Computational Models of the Emergence of Language

Pippin Barr

February 11, 2002

Abstract

An introductory look at the work of computational modellers on the origins of syntax and semantics. We begin by examining current theories on the emergence of language in general and then questioning the utility of computational models. The large part of the work is then spent explaining and analysing the models of various syntax researchers, as well as introducing a new look at language models generally. A final section is provided on the work of Luc Steels on semantics modelling after which we conclude.

Contents

1 Introduction

This essay takes an introductory look at the use of computational modelling to understand the emergence of language. The topic is large, and so the focus will be on the analysis of just a few models and some thoughts for future directions.

First we will look at the two components of our topic in isolation. In section 2 we will review current theories of language evolution, while in section 3 we will focus on the utility of computational modelling in general. Following this, in section 4, we outline the work performed by four syntax emergence researchers and consider some of the common factors. A significant portion of the essay is then spent in section 5 analysing in some depth what it is that the modellers have found. Of particular interest is a detailed outline of a way of viewing these models in section 5.7 which leads to sketching out the basics of a new model. Finally, in section 6, we will touch on at the work of Luc Steels on the modelling of semantics.

2 The Evolution of Language

2.1 Introduction

In this section of the essay we will very briefly look at some current theories concerning the evolution of language as well as address what language actually is. We will not discuss computational models here, but instead present background knowledge so we can analyse what researchers into computational modelling of language have achieved.

2.2 How Did Language Start?

One important divide in the study of the origins of language is between catastrophists and gradualists. Catastrophists claim that some major event in our evolutionary history suddenly made language possible. The gradualists, on the other hand, believe that language came about slowly, through a series of adaptations, in the more traditional mode of evolution.

Another dichotomy is singularism versus pluralism. Singularists claim that natural selection was the only driving force, while pluralists assert that there were a variety of factors and that language emergence was not driven solely by evolution. In particular, they dispute that language was specifically an adaptation, instead claiming that large parts of it probably came about by accident.

Many factors in early hominid history have been suggested as the starting point of language. The usual picture is of animal grunts and cries becoming language over time as they were associated with objects in the world. Another explanation suggests that grunts and cries might have been used as part of a "vocal grooming" system: a way of maintaining relationships within the group as the groups got larger and relationships became more complex to deal with. Over time this might have become more and more expressive, eventually developing into a proto-language. Perhaps language is an exaptation as Bickerton suggests, arising as a side-effect of a social calculus in early hominids [12, p.267].

An interesting perspective on the use of language draws on how intertwined it was with the development of what might be called a language of thought: the capability for complex concepts. Some researchers claim that language developed as a result of this, while others believe the two came about together, developing over time.

In all, then, the origins of language remain mysterious. It is clear that language use is an adaptive trait, but that tells us more about why it has remained than why it began.

2.3 What Does Language Involve?

It is important to have an idea about what counts as language, and why it is useful, so that we can critically examine the models later in the essay. Language as a whole can be considered the conjunction of syntax and semantics.

2.3.1 Syntax

The use of various structural properties of a language to indicate more information than is contained in just the symbols or sounds involved is a good way to think of syntax. A classic example of this is word order. "Jim loves Mary" and "Mary loves Jim" contain exactly the same symbols, but because of syntactic properties they mean very different things: the word order indicates the subject and object of the verb. At a broad level this is the function of syntax.

Syntax is a useful quality in a language because it gives the ability to put more information into communication that would be possible otherwise. Without syntax, communication is just a set of symbols on what is basically a serial channel. This makes the number of possible symbols the limiting factor in the process. Once we can put some structure on top of this, we can map far more complex information into the channel using the *same symbol set*. This is the power of syntax.

2.3.2 Semantics

Semantics refers to the meanings that the symbols in a language have and, as such, is really the reason language works at all. Given that the point of language is to communicate information, the symbols must have meaning or nothing is transmitted.

How semantics come about is a very interesting question, and we have looked a little at this in section 2.2. The key idea is that the meanings of words are generally *grounded* in real world phenomena. While there is an exception in abstract concepts, in general, words are used to directly refer to things or actions in the world. Thus, the key duty of semantics is to indicate *what* we are talking about. Later in this essay we will briefly look at some computational research into grounded semantics.

2.4 Important Questions

The discussion in this section raises interesting questions. While not all of these are answered in the following essay, it is worthwhile listing them as a focus for further thought.

- *How can language change?* In particular, how do syntax and semantics change. In some ways it seems obviously maladaptive to use a language different to that of your peers, and yet that seems the only way for language change to occur.

