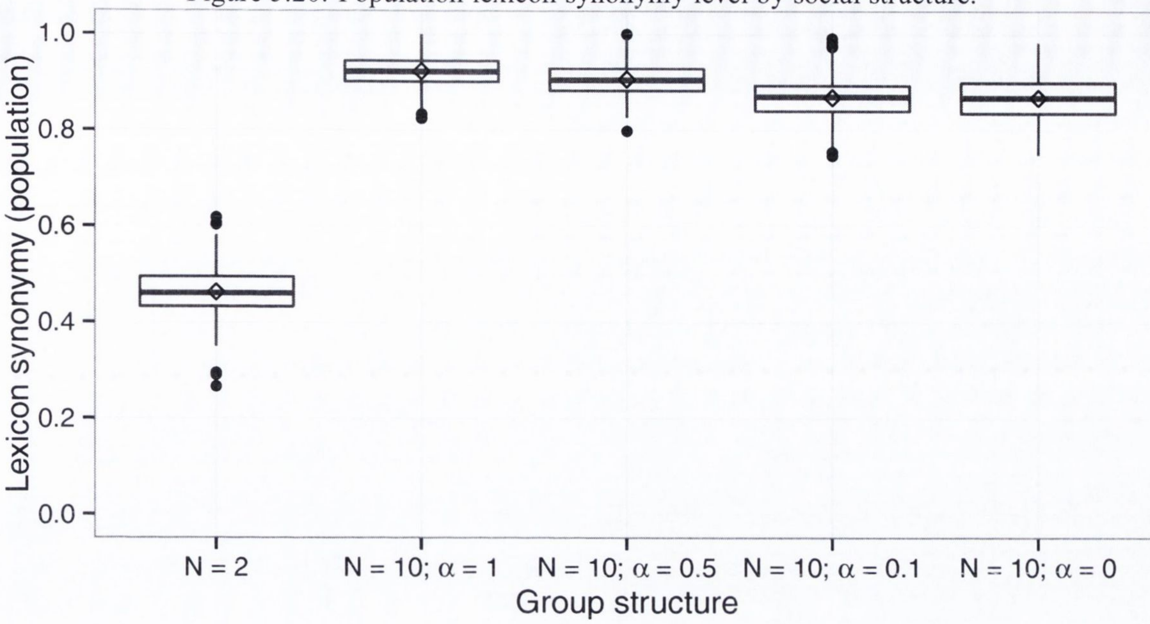
Given the observations from the previous subsection, it is interesting to see if the emerging agent lexicons in the conducted simulations with adjustable social ties will be also any different in terms of the amount of ambiguity and duplicate meaning-form or form-meaning representations. Figure 5.19 indicates that, while the average trend of agent lexicon synonymy monotonically decreases for successive higher levels of the α parameter (in groups of ten agents), falling exactly on the 0.38 level observed in human languages for $\alpha = 1$, the distribution of the synonymy property across individual agent lexicons actually becomes markedly wider with every step. This effect is mainly due to the skewed distribution of the average number of times that different agents from the same group act as hearers throughout the simulations. Given that, in the LEW, synonyms are created when guessing a meaning of a form during interpretation, the consequence of some agents not being selected as anybody's 'friend' and thus not being *talked to* at all, is that these agents will develop lexicons with no synonyms, yet with over 60 different meanings represented in them.

Figure 5.19: Agent lexicon synonymy level by social structure.



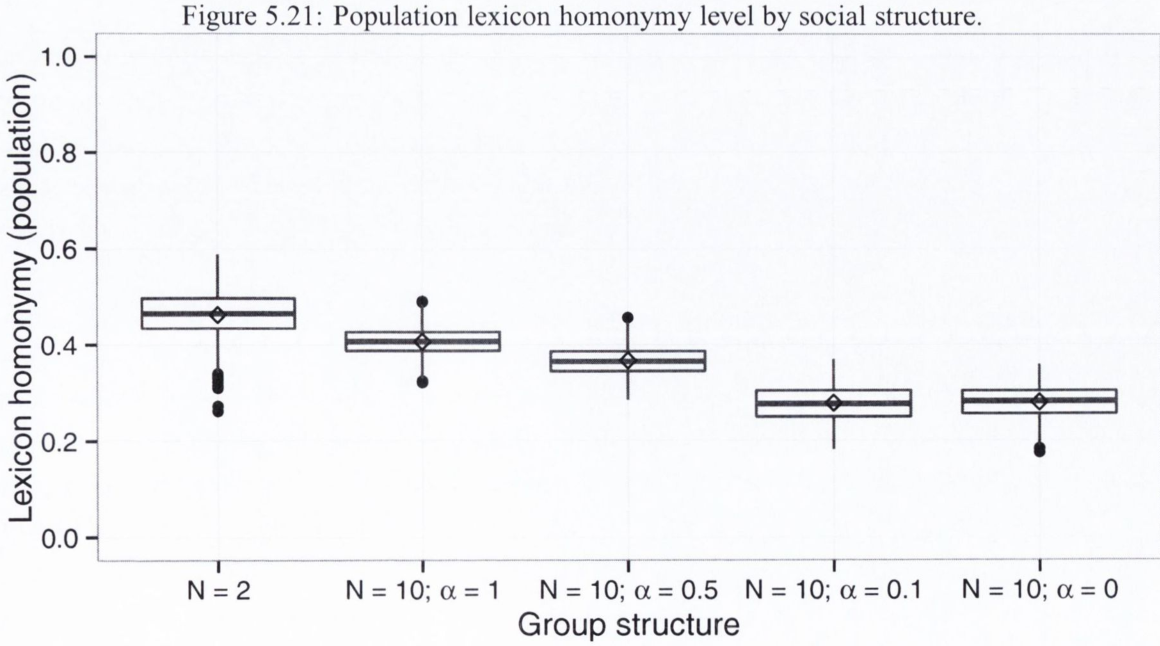
When the lexicons of individual agents are combined, figure 5.20 indicates that the resulting population lexicons in simulations with the two higher levels of α exhibit even higher levels of synonymy than lexicons from the fixed-connection populations. Based on the discussion from the preceding quantitative analysis of agent lexicons, this effect is not surprising, given that in a world with a fixed set of meanings, as well as the probabilities of encountering any of these, agents from different subgroups will generally learn to express similar sets of meanings in the different dialects, thus naturally resulting in high levels of population synonymy.

Figure 5.20: Population lexicon synonymy level by social structure.



The levels of population homonymy depicted by figure 5.21 for the dynamic social network experiment are representative of exactly the same effects that were observed in the preceding group experiment with a varying intra-group communication rate. As was the case in the earlier experiment, agents who end up talking in isolated pairs or small groups end up developing lexicons that have a higher amount of homonymous forms on the population level due to misinterpretations between these agents basically not being rectifiable once enough parts of the corresponding event partition have been agreed on by the agents for a payoff to be obtained in all future interactions, even if other parts are being persistently miscommunicated, thus keeping homonymy levels up.

Figure 5.21: Population lexicon homonymy level by social structure.

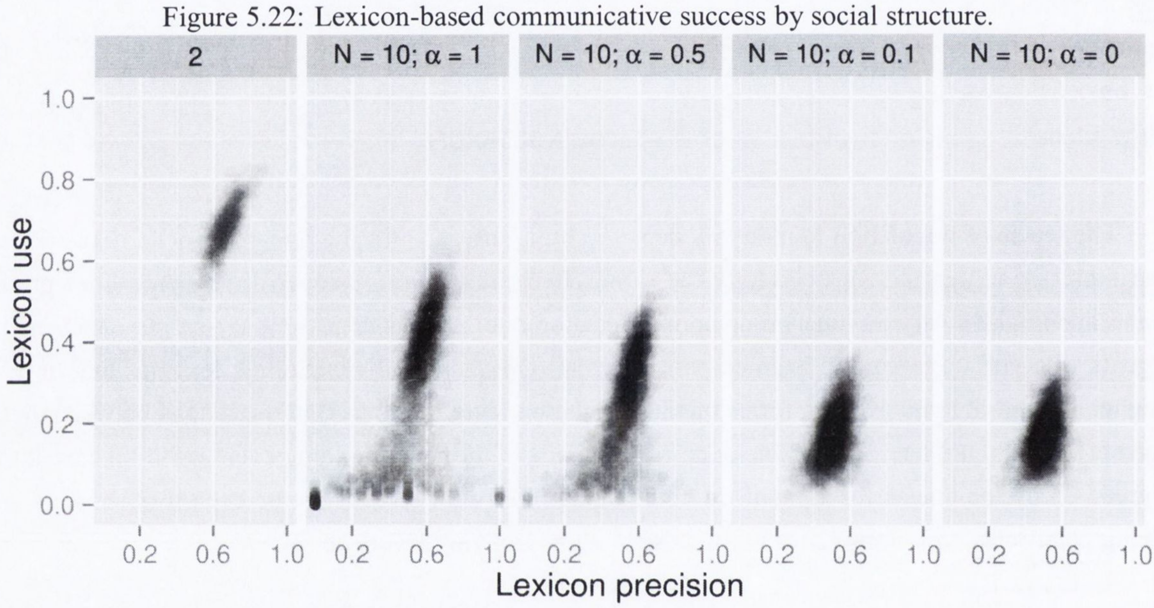


In summary, the possibility of complete isolation of pairs or small subgroups of agents as a consequence of the adjusted social ties certainly contributes to higher levels of ambiguity on the population level through increased levels of lexical synonymy and homonymy. However, the possibly most intriguing observation from the lexical analysis is the wide range of agent lexicon sizes, as well as the synonymy levels within the individual lexicons. In particular, the latter can be as low as zero for agents who are only acting as speakers as a result of not being picked by any other agent as the favoured interaction partner.

Communicative Success

In terms of the lexicon use and lexicon precision levels that are reached by the agents in the dynamic social groups, the overall trend is clearly of a significant improvement in terms of both these measures, again due to the separation of the ten agents into partially or fully isolated subgroups, in which it becomes easier to agree on the lexical conventions.

Figure 5.22: Lexicon-based communicative success by social structure.



However, figure 5.22 also suggests that there is a sizeable portion of agents in the two more dynamic conditions that achieve only very moderate levels of lexicon use and a further group of agents that have their lexicon use fixed at zero. As could be anticipated, these agents are again those who have not been picked by any of their conspecifics as 'friends' and thus never utilise their lexicons during interpretation simply because they (almost) never act as hearers.[10]

5.3.3 Conclusions

In summary, it appears that in order for a non-random social structure to emerge, agents need to make their decisions in a decisive and resolute manner. In effect, the fact that social dynamics of agents in our model so closely resemble those of modern humans for certain parameter levels certainly suggests that, if our predecessors were anything like us socially, it is likely that one dominant characteristic of early hominids was a quickness to bond based on successful communication. However, one issue that arises with such individuals who become 'friends' with a few of their conspecifics and mainly ignore the majority of the population is that it is hard to imagine how they could have agreed on a highly conventionalized communal communication system such as language.

[10] A reminder that lexicon use is a measure of how many of the speaker's words were found by the hearer in his lexicon, i.e. only relevant on the interpretation side of an interaction (see section 3.4).

On the other hand, however, if 'making friends' implies forming exclusive relationships with one particular agent or a very small set of agents and consequently ignoring the rest of the group when it comes to communicating with others, it is becomes questionable how such a community could keep on functioning in the long run. Additionally, as could be seen above, agents who will not be picked as 'friends' by anyone will never be able to test themselves as addressees of communication attempts, even if admittedly this results in them developing a fully unambiguous and actually quite efficient lexicon. In general, it would seem that one should also investigate the possibility of the existence of either universally esteemed and popular agents that are constantly being communicated with regardless of the social connections of others, or a more uniform distribution of agents into different 'social characters', i.e. from ones that make friends almost instantly to others who take their time in developing their relationships with others.

5.4 Summary of Social Experiments

The goal of the current chapter was to evaluate the effects of different social configurations and dynamics on the performance of the simulated agents in the lexicon acquisition task. Two experiments have been conducted with a more extensive setup of the LEW model parameters, in particular including 50 instances of 10 different events and 25 distinct entities. Furthermore, in each of the two experiments, the groups of ten agents have been divided into a variety of structures and assigned a range of interaction patterns that could be realistically imagined to have occurred naturally in the early stages of language evolution. In particular, the main focus of the experiments was on determining the possibility of leveraging the significantly better learning results observed in pairs of agents throughout the establishment of a joint population lexicon.

The first of the two experiments had a group of ten agents arranged in five pairs, with a varying pair-internal communication rate imposed on the participants in the first instance of the experiment and a successive partner rotation between otherwise isolated pairs being executed in the second version of the experiment. The main observation from the results of these experiments was that pair-specific dialects have clearly started to emerge between agents, even if their pair-internal communication rate was set to just 50%. However, in the rotated pair setting, the dialects have been able to fuse better into a single population lexicon, with the agents in these groups performing notably better than those from a fully homogeneous group. Since the average distribution of agent interactions in both these latter group is exactly the same, it can be confidently claimed that the effect of improved lexicon performance is due to the more organized, sequential nature of the interactions in the alternating pair setting.

The following experiment looked to build further on the findings from the rotating pairs setting. For this purpose, having been placed in an equally interconnected group at the beginning of a simulation as usual, the agents in this experiment were allowed to adjust their social ties, and thus the interaction probabilities with other agents based on the level of communicative success experienced with any given partner. The posited hypothesis hereby was that agents would be able to achieve higher levels of success with their favourite partner(s), yet still be able to spread their lexical conventions across the population, assuming that isolated pairs or subgroups of agents do not form here.

Results obtained from this last experiment have produced a number of interesting observations. Firstly, agents with a high social learning rate, i.e. those who strongly reinforce or punish their communicative connections with other agents based on experienced success or failure, end up having one so called 'best friend' after about 200 interactions, in effect only ever interacting with that single partner from there on in. The consequence of this feature is that while some agents are 'befriended' by many of their cohabiters, others do not get picked by anyone at all, resulting in a skewed distribution of hearer-role interactions between agents in these simulations. As an effect, popular agents end up with much bigger lexicons that also include a much larger array of redundant meaning-form mappings, while agents who are never spoken to can have absolutely no synonymy or homonymy in their lexicons, while at the same time still learning a word for at least 60 different meanings. Finally, it has to be noted that while population homonymy levels obtained in the conducted experiments were just slightly higher than those observed in human languages, having synonymy at almost 100% is clearly not natural in any way.

In summary, the way early communicators are connected in a population clearly has significant ramifications for the outcome of the lexicon acquisition task. If one assumes that the use of language should have exhibited obvious advantages from the very early stages of its evolution, then given the results from this and the previous chapter, one would have to assume that early lexicons were developed in small subgroups of a population, possibly even as small as just within individual pairs. Having understood the value of communicating with the help of language, agents might have then slowly agreed on the lexical terms across their population, either via cooperation or as a consequence of group-internal natural selection.

Chapter 6

Conclusions

In the beginning of this manuscript, I have outlined some of the major challenges that the field of language evolution is facing and that it most certainly needs to overcome in order to make significant progress. In particular, obtaining any kind of meaningful empirical data in this field that could have ramifications for our understanding of the human nature, as well as provide invaluable insights in the development of the next stages of artificial intelligence, has always been one of its main challenges. In fact, up until twenty years ago, researchers in the field could solely rely on data from the fields of primatology, experimental pragmatics and language acquisition when evaluating their ideas. While these are all very closely related disciplines, it has already been mentioned in the introduction that the preconditions of the experiments performed in these fields are significantly different from those of the early stages of language evolution.

Since the advent of computing power in the late 20^{th} century, a new avenue has appeared for researchers from all disciplines to obtain additional empirical data with the help of computational models and simulations. The field of language evolution, in particular, has immense potential to benefit from this latest development, considering that computer models provide a long sought for way of obtaining (reasonably) realistic data when researching the emergence of language. As could be expected, a large number of models of the different aspects of language evolution have been proposed over the last two decades, with computational approaches basically dominating the last few editions of *"EvoLang"* – the main conference in the field.

Despite the positive nature of the developments described above, a potentially serious limitation of the way experiments with computational models of language evolution are currently conducted has been observed in section 1.2 and discussed in more detail in section 2.3. What is referred to here is the tendency of the different research groups to develop more and more new models that are usually based on strongly contrasting assumptions, making the results obtained with these largely incomparable. However, how can one make a decision as to which configuration of which model is providing the most accurate results without being able to appropriately compare the empirical data?

6.1 Computational Framework

Motivated by the issues described above, a generic computational framework for simulating experiments that deal with the emergence of language – the Language Evolution Workbench – has been proposed in this research project, described in detail in chapter 3. A particular focus within this framework has been placed on the need for any assumptions that are made within a computational model to be well situated in the corresponding field's literature. In effect, the intention is to assume nothing that is not fully agreed on by the scientific community. Instead, properties, the nature of which remains even remotely uncertain, have been provided as parameters of the model. Notably, even if certain values of some such parameters have been utilised within all of the known experiments conducted with the help of the model so far, this does not mean that these are necessarily considered to be the right values, but rather that the current state of research suggests that these settings correspond most closely to reality.