- *How does syntax emerge?* Syntax creates structure in our communicative systems. It seems highly unlikely that it was present the moment we began communicating with vocal noises, so how did syntax arise?

- *How does semantics emerge?* It seems unlikely that the noises we made with our vocal cords *always* possessed meaning. How is it then that sounds (and, indeed, signs in general) came to possess meaning?

- *Is language intrinsically a representational or a social tool?* Did language come about because we needed a way to easily represent concepts and their combinations in our heads, or did it come about as a means to better social coordination? Or was it both?

3 Computational Models

Before we look at the techniques used to examine the evolution of syntax and semantics, we ought to briefly question the overarching technique being used. Computational models are certainly an interesting avenue for exploring various concepts, but are they appropriate here?

The chief task when implementing computational modelling is to make justifiable assumptions about the experimental domain. If we do not model the domain within a reasonable degree of accuracy, it becomes questionable as to whether our model's results will reflect anything on the real world. Essentially, we need the ability to pick out those aspects of the world that are relevant to our model, while whittling away any superfluous details:

> The mark of a good simulation is that it separates the essential from the incidental, cutting through what is deemed irrelevant detail to get at the heart of a problem. This involves making instinctual judgements about which details are crucial and which can be ignored [8, p.244]

Making such decisions shows us why we might regard computational modelling as something of an art form. Clearly, though, we can never create a model that is completely true to the real world in all the right ways. As Byrne notes in [2], "simulations cannot and must not be regarded as equivalent in a homologous sense to experiment". Given how chaotic the real world is, it is presumptuous to think we could ever pick out all and only the relevant considerations to include in our model. That being the case, the best we can hope for is that our models indicate *ideas* and *salient possibilities* about the target domain, rather than simulate exactly what would or did happen.

An important consideration when it comes to modelling is whether we are in the business of prediction or theory building. We can either set up a model we believe reflects the world, and see what happens; or we can set up some parameters we think are interesting, see what they generate, and from this note that these parameters have some connection with the result. In the case of language modelling, the latter is generally the case. As Batali says:

> Models are thus valuable to the degree that they explicitly illustrate the consequences of the set of assumptions they embody. This may be even more important than whether those assumptions are correct. [1, p.3]

Thus, computational modelling of language generally seeks to establish what sorts of conditions might lead to language, rather than attempting to evolve language in a realistic scenario.

A final consideration about computational models concerns the issue of inevitability. It is clearly the case that, if one could say with confidence what the results would be without running the model, then the model is essentially pointless. According to Gilbert, "emergent behaviour is that which cannot be predicted from knowledge of the properties of agents, except as a result of simulation" [7, p.150]. Thus we must be on the lookout for models that give their results "trivially" from the initial properties of the model.

Computational models have been clearly demonstrated as useful analytical tools in various fields, and they are likely to have the same impact in the field of language evolution. Nevertheless, we have seen that the process is difficult, and we must be on our guard to watch for faulty assumptions, and trivial results.

4 Syntax Evolution

4.1 Introduction

This section of the essay will discuss the techniques used by several people to model the evolution of syntax in populations. Once we have broadly outlined what each of the models involves, we will go on to examine some of the factors they have in common. The models to be examined are those of Simon Kirby, John Batali, an honours project by Stephen Chong for Victoria University, and, finally, the work of Zuidema & Hogeweg.

4.1.1 Simon Kirby

Simon Kirby's aim is to "explain the origins of syntax in language by looking at general properties of the transmission of learned behaviour." [10, p.1] This involves simulating a population of induction algorithms which teach and learn from each other in order to develop a shared language system.

In Kirby's model each agent consists of an induction algorithm and a grammar. The environment is either just two agents, with one teaching and one learning [10], or a ring of ten agents, where each agent teaches and learns from its immediate neighbours [11].

The process of learning is very simple. The teacher agent is given some number of randomly selected meanings which it is to express to the learner. It creates strings (or utterances) for each of these based on its grammar and then presents them to the learner. The learner then creates new rules based on what it heard, and what the meaning was meant to be. The induction algorithm is applied to the rules in the grammar to find generalisations, substitutions, rules that could be joined and so forth. If the teacher has no rule to create an utterance for a meaning, it can invent a random string of characters, such as "abababbc" which it then adds to its grammar as representing that meaning.

Using just this simple set-up Kirby observed the emergence of syntactic structure in the population. The agents begin with empty grammars and so teachers have to randomly invent utterances. This means that strings for meanings have no relation to each other at all. Over time, though, a distinctive structure emerges, with tokens of utterances being repeated, and an overall "word"-order being displayed. Thus we see a common set of characters, say "aba" which means "John" across all utterances, for example. We also see a consistent verb-object-subject word ordering.