For instance, one thing that has remained at the core of the proposed framework are interactions between agents, which form the main part within the simulations conducted with the help of the LEW. However, every care has been taken not to introduce any kind of telepathic meaning-form transmission within these interactions, as has been done by Steels and Kaplan (2002) and Kirby and Hurford (2001) among others, with the lack of logic about having telepathy as the driving mechanism behind the emergence of a new communication system not requiring any further discussion. Having said that, in order for any kind of conventional system to emerge in a simulation (or in the real world), the actors that are involved in the process need to be able to learn from their observations, the actions that they make in response and the results of these actions. Otherwise, if there is no perceivable value to a new invention, it is hard to see how it could emerge other than by pure chance.

When looking for a way of introducing a telepathy-free learning mechanism into the LEW, the notion of the signaling game introduced by Lewis (1969) has served as the main inspiration, together with the scavenging-based account of language evolution proposed by Bickerton and Szathmáry (2011). In the signaling game, one agent produces a signal based on his view of the world and another agent performs an action in response to this signal. Having done so, a payoff is issued to both agents if the performed action was appropriate in the current world state, which the actor agent himself is not able to observe. In effect, the design of the signaling game situates a communication system within a real-world success-based cooperation scenario, such as, but not in any way limited to, scavenging or hunting. Within this scenario, success or failure of the interaction itself can be then deduced from the desirability of the changes in the world that were caused with the action that was performed in response to the communication attempt, without actually realising which parts of the interaction were (mis-)understood.

In summary, with the help of the signaling game concept, agents in the LEW model can be provided with information about the level of their communicative success without the need for telepathy, which was one of the primary goals of this project. In doing so, the model enables agents to learn a set of lexical items that refer to the meanings from the simulated world, effectively setting them off on the path to learning a conventional language. Having said that, the focus throughout the model remains on the task of lexicon acquisition. In particular, it is important to understand the suitability of different potential learning strategies and constraints in a variety of meaning spaces and social configurations. In order to perform such an evaluation, a small number of arguably highly influential parameters have been experimented with throughout the project while the majority of LEW's parameters have been kept fixed, as summarized further in the following sections. This approach allows one to establish the most realistic setup for the emergence of language in a reasonably small and overviewable empirical space, though without the need to fully commit to all of its properties from the start.

6.2 Lexicon Acquisition Game

For the purpose of performing a fundamental analysis of the proposed model, a set of experiments has been performed and evaluated in chapter 4 with a very basic set of settings and a limited world model. In particular, experiments have been performed with two agents and a single instance of either the one-, two- or three-argument event type. Notably, an analysis of a multi-state signaling game by Barrett (2009) has shown that the complexity of the learning task grows tremendously, even if just a third state is added to the world, with a significant number of simulations not ending up with a perfect signaling system, as opposed the 2-state version of the game, in which reaching a signaling system is guaranteed. Accordingly, the purpose of conducting experiments with a very limited number of meanings, while not necessarily reflecting reality in the best possible way, was to estimate the chances of successful communication systems emerging in the LEW model under such conditions. An additional goal was to determine an optimal level of minimum success requirement that agents (and by extension, humans) should follow at the early stages of lexicon acquisition in order to develop a lexicon that would maximize their chances of communicative success.

First of all, as presented in section 4.2, in a population of just two agents, the limits of the majority of lexical properties could be analytically established with respect to the number and types of events involved, i.e. the size of the meaning space, and the minimum success threshold that the agents followed. Effectively, these limits provided important guidance when evaluating the actual outcomes of the simulations, as they allow one to tell if the results were spaced out between the two limits, or rather skewed towards either the minimum or the maximum levels of the corresponding lexical properties.

The analysis of the conducted experiments has indicated some important findings. In particular, a strong effect of the level of minimum success has been observed on all of the lexical properties. If the success threshold was set too low, the agents tended to acquire increasingly large lexicons, without there actually being a need for it when compared to the number of distinct meanings available to them in the corresponding simulations. In these cases, agent lexicons were exhibiting synonymy levels of between 0.25 and 0.28, on average, depending on the selected event type, with the synonymy in the combined lexicons of the population as high as 0.62 and 0.77 on average, reaching 1.0 in some simulations. At the same time, population homonymy levels were significantly lower, averaging between 0.4 and 0.55, somewhat resembling the synonymy-homonymy relationship in real languages.[1]

On the other hand, if the agents required a high level of success to be experienced in order to consider that an interaction has yielded a payoff, the trends reverse dramatically, with some of the agents learning to express just one meaning out of three, six or ten available, after as many as 1000 interactions. However, in experiments with a minimum success threshold set to $s_{min} = 0.5$ and $s_{min} = 0.6$, most agents actually learn to express every meaning in their world, with the levels of population lexicon synonymy averaging at 0.29 and 0.22,[2] and homonymy between 0.13 and 0.19, respectively. Given that synonymy and homonymy levels in the English language have been estimated at 0.38 and 0.17, respectively, in section 3.4, these results clearly suggest a high degree of similarity to real languages.

Finally, in terms of actual lexicon use and precision in interactions, lower levels of minimum success threshold result in almost definite lexicon use after the initial few simulation rounds, courtesy of

[1] Agent homonymy is not possible in the utilised configuration of the LEW, as discussed at length in section 4.2.

[2] With the exception of the smallest event type, whereby any level of minimum success threshold above 0.5 fully prevents synonymy.

the large lexicons. High requirements for success, on the contrary, make sure that only the absolutely reliable lexical items are being remembered by the agents, resulting in lower levels of lexicon use, but nearly perfect lexicon precision. While it could be argued either way, which of these properties is actually more significant, the most optimal combination between the two appears to manifest itself for a minimum success level of 0.5, in a population of two agents.

Contrary to simulations with two agents, interactions performed between three or more agents have been shown to have the potential to result in an endlessly looping accumulation of lexical items (see section 4.3), making an estimation of the limits of either quantitative or qualitative lexical properties possible only with respect to the length of the conducted simulations, which was not considered particularly meaningful.[3] Regardless of that, the results of simulations with either three or ten agents, as discussed in sections 4.3 and 4.4, respectively, clearly exhibit the same trends with respect to the minimum success parameter as discussed above, albeit with some additional effects. Most notably, the lexicons of agents and the populations that they inhabit appear to grow proportionally to the population size, with some population lexicons reaching up to 600 items in simulations with just ten different meanings as a consequence of the limit-free lexicon accumulation process mentioned above.

In terms of the observed synonymy and homonymy levels, the relation between the two remains the same in simulations with three or more agents, i.e. synonymy still outweighs homonymy in all examined lexicons. Furthermore, population synonymy rates are again closest to the corresponding human language levels if the level of minimum success observed by the agents is either $s_{min} = 0.5$ or $s_{min} = 0.6$. However, lexicon homonymy in simulations with ten agents is basically non-existent for $s_{min} \geq 0.5$. While this appears to be a positive development, seeing as homonymy is more likely to impair agents in interactions rather than anything else, it does not truly correspond to the levels of homonymy observed in English e.g., which are still significantly above zero. The intuition about the impact of lexical homonymy is confirmed by the analysis of lexicon precision in the corresponding simulations, with precision going up to nearly 100% with the disappearance of homonymy.[4]

The main takeaway from the experiments performed with the extended groups of agents is that it is extremely hard for them to agree on a population-wide conventional signaling system, no matter what level of minimum success they consider as appropriate within interactions. In general, the lexicons of the simulated populations grow excessively large if the success requirement is too low, and do not manage to spread the learned words throughout the population, in case a high level of success is deemed necessary. These observations appear to put one of the following two options into doubt. Either language did not evolve through one-on-one dialogue, but rather via one-to-many group-internal monologue-like communication attempts; or it was the particular social organisation of early hominids that aided them in learning a reliable lexicon within a group. Finally, despite what has been concluded above, lexicons that strongly resemble those of real languages can certainly emerge within the proposed framework, notably without any need for telepathy, indicating that it could provide valuable insights into the language evolution process, naturally subject to further investigation.

[3] A calculation of limit values based on the number of interactions in a simulation would result in extremely high lexicon size limits, which would not be at all representative of the actual probability distribution of the lexicon sizes developed throughout such simulations.

[4] Admittedly, a higher success requirement also plays its role in ensuring that only the more reliable lexical items are stored in the agents' lexicons.

6.3 Social Structures

In the following set of experiments, the latter of the two options presented above was investigated from two perspectives. The first of these considered if arranging agents into pairs throughout the simulation runs would help them fully develop reliable lexicons within the pairing, while at the same time enabling them to distribute the lexical items across the group with the help of increased confidence levels obtained in pair-internal interactions. Within this experiment, the pairings have been either defined from the start, with three different pair-internal interaction rates (25%, 50%, 75%), or rotated between successive agents throughout the simulations runs.

Under both conditions outlined above, the first observation that has been made is that, perhaps expectedly, pair-specific dialects appear to be emerging during the simulations. Naturally, individual dialects have a direct influence on the size of the combined population lexicon, as well as the levels of synonymy in it. The levels of lexicon use and precision observed in these experiments are also quite low, with the effects of having a preferred interaction partner clearly not being carried over onto interactions between agents from different pairings, driving overall success rates down. However, if agents are allowed to exclusively communicate within a pairing for a certain period of time before moving on to the next partner, the pair-internal success is beginning to influence the population as a whole, indicating that agents need more time to agree on a proper lexicon with a small set of partners before they can move on to communicating with other members of the group. Otherwise, if one is constantly talking to someone new, chances are that, as misunderstandings keep constantly occurring, agents will never be able to achieve sufficient confidence levels in their lexical items to persist with these and eventually reach agreement about their use with others.

Motivated by the encouraging results from the alternating pair condition, the second option that was evaluated in this set of experiments turned to the possibility of agent pairings or subgroups emerging dynamically within the individual simulation runs. This experiment is based on an observation made throughout the research project that most simulations of agent interactions have participants arranged in fixed social networks from the very beginning. In some models, agents can be added to and eliminated from the population, but the social connections, i.e. the interaction rates between agents never change. It is understandable that integrating a feature like this into a model is not necessarily straightforward, especially considering the increased complexity of evaluating the results. However, reciprocity is certainly one of the dominant features of cooperative behaviour in many species (cf. Nowak, 2006; Gärdenfors, 2007), suggesting that it could have at least implicitly influenced the process of language evolution as well. Accordingly, it is postulated that it is natural for agents to adjust their social ties with others based on the success of mutual interactions, with three different rates of such adjustment being tested.

Simulations of this final experiment have resulted in some noteworthy observations. Agents who are quick to adjust their connections with other agents end up having one so called ‘best friend’ in the very early stages of the simulations already. However, since friendships are defined as unidirectional relationships in the current version of the experiment, and since any agent can end up picking any other as his friend, some agents may get ‘befriended’ by many others, while some may not be anybody’s friend. Consequently, the distribution of the number of hearer-role interactions per agent becomes skewed in such groups, with popular agents ending up having an overflow of lexical items in their lexicons, including multiple redundant meaning-form mappings. At the same time, the solitary agents who are never spoken to can develop lexicons representing a significant portion of the world’s different meanings, yet without any synonymy or homonymy in them.

Remarkably, the dynamics and the final states of the social structures that have developed in the simulations with adjustable agent connections clearly resemble those observed in human social groups. In particular, just as in human experiments conducted by Panzarasa et al. (2009), agents in the conducted simulations get quickly familiar and build up their relationships with one or two of their conspecifics, without significantly adjusting their social ties at the later stages. Furthermore, the overall emerging distribution of social connections between agents exhibits a distinctive scale-free character, which has been observed by Barabási (2009) in numerous human social groups.

In summary, the way in which early adopters of language are connected within their populations clearly has significant ramifications for the success of the lexicon acquisition task. In particular, given that it is unreasonable to assume that pioneers of language would have kept trying to make sense of it for too long if it did not provide them with any clear benefits from very early on, the conclusion that has to be made based on the observations from the experiments conducted in this research project is that, in its first stages, language was a tool used by very small subgroups of a population, possibly involving as few as two or three members. Having established itself as a reliable tool for communication between these inventive individuals, language has then spread throughout the whole population, possibly as a result of close cooperation between certain groups, or as a consequence of group-internal natural selection and generation change.

In effect, the above conclusion suggests that the sudden growth of hominid group sizes at one stage in their evolution was just the initial trigger for the invention of a new communication system. Having been set off by this trigger, the first forms of human language then emerged in small unions within the large groups and only afterwards gradually spread out to whole populations. At the same time, there still exists a significant number of alternative scenarios that could have taken place at the early stages of language evolution, some of which that are directly related to the experiments presented in this project and that could thus be seen as direct extensions of it, are presented in section 6.5.

6.4 Model Comparison

The introduction of yet another computational model of language evolution was motivated by a number of relatively questionable assumptions made in some of the models that exist to date, as discussed in detail in section 2.3. The LEW was thus designed to allow its users to configure an evolutionary scenario based on the assumptions that one deemed acceptable, leaving the remaining parameters of the model open to experimentation. In the case of the current research project, this meant replacing explicit meaning-form transmission with a task-oriented success-based learning strategy, while keeping the population size and the meaning space relatively small. These latter two settings keep the model in the same class as the ILM and the Language Games, allowing one to compare the results between the different approaches.

At first glance, judging by the results reported in chapter 4, it would seem that the communicative ability of agents in the LEW, at least in the presented configurations, is significantly inferior to the nearly perfect linguistic systems that emerge either in the ILM or the Language Games. While this is certainly the case in a direct comparison of the levels of communicative success and ambiguity in the emergent languages (in the form of synonymy and homonymy), this does not automatically imply that the setup of the LEW is inferior to the other two models. As has been remarked by Smith (2005) among others, language variation and change are two quite integral (if not unequivocally efficient) features of any human language. What this implies is that if one wishes to evaluate a model of the emergence of *human* language, then one should not compare it to flawless information transmission systems, but rather to the quite ambiguous and ever changing communication systems that human languages are.

When doing so, it turns out that for certain configurations of the LEW, e.g. $N = 2; 0.5 \leq s_{min} \leq 0.6$, identified in section 4.2, the observed levels of communicative success, as well as synonymy and homonymy are much more resemblant of a human language than those from the final states of the ILM or the Language Games, where the communication is nearly perfect and there is no ambiguity at all. It cannot be denied that the majority of the presented LEW configurations result in communication systems with characteristics that are unnatural for human languages. The most common issue hereby appears to be finding the balance between knowing how to express a sufficient amount of meanings and being confident enough about the acquired linguistic conventions. However, the fact that certain configurations of the LEW lead agents to develop lexicons, the properties of which are quite similar to those of modern human languages, in contrast to the ILM and the Language Games, suggests that the LEW is closer to simulating the evolution of our seemingly imperfect, often ambiguous, and wonderfully unpredictable communication tool than its current competitors.

6.5 Future Directions

Based on the results from the current research project, it is clear that the underlying organization of agents can play a significant role on the efficiency of the learning process, as well as the qualities of the emergent lexicons within the community. Accordingly, two further extensions that appear to have high potential within this experimental path are proposed here, both of which rely on a natural separation of members of a group into different types with respect to their social status. The first extension is based on the observation made by Barabási and Bonabeau (2003) that human social groups usually have a number of highly interconnected agents that are either the most popular targets or the most active sources of all interactions within the group (or both). In evolutionary terms, such members could be naturally selected based on their alpha-male status, or as a consequence of certain character features, i.e. being generally friendly and likeable.

The claim that is being put forward here is that having one or several agents in a group that are involved in significantly more interactions than others would allow them to implicitly guide the entire population towards a more uniform lexicon. However, there is no reason why the influence of the popular group members has to be necessarily implicit. In fact, Latané (1981) has shown that dominant individuals within a group can actually directly influence the behaviour of others.[5] Accordingly, it would be plausible to assume that 'average' agents would perceive success or failure in interactions with an influential agent as more pronounced than usual. As a consequence, it can be expected that a single dominant agent should be able to basically impose a set of more or less unique lexical items on a small group of around ten agents.[6] Finally, while the two avenues described above appear to have particularly high potential, numerous other social configurations could also account for the emergence of language and should most certainly be considered in future research.

[5] Again, either by being in an actual position of power or simply by having superior reasoning and oratory skills.

[6] While this particular configuration is similar to a 'dictatorial' power structure, as proposed by Gärdenfors (1993), the influential agents are in no way capable of actually imposing the 'correct' use of lexical items on other members of the group, meaning there is never a 'semantic arbiter'.