4.1.2 John Batali

In John Batali's model an agent is represented by a recurrent neural network. These networks have a feedback loop from their hidden layer into a set of "context" nodes which allow the network to track some temporal properties on the

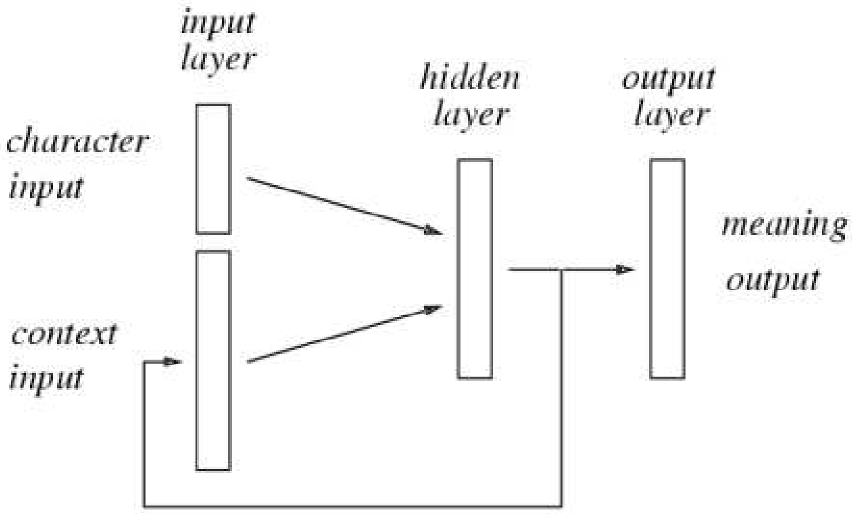

Figure 1: Recurrent Neural Networks in Batali's model [1, p.5]

input. As we see in figure 1 the network has inputs for the characters used to represent meanings and outputs which give the meaning of the entered string for the agent. There is one input for each character used, so a string is entered one character at a time, with the meaning output being final once the last character has been processed.

Meanings are represented by the agents as vectors of real numbers between zero and one, while a "true" meaning is strictly a binary vector. To create an utterance the agent successively tries out characters and builds up a string that moves its meaning outputs as close as possible to the meaning vector to be expressed. To understand an utterance from another agent, an agent feeds the string into its character inputs and takes the meaning as the vector generated at the outputs.

In the population of networks an agent is randomly selected as the learner and then ten random agents are randomly selected to be teachers. The teachers train the learner network on a set of meanings using back-propagation to modify the connection weights. This is done in the traditional manner of calculating the error of the learner's output compared to the correct meaning vector and then adjusting the weights appropriately with the back-propagation algorithm. As the learners are neural networks, back-propagation seems the most obvious choice for the learning algorithm.

Batali obtained extremely similar results to Kirby, with a clear syntactic structure emerging. In Batali's case this structure was interpreted as a root plus a suffix being used to express the simple subject and predicate meanings.

4.1.3 Stephen Chong

In Stephen Chong's model described in [6] an agent simply consists of a grammar which is modified using a genetic algorithm while meanings are represented as directed chain graphs. Each rule in an agent's grammar is simply a graph pattern and a string of characters.

Change is achieved via the genetic algorithm. The fitness of the population is assessed and the top percentage of the population is copied across to the next generation. The remaining places in the population space are filled by offspring created randomly from the available parents. Offspring are created asexually by mutation of the parent's grammar. Mutation consists of random deletion, addition and alteration of rules in the grammar. A similar process of deletion, addition or alteration is applied further down to the graph patterns and the character strings of the rules, if the rule has been selected for alteration.

Chong uses several measures to examine fitness through the language game played. When a speaker is unable to produce any utterance for a meaning, they are penalised. When a speaker produces an utterance that the listener does not understand, the speaker is rewarded based on brevity, while the listener is penalised. When a successful exchange takes place both are rewarded based on the brevity of the utterances and how well the utterance was understood.

The results of running this model over 20,000 generations for different concept sets were quite promising. In particular, when Chong used a concept set equivalent to Batali's meanings, he achieved very similar results, with a clear syntactic structure emerging.

4.1.4 Zuidema & Hogeweg

Zuidema and Hogeweg's model in [22] is very similar to Chong's. Agents consist solely of context-free grammars and again mutation is the sole reproduction operator. Fitness is determined by playing a language game in which everyone speaks once and tries to understand everyone else. Scoring is based on success in speaking, understanding and being understood.