Appendix A

LEW Parameters

This section provides a more detailed and technical description of the parameters of the LEW model, based on the description of selected model assumptions, data structures and mechanisms provided in chapter 3. In effect, a complete enumeration of the different parameters involved in the model is meant to exhibit its possibilities as well as help an interested reader understand how to operate the workbench, which is also publicly available under the Creative Commons Attribution-ShareAlike 3.0 Unported (CC BY-SA 3.0) license at http://github.com/arski/LEW. At the same time, the list below underlines the level of detail that has been already reached within the proposed framework and that can be hopefully increased even further in the future.

Population

Agents: number of interacting agents in the simulated population. *Options:* any positive integer.

Speaker selection: strategy employed when selecting the speaker in an interaction. *Options:* random selection or turn-based selection, with the latter guaranteeing that all agents enter exactly the same number of interactions as speakers.

Hearer selection: strategy employed when selecting the speaker in an interaction. *Options:* random selection or turn-based selection. The second option assigns a new fixed partner to a speaker at the turn of an 'epoch' (see below).

Self-talk: defines if an agent can act as both speaker and hearer in an interaction. *Options:* on or off.

Interaction type: specifies how many agents can act as hearers in an interaction. *Options:* dialogue mode, implying a single hearer, or monologue mode, whereby all agents bar the speaker are present as hearers in every interaction.

Overhearers: similar to the above, but allows to define how many additional hearers should be present in an interaction next to the addressee. Overhearers can also differ from the addressee in terms of their learning strategy (see next parameter). *Options:* any non-negative integer value.

Overhearer success: defines how overhearers should calculate success of an interaction. *Options:* self-based, i.e. based on the accuracy of own interpretation of the speaker's utterance; or hearer-based, utilising the success of the speaker-hearer interaction that was overheard.

Group size ratios: defines the size distribution between groups, as well as the actual number of groups in the population. *Options:* any list of positive integers, with the length of the list not exceeding overall population size.[1]

Group speak ratios: assigns a probability of an agent from the corresponding group being selected as a speaker, used for the definition of various types of agents and agent groups in a population. *Options:* any list of non-negative integers, with the length of the list being equal to the number of groups specified by the preceding parameter.

Group interaction ratios: specifies the probability distribution of hearer group selection for every potential speaker group. *Options:* a nested list of lists, both requiring to be the size of the number of groups and the inner list, taking non-negative integers.

Influence ratios: meant to define the influence, i.e. an additional payoff multiplicator in success-based learning, between individual groups. Currently unused and untested.

World Model

Phones: number of distinct phones in the model. *Options:* any positive integer value, whereby an extremely high value can be used to ensure that the same phoneme is never generated from the set of phones.

Entities: number of entities in the world. *Options:* any positive integer.

Event types: which of the available event types should be used when building up the world. *Options:* a list of any of the five event indexes. In particular, 1-3 represent the three basic types, while 4 and 5 represent types 2 and 3 with a recursive second argument slot, respectively.

Events: number of different (uninstantiated) events of the above types. *Options:* any positive integer.

Event instances: actual number of event instances generated in the world model. *Options:* any positive integer.

Zipfism: defines if the event selection probabilities should be distributed along a Zipfian distribution. *Options:* yes or no.

Interactions

Memory lookup: strategy that should be used if one of multiple acceptable mappings needs to be selected from one's lexicon. *Options:* frequency-based, recency-based.

Random form generation: defines the interval at which a new form should be generated by the speaker, even if he has one for the particular meaning in his lexicon. *Options:* any non-negative integer.

[1]This and the following two ratio parameters all use proportions to calculate probabilities, mainly as a consequence of this being the easiest way to compute probabilities in Prolog. It is conceded that the human readability of these parameters suffers from this approach.

Force segmentation: should events and utterances be divided into at least two parts, thus forbidding one-part event partitions and one-word utterance segmentations. *Options:* yes or no.

Synchronous segmentation: defines if the hearer has direct access to the intended segmentation of the speaker's utterance into words. *Options:* on or off.

Synchronous segmentation (self): same as the above, but applies to self-talk instances.

Interpretation rule: defines additional constraints on the hearer's interpretation of an utterance. *Options:* random, principle of contrast, omniscient, wrong. Last two options are provided for testing purposes only.

Interpretation rule (self): same as the above, but applies to self-talk instances.

Event role: role of event during interpretation. *Options:* ignore, prefer, event-only.

Reformulations: should the hearer be allowed to ask for an alternative formulation of the utterance. *Options:* yes or no.

Learning

Success matters: if turned on, only interactions exceeding the level of minimum success (see next) will be recorded by the agents. *Options:* yes or no.

This parameter is obsolete and is replaced by the lexicon update parameter set.

Minimum success: the minimum level of success that needs to be reached by an interaction for it to be recorded by the agents, in the case that the preceding parameter is enabled. *Options:* any decimal value between 0 and 1.

This parameter is obsolete and is replaced by the lexicon update parameter set.

Lexicon update: payoff (update value) calculation strategy used for updating lexical mapping weights, i.e. learning the lexicon. *Options:* none, fixed, fixed with s_{min}, relative to s_{min}.

Lexicon update threshold: the minimum success threshold s_{min} used when calculating the update value in the latter two strategies listed for the preceding parameter. *Options:* any decimal value between 0 and 1.

Lexicon update rate: a multiplier applied to the payoff value for the purpose of increasing or reducing the learning rate of the above two parameters. *Options:* any positive value.

Interaction ratio update: payoff (update value) calculation strategy used for updating interaction ratios between groups. *Options:* none, fixed, fixed with s_{min}, relative to s_{min}.

Interaction ratio update threshold: the minimum success threshold s_{min} used when calculating the update value in the latter two strategies listed for the preceding parameter. *Options:* any decimal value between 0 and 1.

Interaction ratio update rate: a multiplier applied to the payoff value for the purpose of increasing or reducing the learning rate of the above two parameters. *Options:* any positive value.

Influence update: payoff (update value) calculation strategy used for updating influence levels between groups. *Options:* none, fixed, fixed with s_{min}, relative to s_{min}.

Influence update threshold: the minimum success threshold s_{min} used when calculating the update value in the latter two strategies listed for the preceding parameter. *Options:* any decimal value between 0 and 1.

Influence update rate: a multiplier applied to the payoff value for the purpose of increasing or reducing the learning rate of the above two parameters. *Options:* any positive value.

Extras

Forgetting function: a decay function that should be applied between epochs to reduce the weights of all agent mappings. The inverse of the value returned by the function is subtracted from the mapping weights. *Options:* none, constant, linear and polynomial.

Forgetting factor: used as the main value in the above functions. *Options:* normally, 1 is used for constant and linear forgetting, 2 is used for polynomial forgetting.

Forgetting offset: value that should be added to the denominator of the forgetting element, consequently reducing the effect of forgetting. *Options:* any value, though supplying a negative value might have adverse effects.

Agent addition: should agents be added between epochs. *Options:* yes or no.

New agents: maximum number of agents to be added, provided previous parameter is enabled. *Options:* any positive integer.

Agent elimination: should agents be eliminated between epochs. *Options:* yes or no.

Agent fitness: should a fitness value be calculated and utilised during agent elimination. *Options:* on or off.

Fitness threshold: defines the level of fitness, under which the agents are considered to be unfit and are subject to elimination. *Options:* any non-negative decimal.

Age range: number of interactions that the agent should be able to perform before falling below the fitness threshold. *Options:* any positive integer.

System

Epochs: number of epochs that a simulation run should be divided into, with every epoch consisting of as many interactions as there are agents, multiplied by the epoch factor (see next). *Options:* any positive integer.

Epoch multiplier: multiplies the length of an epoch, thus increasing the number of interactions within a simulation. *Options:* any positive integer.

Epoch span: number of epochs in one ‘span’, after which intermediate tasks such as agent addition and removal, as well as forgetting, are executed, provided any of these is enabled. *Options:* any positive integer that is a factor of the epoch count.

Stats span: number of epochs in a special 'span' used for the generation of intermediate statistics only. *Options:* any positive integer that is a factor of the epoch count.

Group stats: should group-specific lexicon statistics be calculated or not. *Options:* on or off.

Print trace: print trace of the experiment on screen. *Options:* on or off.

Appendix B

Experiment Configurations

The following appendix provides an overview of the parameter values employed in the experiments that have been presented throughout this manuscript, as presented in chapters 4 and 5. The main purpose of providing this information is to allow anyone to reproduce the experiments with the help of the LEW model, as well as to build on top of the discussed configurations with ideas of one's own. First, table B.1 lists all of the LEW parameters, with the levels of those parameters that have remained fixed across all of the presented experiments stated directly, the parameters that have varied as experiment factors or which have simply taken up different values as a consequence of the selected configurations being marked as 'VAR' and the unused/unneeded parameters filled with 'N/A'. Following that, the specific parameter values selected in the individual configurations are outlined, together with the ranges of the factors that have been experimented with in the corresponding simulations.

Section	Parameter	Value
Population	Agents	VAR
	Speaker selection	random
	Hearer selection	VAR
	Self-talk	off
	Interaction type	dialogue
	Overhearers	0
	Overhearer success	N/A
	Group size ratios	VAR
	Group speak ratios	VAR
	Group interaction ratios	VAR
	Influence ratios	N/A
World model	Phones	500
	Entities	VAR
	Event types	VAR
	Events	VAR
	Event instances	VAR
	Zipfism	off

Section	**Parameter**	**Value**
Interactions	Memory lookup	frequency-based
	Random form generation	off
	Force segmentation	no
	Synchronous segmentation	on
	Synchronous segmentation (self)	N/A
	Interpretation rule	random
	Interpretation rule (self)	N/A
	Event role	ignore
	Reformulations	no
Learning	Lexicon update	relative to s_{min}
	Lexicon update threshold	VAR
	Lexicon update rate	1
	Interaction ratio update	VAR
	Interaction ratio update threshold	0.35
	Interaction ratio update rate	VAR
	Influence update	none
	Influence update threshold	N/A
	Influence update rate	N/A
Extras	Forgetting function	none
	Forgetting factor	N/A
	Forgetting offset	N/A
	Agent addition	off
	New agents	N/A
	Agent elimination	off
	Agent fitness	N/A
	Fitness threshold	N/A
	Age range	N/A
System	Epochs	500
	Epoch multiplier	1
	Epoch span	50
	Stats span	50
	Group stats	on
	Print trace	off

Table B.1: Fixed parameters

Interacting in Pairs

The first experiment presented in section 4.2 featured just two agents in a world with a single event. The first factor of the experiment was the type of the single event, with the three basic, non-recursive types being experimented with. Since the three event types take one, two or three arguments, respectively, the number of entities was varied accordingly, though always kept to the minimum required for the single event. Furthermore, the minimum success threshold s_{min} that the agents needed to achieve in their interactions in order to obtain a payoff was varied between 0 and 0.9, in steps of 0.1, resulting in a total of 30 different configurations. Table B.2 lists the values utilized for the parameters that were experimented with across all of the experiments, with the factors explored in the particular experiment being highlighted in italics.

Table B.2: Fixed parameters

Parameter	**Value**
Agents	2
Hearer selection	random
Group size ratios	[1]
Group speak ratios	[1]
Group interaction ratios	[[1]]
Entities	*1 -vs- 2 -vs- 3*
Event types	*[1] -vs- [2] -vs- [3]*
Events	1
Event instances	1
Lexicon update threshold	*0-0.9 in steps of 0.1*
Interaction ratio update	none
Interaction ratio update rate	N/A

Interacting in Triads

The experiment presented in section 4.3 was an exact copy of the preceding one, with the exception of the simulations being conducted with three different agents, as opposed to two. Table B.3 presents the model configuration used for this experiment, along with the varied parameters.

Table B.3: Fixed parameters

Parameter	**Value**
Agents	3
Hearer selection	random
Group size ratios	[1]
Group speak ratios	[1]
Group interaction ratios	[[1]]
Entities	*1 -vs- 2 -vs- 3*
Event types	*[1] -vs- [2] -vs- [3]*
Events	1
Event instances	1
Lexicon update threshold	*0-0.9 in steps of 0.1*
Interaction ratio update	none
Interaction ratio update rate	N/A

Interacting in Small Groups

The final experiment from the fundamental analysis chapter, discussed in section 4.4, extends the number of agents even further to ten, as can be seen from table B.4.

Table B.4: Fixed parameters

Parameter	**Value**
Agents	10
Hearer selection	random
Group size ratios	[1]
Group speak ratios	[1]
Group interaction ratios	[[1]]
Entities	*1 -vs- 2 -vs- 3*
Event types	*[1] -vs- [2] -vs- [3]*
Events	1
Event instances	1
Lexicon update threshold	*0-0.9 in steps of 0.1*
Interaction ratio update	none
Interaction ratio update rate	N/A

Population Subroups

The first of the experiments involving slightly more complex configurations involves two different approaches, as presented in section 5.2. The first of these has a group of ten agents paired up, with the group interaction rates defining the probability of these agents talking within these pair, as depicted by table B.5.

Table B.5: Fixed parameters

Parameter	**Value**
Agents	10
Hearer selection	*random*
Group size ratios	[1,1,1,1,1]
Group speak ratios	[1,1,1,1,1]
Group interaction ratios	*[[12,1,1,1,1],[1,12,1,1,1],[1,1,12,1,1],[1,1,1,12,1],[1,1,1,1,12]]* *-vs-* *[[4,1,1,1,1],[1,4,1,1,1],[1,1,4,1,1],[1,1,1,4,1],[1,1,1,1,4]]* *-vs-* *[[4,3,3,3,3],[3,4,3,3,3],[3,3,4,3,3],[3,3,3,4,3],[3,3,3,3,4]]*
Entities	25
Event types	[1,2,3]
Events	10
Event instances	50
Lexicon update threshold	0.35
Interaction ratio update	none
Interaction ratio update rate	N/A

The second approach from section 5.2 is quite similar in nature, though in this case, the pairings were not fixed throughout the simulation runs, but successively alternated for each epoch, based on the configuration shown in table B.6.

Table B.6: Fixed parameters

Parameter	**Value**
Agents	10
Hearer selection	*turns*
Group size ratios	[1,1,1,1,1]
Group speak ratios	[1,1,1,1,1]
Group interaction ratios	*[[1,0,0,0,0],[0,1,0,0,0],[0,0,1,0,0],[0,0,0,1,0],[0,0,0,0,1]]*
Entities	25
Event types	[1,2,3]
Events	10
Event instances	50
Lexicon update threshold	0.35
Interaction ratio update	none
Interaction ratio update rate	N/A

Friendships

Finally, the experiment presented in section 5.3 allowed for the social ties between agents to be adjusted between interactions, based on the level of success obtained with the partner. The model settings utilised for this experiment are presented below in table B.7.

Table B.7: Fixed parameters

Parameter	**Value**
Agents	10
Hearer selection	random
Group size ratios	[1,1,1,1,1,1,1,1,1,1]
Group speak ratios	[1,1,1,1,1,1,1,1,1,1]
Group interaction ratios	[[0,1,1,1,1,1,1,1,1,1],[1,0,1,1,1,1,1,1,1,1],[1,1,0,1,1,1,1,1,1,1],.. [1,1,1,0,1,1,1,1,1,1],[1,1,1,1,0,1,1,1,1,1],[1,1,1,1,1,0,1,1,1,1],.. [1,1,1,1,1,1,0,1,1,1],[1,1,1,1,1,1,1,0,1,1],[1,1,1,1,1,1,1,1,0,1],.. [1,1,1,1,1,1,1,1,1,0]]
Entities	25
Event types	[1,2,3]
Events	10
Event instances	50
Lexicon update threshold	0.35
Interaction ratio update	relative to s_{min}
Interaction ratio update rate	*1.0 -vs- 0.5 -vs- 0.1*

References

Aiello, L. C., & Dunbar, R. (1993). Neocortex Size, Group Size, and the Evolution of Language. *Current Anthropology*, *34*(2), 184–193.

Argiento, R., Pemantle, R., Skyrms, B., & Volkov, S. (2009). Learning to signal: Analysis of a micro-level reinforcement model. *Stochastic Processes and their Applications*, *119*(2), 373–390.

Atkinson, Q. D. (2011). Phonemic Diversity Supports a Serial Founder Effect Model of Language Expansion from Africa. *Science*, *332*(6027), 346–349.

Bachwerk, M. (2011). Reinforcement Learning in Language Evolution. In *Proceedings of the ESSLLI 2011 Student Session.*

Bachwerk, M., & Vogel, C. (2010). Modelling Social Structures and Hierarchies in Language Evolution. In M. Bramer, M. Petridis, & A. Hopgood (Eds.), *Research and Development in Intelligent Systems XXVII* (pp. 49–62). Springer.

Bachwerk, M., & Vogel, C. (2011). Establishing Linguistic Conventions in Task-Oriented Primeval Dialogue. In A. Esposito, A. Vinciarelli, K. Vicsi, C. Pelachaud, & A. Nijholt (Eds.), *Analysis of Verbal and Nonverbal Communication and Enactment: The Processing Issues* (pp. 48–55). Springer.