Both composition and recursion emerged over time in the model. Interestingly, Zuidema and Hogeweg discovered that sometimes, even with all the same parameter settings one could end up with *either* composition or recursion on different runs. We will discuss this further below.

4.2 Common Features of the Models

4.2.1 Compositionality and Recursion

An important point to explore is what these researchers think was a successful result. Their criteria tend to be similar and straightforward, but fascinating nonetheless: the emergence of compositionality and recursion.

What they are looking for, then, is for structure to emerge from the random chaos that the agents begin with. Utterances would not be just a jumble of characters, but composed of tokens with specific meanings in the utterance. As opposed to the random string "a" meaning "John loves Mary" we might see "aac bba ca" where each of those tokens stands in for one of the words in the phrase. We would be able to tell this was the case by looking for consistency across the various utterances created. If the agent produces "aac bba ca" for "John loves Mary" and "aac ca bba" for "Mary loves John" we can begin to predict that "aac" means "loves", for example. With further strings we could find which of "ca" and "bba" means "Mary", and so forth. In this way we can see that the utterance would be *composed* of tokens, hence: compositional grammar.

To get this result from initially random utterances seems, at least at first, impressive. Recursion adds the ability to string together long sentences such as "John believes that Mary believes that John loves Mary" in a consistent manner.

4.2.2 Speaking and Meanings

In addition to the ways that modellers handle transmission of language, there are major similarities when it comes to how the agents in their simulations "communicate". In all of these models the meanings that the agents can express are all *predetermined*. That is, a set of meanings exist, and those are all the things that can be meant in the agents' universe.

All models also have in common the way that agents express themselves. In each case the agent generates a string of characters which, for it, represents the meaning under consideration.

Another common feature is the nature of the meanings expressible. In all cases these meanings are compositional in that they can be split into parts. In some cases, such as Simon Kirby's model, they are compositional on subject-object-verb distinctions, such as "John loves Mary"; in others the composition comes in pronoun-verb distinctions, such as "I am happy". Note that this compositionality is in the *meanings* and is not explicitly built into the languages the agents learn. Additionally, some models have the possibility of recursive meanings which involve verbs such as "believes that", which can be strung together indefinitely.

In all but one of the models the agents represent their languages as some kind of a grammar. Kirby uses definite-clause grammars, Chong uses directed graphs and Zuidema and Hogeweg use simple context-free grammars. Note that all of these types of grammars innately possess the ability to express compositionality and recursion, though it is not compulsory for them to do so. John Batali is the only modeller not to represent languages as grammars, and this is because he is using neural networks for his agents.

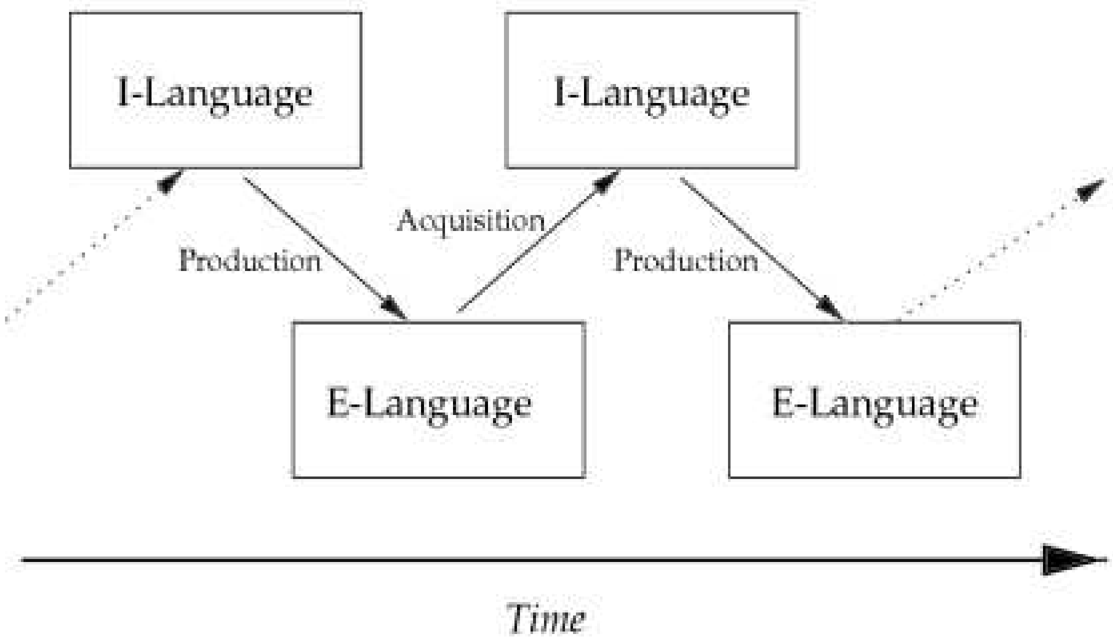

Figure 2: E- and I-language diagram from Kirby [11, p.15]

4.3 Transmission

The four models differ as to how to model the actual transmission of language in a population. Two of the researchers examine cultural or social transmission, essentially through learning; the other two explore the idea of genetic transmission of languages.