Bachwerk, M., & Vogel, C. (2012). Language and Friendships: A Co-Evolution Model of Social and Linguistic Conventions. In T. C. Scott-Phillips, M. Tamariz, E. A. Cartmill, & J. R. Hurford (Eds.), *The Evolution of Language, Proceedings of the 9th International Conference (EVOLANG9)* (pp. 34–41). World Scientific.

Barabási, A.-L. (2009). Scale-Free Networks: A Decade and Beyond. *Science*, *325*(5939), 412–413.

Barabási, A.-L., & Bonabeau, E. (2003). Scale-free networks. *Scientific American*, *288*(5), 50–59.

Baronchelli, A., Loreto, V., & Steels, L. (2008). In-Depth Analysis of the Naming Game Dynamics: The Homogeneous Mixing Case. *International Journal of Modern Physics C*, *19*(5), 785–812.

Barrett, J. A. (2006). *Numerical Simulations of the Lewis Signaling Game: Learning Strategies, Pooling Equilibria, and the Evolution of Grammar.*

Barrett, J. A. (2009). The Evolution of Coding in Signaling Games. *Theory and Decision*, *67*(2), 223–237.

Barrett, J. A., & Zollman, K. J. S. (2009). The Role of Forgetting in the Evolution and Learning of Language. *Journal of Experimental & Theoretical Artificial Intelligence*, *21*(4), 293–309.

Bickerton, D. (1990). *Language and Species.* University of Chicago Press.

Bickerton, D. (2002). Foraging Versus Social Intelligence in the Evolution of Protolanguage. In A. Wray (Ed.), *The Transition to Language* (pp. 207–225). Oxford University Press.

Bickerton, D. (2003). Symbol and Structure: A Comprehensive Framework for Language Evolution. In M. H. Christiansen & S. Kirby (Eds.), *Language Evolution: The States of the Art* (p. 77-93). Oxford University Press.

Bickerton, D. (2007). Language evolution: A brief guide for linguists. *Lingua*, *117*(3), 510–526.

Bickerton, D., & Szathmáry, E. (2011). Confrontational scavenging as a possible source for language and cooperation. *BMC Evolutionary Biology*, *11*(1), 261.

Burling, R. (2005). *The Talking Ape: How Language Evolved.* Oxford University Press.

Byrne, R. W., & Whiten, A. (1988). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes, and Humans.* Oxford University Press.

Cangelosi, A., & Parisi, D. (1998). The Emergence of a 'Language' in an Evolving Population of Neural Networks. *Connection Science*, *10*(2), 83–97.

Carstairs-McCarthy, A. (1999). *The Origins of Complex Language: An Inquiry into the Evolutionary Beginnings of Sentences, Syllables, and Truth.* Oxford University Press.

Chase, H. W., & Clark, L. (2010). Gambling severity predicts midbrain response to near-miss outcomes. *The Journal of Neuroscience*, *30*(18), 6180–7.

Clark, E. V. (1987). The principle of contrast: A constraint on language acquisition. In B. MacWhinney (Ed.), *Mechanisms of Language Acquisition* (pp. 1–33). Erlbaum.

Clark, H. H., & Wilkes-Gibbs, D. (1986). Referring as a collaborative process. *Cognition*, *22*, 1–39.

Corballis, M. (2000). *The Descent of Mind: Psychological Perspectives on Hominid Evolution.* Oxford University Press.

de Boer, B., & Zuidema, W. (2009). Models of Language Evolution: Does the Math Add Up? In *ILLC Preprint Series (PP-2009-49).*

de Lara, J., & Alfonseca, M. (2002). The role of oblivion, memory size and spatial separation in dynamic language games. *Journal of Artificial Societies and Social Simulation*, *5*(2).

Dessalles, J.-L. (2007). *Why We Talk: The Evolutionary Origins of Language.* Oxford University Press.

Deuchar, M., & Quay, S. (2001). *Bilingual Acquisition: Theoretical Implications of a Case Study.* Oxford University Press.

Dunbar, R. I. M. (1997). *Grooming, Gossip and the Evolution of Language.* Harvard University Press.

Fitch, W. T. (2010). *The Evolution of Language.* Cambridge University Press.

Fomin, F. V., Grandoni, F., Pyatkin, A. V., & Stepanov, A. A. (2008). Combinatorial bounds via measure and conquer: Bounding minimal dominating sets and applications. *ACM Transactions on Algorithms*, *5*(1), 9:1–17.

Galantucci, B. (2005). An Experimental Study of the Emergence of Human Communication Systems. *Cognitive Science*, *29*, 737–767.

Galantucci, B. (2009). Experimental Semiotics: A New Approach for Studying Communication as a Form of Joint Action. *Topics in Cognitive Science*, *1*(2), 393–410.

Gärdenfors, P. (1993). The Emergence of Meaning. *Linguistics and Philosophy*, *16*(3), 285 – 309.

Gärdenfors, P. (2004). Cooperation and the Evolution of Symbolic Communication. In K. D. Oller & U. Griebel (Eds.), *Evolution of Communication Systems: A Comparative Approach* (pp. 237–256). MIT Press.

Gärdenfors, P. (2007). The Cognitive and Communicative Demands of Cooperation. In T. Rø nnow Rasmussen, B. Petersson, J. Josefsson, & D. Egonsson (Eds.), *Hommage à Wlodek. Philosophical Papers Dedicated to Wlodek Rabinowic.*

Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition*, *27*, 181–218.

Garrod, S., & Doherty, G. (1994). Conversation, co-ordination and convention: an empirical investigation of how groups establish linguistics conventions. *Cognition*, *53*, 181–215.

Garrod, S., Fay, N., Lee, J., Oberlander, J., & Macleod, T. (2007). Foundations of Representation: Where Might Graphical Symbol Systems Come From? *Cognitive Science*, *31*, 961–987.

Gathercole, V. C. M. (1998). Lexical Constraints in Acquisition: Are They Viable Any Longer? In C. M. Gruber, D. Higgins, K. S. Olson, & T. Wysocki (Eds.), *Proceedings of the Chicago Linguistics Society* (pp. 481–492). Chicago Linguistic Society.

Gong, T. (2010). Exploring the Roles of Horizontal, Vertical, and Oblique Transmissions in Language Evolution. *Adaptive Behavior*, *18*(3-4), 356–376.

Gong, T., Ke, J., Minett, J. W., & Wang, W. S.-Y. (2004). A Computational Framework to Simulate the Coevolution of Language and Social Structure. In J. Pollack, M. Bedau, P. Husbands, T. Ikegami, & R. A. Watson (Eds.), *Artificial Life IX* (pp. 158–163). MIT Press.

Gong, T., Minett, J. W., & Wang, W. S.-Y. (2008). Exploring social structure effect on language evolution based on a computational model. *Connection Science*, *20*(2), 135–153.

Gong, T., Minett, J. W., & Wang, W. S.-Y. (2009). Coevolution of Language and Intentionality Sharing. In *Proceedings of the Eleventh conference on Congress on Evolutionary Computation* (pp. 1530–1537). IEEE Press.

Gong, T., Puglisi, A., Loreto, V., & Wang, W. S.-Y. (2008). Conventionalization of Linguistic Knowledge Under Communicative Constraints. *Biological Theory*, *3*(2), 154–163.

Gong, T., & Wang, W. S.-Y. (2005). Computational Modeling on Language Emergence: A Coevolution Model of Lexicon, Syntax and Social Structure. *Language and Linguistics*, *6*(1), 1–42.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The Faculty of Language: What Is It, Who Has It, and How Did It Evolve? *Science (New York, N.Y.)*, *298*(5598), 1569–1579.

Hazlehurst, B., & Hutchins, E. (1998). The Emergence of Propositions from the Co-ordination of Talk and Action in a Shared World. *Language and Cognitive Processes*, *13*(2), 373–424.

Healey, P. G. T., Swoboda, N., Umata, I., & Katagiri, Y. (2002). Graphical representation in graphical dialogue. *International Journal of Human-Computer Studies*, *57*, 375–395.

Hedetniemi, S., & Laskar, R. (1990). Bibliography on domination in graphs and some basic definitions of domination parameters. *Discrete Mathematics*, *86*(1-3), 257–277.

Heine, B., & Kuteva, T. (2007). *The Genesis of Grammar: A Reconstruction.* Oxford University Press.

Herrnstein, R. J. (1970). On the law of effect. *Journal of the Experimental Analysis of Behavior*, *13*, 243 –266.