4.3.1 Social Transmission

Both Kirby and Batali seek to model social transmission of language. What this means is that they set up populations of *learning* agents who communicate with each other, and learn language in that way. Kirby suggests thinking of this in terms of Chomsky's concepts of E(xternal)- and I(nternal)-language. The E-language is what is uttered, while the I-language is the syntactic and semantic rules in the head that make up the conceptual version of a language. The goal of social transmission, then, is for a learner to transform the E-languages it hears into its own I-language, hopefully consistent enough with those of its peers to successfully communicate. We can see the general process illustrated graphically in figure 2

The scenario in both Batali and Kirby's work is to have a population of learners. Simply speaking, at some given time there will be a learning agent and a teaching agent. The teacher utters some number of strings to express some randomly chosen meanings. The learner then tries to learn these strings in its I-language. In Kirby's case this consists of induction, while in Batali's it is back-propagation.

4.3.2 Genetic Transmission

Stephen Chong, and Zuidema and Hogeweg, seek to model language evolution more traditionally via genetics. Their models essentially treat language as the expression of genes. Thus they use the mutation of I-languages (sets of rules) to create agents with new grammars.

A parent is one of the fitter members of the population which is then mutated to create a child. The theory is that, with only this simple process, syntax can emerge in a population of grammars.

Naturally, when we start talking about reproduction, there needs to be some way to choose who reproduces and who does not. Generally, fitness is measured by having the agents play some sort of language game with each other, with success in the game being the measure of fitness. Selection consists of taking some portion of the most fit agents into the next generation and allowing them to have children. As we will see in the analysis, the chief difficulties for this kind of model involves the plausibility of the fitness measures.

5 Analysis

5.1 Introduction

In this section of the essay we will critically examine the models of syntactic emergence described in the previous section. This will involve a more detailed look at how the researchers put their models together, along with some analysis of their results, and their interpretations of those results. Additionally, we will consider more generally the aims of such modelling, and think about the major issues involved.

5.2 Simon Kirby

5.2.1 Induction

Probably the most important single feature of Kirby's model is the use of induction as the learning mechanism. Kirby claims induction

> ... does model in a simple way the dual processes of rote learning and search for generalisation that should be the core of any theory of language acquisition...
> [10, sec. 3.1]

While Kirby believes that induction is an accurate portrayal of real-world learning, the central role of induction has been questioned by advocates of Chomsky's Universal Grammar:

> It is not at all clear what role induction plays. Induction does not enable a child to determine what a well-formed sentence is; nor does it explain how children "learn" the meanings of even the simplest words. Children do not have sufficient evidence to induce the meaning of *house*, *book*, or *city*, or of more complex expressions...
> [13, p.63]

The general thinking here is that the real world acquisition of language seems *far* too complex for such a simple method as induction to take care of an agent's learning needs. In particular, it is widely noted that neither the amount nor the structure of information presented to a learner is enough for that learner to develop a good concept of the underlying workings of the language. That being the case, it is hard to believe that syntax came about through inductive processes only.

Nonetheless, what Kirby *has* shown is that induction was sufficient, in this extremely simple case, for agents to develop and discover syntactic structure. The problem is that Kirby's conditions may be so far removed from anything realistic, that it may not show as much about the emergence of syntax as we would like.

5.2.2 The Agents

A claim that Kirby makes about his agents is that they are not constrained to learn a syntactic language at all. This means that Kirby believes that the combination of an inductive learning process and the use of definite-clause grammars does *not* inexorably lead to the creation of syntax. While this might be true, we will withhold judgement until further discussion of Kirby's results. A strong intuition is that his results actually *were* inevitable, but we will have to think carefully before we make such a pronouncement.

Another particularly noteworthy feature of the agents is the invention algorithm. This allows the agent to say something, even though it might not have a grammatical structure in its I-language to create the appropriate utterance. The value of Kirby's work here is that the agents use their internal grammar *as far as they can* before putting in invented characters. This simulates the idea of a partial understanding of the language, and is one of the nicer features of the model.

A final point to be made is that all of Kirby's agents are "genetically identical". That is, every agent starts exactly the same as every other agent: with an empty grammar, and identical induction and invention algorithms. Thus the only distinguishing feature of agents becomes the grammars they have learned. This is effectively identical to Avital and Jablonka's notion of "Tarbutnicks", which are discussed in [3]. The idea is, like Kirby's, that evolution can take place *without* genetic variation, solely via the transmission of learned behaviour. Indeed this can be thought of as a kind of Lamarckian inheritance: a quality gained by the parent during their life is passed on to the child not genetically, but socially.

5.2.3 The Environment

The environment in Kirby's models is intentionally kept exceedingly simple. For example there is no social interaction between agents, no sexual reproduction, no real death other than random replacement and no sense of family groups or such things. Instead we have the situation of agents living in a vacuum and existing only to say random phrases to each other. We must ask: is this so far removed from anything remotely realistic as to be undesirable? As we discussed earlier, the aim of these models is to discover what sorts of assumptions might lead to the emergence of syntax. Since Kirby has elicited this emergence, we should look at what assumptions he *does* make, rather than question whether other assumptions might have been better. That is not to say, however, that it would not be interesting to work some of those more realistic aspects of life into simulations, and we will discuss this pointin the last section of the essay.

One aspect of the model that is highly unrealistic is the lack of semantic confusion. All agents immediately know exactly what the other agents "meant" by a particular utterance, even if that string is not anywhere in their grammar.

The idea that every agent always understands what another means is so far from reality that it seems it must hinder the model. We will see when we discuss Batali, though, that perhaps this is an assumption we need not concern ourselves with, as Batali *does* have the possibility of misunderstandings, and still gets strikingly similar results.

A final aspect of the learning environment that we might wonder about, is the fact that any given learner only ever has one or two agents who speak to it. As we know, in the real world we are exposed to huge amounts of information from vastly diverse sources. We cannot really consider Kirby's condition of having only one or two "teachers" at a given time a useful simplification. That is because, as we mentioned earlier, a key argument against induction being the key ingredient in the development of syntax and language in general is the huge diversity of input. The claim is that the input to a learner is actually *too* diverse for induction to be effective. In that case, simplifying to just one or two inputs can be seen as completely ignoring this aspect of language learning.

One reply that could be offered to these criticisms is to cast Kirby's induction algorithm as something like Chomsky's Language Acquisition Device. Given that the algorithm is clearly for towards discovering patterns in language, and thus towards creating syntax over time, surely this bias might combat the diversity of inputs in the real-world environment. This would give some support to the use of specialised induction in a language learning environment, at least. Nevertheless, we cannot avoid the point that Kirby's environmental set-up is highly unrealistic, as it is not nearly complex enough to even approximate everyday conditions in the world.

5.2.4 Semantics

The fact that semantics is effectively ignored is an issue with Kirby's model. In the model, on hearing an utterance, the learner immediately knows precisely what the speaker means every single time. This is nothing like the way the real world works, where we frequently misunderstand what others are trying to tell us. Perhaps adding in misunderstanding, though, would just slow down the development of syntax, rather than create any interesting side-effects. As we will see in section 5.3, John Batali has a model which can involve misunderstanding, and achieves strikingly similar results to Kirby.

The set-up of the model also reflects a lack of what Luc Steels calls "grounded word semantics". That is, the meanings that the agents are using are not related in any way to their environments; they just exist. Angelo Cangelosi believes this is a deficiency in most syntax emergence models. Clearly this is because the agents lack any sensory functioning other than speaking and hearing, they have nothing to ground semantics on. Is this something we should be concerned about? Cangelosi thinks so and emphasises "[t]he importance of providing a model with a mechanism for grounding symbols"[4, p.100].

It is an interesting question to ask how much difference it would make to also model the grounding of the symbols in the environment. It would complicate the model to a large degree, but would interesting results follow? Cangelosi believes the answer is yes, and has performed various experiments of his own to show how combining grounded semantics and language evolution can yield intriguing discoveries. In particular, Cangelosi notes that "unexpected sets of semantic categories can emerge"[4, p.100]. While this is certainly an interesting effect of including a grounded semantics in models, it is still questionable as to whether it is really a *problem* for pure syntax modellers.

Cangelosi also observes that, by not allowing grounding, modellers can miss out on simulating the very early stages of language development. By assuming an already fully fledged semantic system, surely the development of syntax is at the very least contrived. This raises the difficult question of just how intertwined syntax and semantics really are. Cangelosi's experiments suggest there are some relations, but not to such a degree that we should disregard the results of Kirby and others, who ignore semantic grounding.