Hurford, J. R. (1989). Biological Evolution of the Saussurean Sign as a Component of the Language Acquisition Device. *Lingua*, *77*(2), 187–222.

Hurford, J. R. (1991). The evolution of the critical period for language acquisition. *Cognition*, *40*(3), 159–201.

Hutchins, E., & Johnson, C. M. (2009). Modeling the Emergence of Language as an Embodied Collective Cognitive Activity. *Topics in Cognitive Science*, *1*(3), 523–546.

Huttegger, S. M., & Zollman, K. J. S. (2011). Signaling Games: Dynamics of Evolution and Learning. In A. Benz, C. Ebert, G. Jäger, & R. van Rooy (Eds.), *Language, Games, and Evolution* (pp. 160–176). Springer.

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends in Cognitive Sciences*, 272–279.

King, B. J. (1999). *The Origins of Language: What Nonhuman Primates Can Tell Us.* SAR Press.

Kirby, S. (2001). Spontaneous Evolution of Linguistic Structure: An Iterated Learning Model of the Emergence of Regularity and Irregularity. *IEEE Transactions on Evolutionary Computation*, *5*(2), 102–110.

Kirby, S. (2002). Learning, Bottlenecks and the Evolution of Recursive Syntax. In T. Briscoe (Ed.), *Linguistic Evolution through Language Acquisition: Formal and Computational Models* (pp. 173–203). Cambridge University Press.

Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(12), 5241–5.

Kirby, S., & Hurford, J. R. (2001). The Emergence of Linguistic Structure: An Overview of the Iterated Learning Model. In *Simulating the Evolution of Language* (pp. 53–78).

Kudo, H., & Dunbar, R. (2001). Neocortex size and social network size in primates. *Animal Behaviour*, *62*(4), 711–722.

Latané, B. (1981). The Psychology of Social Impact. *American Psychologist*, *36*(4), 343–356.

Lewis, D. (1969). *Convention: A Philosophical Study.* Harvard University Press.

Lindh-Knuutila, T., Honkela, T., & Lagus, K. (2006). Simulating Meaning Negotiation Using Observational Language Games. In P. Vogt, Y. Sugita, E. Tuci, & C. Nehaniv (Eds.), *Symbol Grounding and Beyond* (pp. 168 – 179). Springer Berlin Heidelberg.

Longmore, L. (2008). *Talk about Talk, Population Dynamics and Agent Fitness.* Unpublished master's thesis, Trinity College Dublin.

MacWhinney, B. (2005). A Unified Model of Language Acquisition. In J. Kroll & A. de Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches.* Oxford University Press.

Markman, E. M. (1989). *Categorization and naming in children: Problems of induction.* MIT Press.

Mirolli, M., & Nolfi, S. (2010). Evolving communication in embodied agents: Theory, Methods, and Evaluation. In S. Nolfi & M. Mirolli (Eds.), *Evolution of Communication and Language in Embodied Agents* (pp. 105–121). Springer.

Nettle, D. (1999). Using Social Impact Theory to simulate language change. *Lingua*, *108*, 95–117.

Nowak, M. A. (2006). Five Rules for the Evolution of Cooperation. *Science*, *314*(5805), 1560–1563.

Oliphant, M. (1999). The learning barrier: Moving from innate to learned systems of communication. *Adaptive Behavior*, *7*(3/4), 371–384.

Panzarasa, P., Opsahl, T., & Carley, K. M. (2009). Patterns and Dynamics of Users Behaviour and Interaction: Network Analysis of an Online Community. *Journal of the American Society for Information Science and Technology*, *60*(5), 911–932.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, *27*, 169–190.

Pika, S., & Mitani, J. C. (2009). The directed scratch: evidence for a referential gesture in chimpanzees? In R. Botha & C. Knight (Eds.), *The Prehistory of Language* (pp. 166–180). Oxford University Press.

Pinker, S. (1994). *The Language Instinct.* W. Morrow and Co.

Quine, W. V. O. (1960). *Word and Object.* MIT Press.

Ragir, S., & Savage-Rumbaugh, S. (2009). Playing with meaning: normative function and structure in play. In R. Botha & C. Knight (Eds.), *The Prehistory of Language* (pp. 122–141). Oxford University Press.

Roth, A. E., & Erev, I. (1995). Learning in Extensive-Form Games: Experimental Data and Simple Dynamic Models in the Intermediate Term. *Games and Economic Behavior*, *8*, 164–212.

Santos, F. C., Pacheco, J. M., & Lenaerts, T. (2006). Cooperation Prevails When Individuals Adjust Their Social Ties. *PLoS Computational Biology*, *2*(10), 1284–1291.

Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in Cognitive Sciences*, *14*(9), 411–417.

Scott-Phillips, T. C., Kirby, S., & Ritchie, G. R. S. (2009). Signalling signalhood and the emergence of communication. *Cognition*, *113*, 226–233.

Skyrms, B. (2008). Presidential Address: Signals. *Philosophy of Science*, *75*(5), 489–500.

Skyrms, B. (2010). *Signals: Evolution, Learning, and Information.* Oxford University Press.

Smith, A. D. M. (2001). Establishing Communication Systems without Explicit Meaning Transmission. *Advances in Artificial Life*, *2159*, 381 – 390.

Smith, A. D. M. (2005). The Inferential Transmission of Language. *Adaptive Behavior*, *13*(4), 311–324.

Smith, K. (2002). Natural Selection and Cultural Selection in the Evolution of Communication. *Adaptive Behavior*, *10*(1), 25–45.

Smith, K., Kirby, S., & Brighton, H. (2003). Iterated Learning: A Framework for the Emergence of Language. *Artificial life*, *9*(4), 371–386.

Steels, L. (1996). Emergent Adaptive Lexicons. In P. Maes, M. Mataric, J.-A. Meyer, J. Pollack, & S. W. Wilson (Eds.), *SAB96.* MIT Press.

Steels, L., & Kaplan, F. (2002). Bootstrapping Grounded Word Semantics. In T. Briscoe (Ed.), *Linguistic Evolution through Language Acquisition: Formal and Computational Models* (p. 53-74). Cambridge University Press.

Steels, L., & Loetzsch, M. (2008). Perspective Alignment in Spatial Language. In K. R. Coventry, T. Tenbrink, & J. A. Bateman (Eds.), *Spatial Language and Dialogue* (p. 70-88). Oxford University Press.

Steels, L., & Spranger, M. (2009). How Experience of the Body Shapes Language about Space. In C. Boutilier (Ed.), *Ijcai'09: Proceedings of the 21st international joint conference on artifical intelligence* (pp. 14–19). Morgan Kaufmann.

Steels, L., & Vogt, P. (1997). Grounding adaptive language games in robotic agents. In I. Harvey & P. Husbands (Eds.), *ECAL97* (p. 474-482). MIT Press.

Tamariz, M., Gong, T., & Jäger, G. (2011). Investigating the effects of prestige on the diffusion of linguistic variants. In L. Carlson, C. Hölscher, & T. Shipley (Eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society* (p. in press).

Tomasello, M. (2003). *Constructing a Language: A Usage-Based Theory of Language Acquisition.* Harvard University Press.

Vogel, C. (2010). Group Cohesion, Cooperation and Synchrony in a Social Model of Language Evolution. In A. Esposito, N. Campbell, C. Vogel, A. Hussain, & A. Nijholt (Eds.), *Development of Multimodal Interfaces: Active Listening and Synchrony* (pp. 16–32). Springer Berlin Heidelberg.

Vogel, C., & Woods, J. (2006). A Platform for Simulating Language Evolution. In M. Bramer, F. Coenen, & A. Tuson (Eds.), *Research and Development in Intelligent Systems XXIII* (pp. 360–373). London: Springer.

Vogt, P. (2005). On the Acquisition and Evolution of Compositional Languages: Sparse Input and the Productive Creativity of Children. *Adaptive Behavior*, *13*(4), 325–346.

Vogt, P., & de Boer, B. (2010). Language Evolution: Computer Models for Empirical Data. *Adaptive Behavior*, *18*(1), 5–11.

Wellens, P., Loetzsch, M., & Steels, L. (2008). Flexible Word Meaning in Embodied Agents. *Connection Science*, *20*(2), 173–191.

Worgan, S. F., & Damper, R. I. (2008). Removing mind-reading from the iterated learning model. In A. D. M. Smith, K. Smith, & R. Ferrer-i Cancho (Eds.), *Proceedings of the 7th International Conference on the Evolution of Language* (pp. 378–385). World Scientific.