5.2.5 Natural Selection

One point of emphasis for Simon Kirby in his model is the absence of any natural selection. There is no reward system for successful communication, and hence no concept of fitness for the agents at all. He did this to emphasise how language can come about purely through social interactions and not necessarily as an adaptation.

In this aim it is fairly clear that Kirby was successful. What is most interesting about eliminating natural selection from the agents in the model, is that it leads us towards considering the possibility of natural selection of the actual *languages* that they use. This is an extremely valuable area to explore, and sheds some light on what is going on behind the scenes. We will look in some detail at language as a complex adaptive system at the end of this section, including some of Kirby's own observations on the matter.

5.2.6 The Results

Kirby's model does indeed result in the emergence of syntactic language. The agents in the model all eventually use a compositional syntax in order to express the various concepts they are able to express. It is also clear that this result came about only through "social transmission" as Kirby intended, i.e. through the interaction of the induction algorithm and the invention algorithm.

One of the most important questions we should ask about Kirby's model is whether the results were inevitable. If so, then there was no point in doing the experiment. It does seem that Kirby's induction algorithm is specifically geared towards creating a syntactic grammar. The process it follows is essentially to look for patterns in the current grammar, and simplify them into smaller rules. This process is exactly what you would use to create a syntactic grammar, which is, after all, just the utilisation of patterns to simplify communication. That being the case, it becomes difficult not to claim that, given the environment Kirby set up, syntax would inevitably emerge. Note, however, that if this is the case, Kirby has at least shown something about the assumptions made in the experiment: namely, that they are sufficient to lead to syntactic development. That in itself is an interesting result, and so, whether syntax was inevitable or not, Kirby's model has useful consequences.

A very interesting aspect of Kirby's results are the implications concerning the idea of catastrophic change. In figure 3 we can see a revealing portrait of the development of the grammars. As we can see during stage I the number of meanings expressible is very low, as is the size of grammar. This means that the agents essentially have few words, and thus can express few concepts. There is a sudden transition to stage II, where we now see a much larger number of meanings expressible, but at the same time much larger grammars. This suggests that the agents use many words to express many concepts, which is impressive, but inefficient. Finally the agents move to stage III, where it is clear that a syntactic grammar has developed: all the meanings can be expressed, but the grammar has actually *reduced* in size.

The transitions from stage I to stage II, and from stage II to stage III are extremely sharp. This suggests that there was some sort of catastrophic change in the way the the agents' language worked. In the first case it caused their grammars and meanings to explode, and in the second case drastically reduced their grammar sizes while maintaining the meanings expressible. This would seem to lend at least some support to catastrophist hypotheses of language change.

Finally, Kirby gives an interesting look at his results in [10] by examining them in terms of what he calls the "learning bottleneck" between the E-languages and the I-languages. What Kirby claims is that non-compositional languages simply cannot survive over time. This is because to learn them you must be exposed to every rule in the grammar, and when the number of presented utterances is fewer than the meanings that can be expressed, no learner will ever see all possible rules. That being the case, such grammars simply cannot be passed through the bottleneck. Compositionality changes this, Kirby says, because now one can generalise to rules by seeing only a couple of applications of them. In this way a learner *can* learn the entire grammar through only limited meaning presentations because of the structure present. This leads us to conclude that syntactic communication is the *only* system that could survive the learning bottleneck. This sounds interestingly like adaptation.

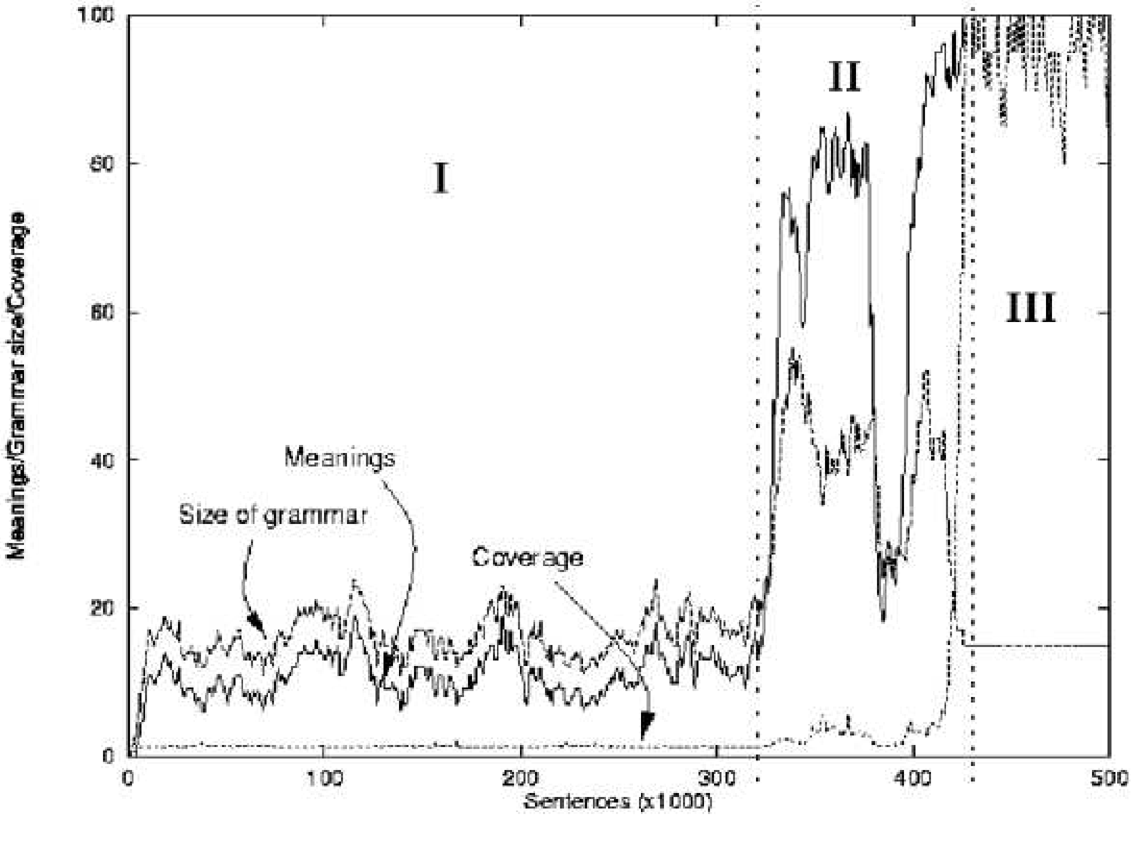

Figure 3: Graph from [11, p.9]

5.3 John Batali

5.3.1 The Agents - Recurrent Neural Networks

Batali's use of neural networks, as opposed to induction algorithms, for agents initially seems an extremely different approach to the issue of syntax emergence. The two techniques do, however, have one major feature in common: they are both effectively pattern-finding techniques. In that respect, then, it is not surprising that Kirby and Batali find the same results. The whole process is effectively the evolution of patterns.

One of the nicer features of Batali's model is that the agents use their own network to gauge the best phrase to elicit the correct response in other agents. Thus, if they wish to express some particular meaning, they work out the string that would be most likely to express that meaning to *themselves*. This is reminiscent of the way in which human beings work, and has an aspect of folk psychology to it. Most of us make the assumption that others are fairly similar to ourselves, and thus will react in similar ways. Having this assumption embodied in the model is a good way to have some real-world accuracy, and may well have an important impact on the emergence of a proper syntax, and particularly a *shared* syntax. In effect, Kirby's model also draws on this sort of folk psychology, as agents simply use their *own* grammar to express meanings and hope that other agents will understand. Batali's version is the more interesting because it involves the agents actually presenting the meaning to themselves and then finding out what would make them understand it.

Once again, we should note that there is absolutely no semantic grounding involved as the agents have no perceptive ability at all beyond "hearing" strings from other agents. The same issues that were discussed with reference to Kirby apply here also. While the meanings are now vectors of real numbers, this is really just so that Batali can use them with neural networks and there is no difference between these and Kirby's meanings. The meanings are compositional in a similar way, and this is again what allows the agents to create a syntactic grammar.

5.3.2 The Environment

An improvement over Kirby's model comes in the set-up of the learning environment for agents. The fact that agents now learn from a far more diverse collection of "teachers" is more similar to the real world. Additionally, the teachers are always random, rather than Kirby's single teacher, or teaching neighbours.

Despite that observation, we must keep in mind that these researchers are attempting to go backwards from the result to examine the conditions. Thus, as Kirby's model did obtain the same results, perhaps his simplification of the learning environment was a reasonable one. Nonetheless, it is reassuring to see a more realistic environment yielding those results as it gives better support to the result being applicable in the real world.

An interesting feature of the environment in Batali's model is the lack of any form of replacement of agents. The entire model consists of random agents negotiating until a syntax is created, rather than new agents being introduced into the population and educated by more experienced agents.

A final aspect of the environment of the agents concerns the actual game that they are playing. Specifically, the accuracy required to be considered to understand another agent's utterance is surprisingly low. In fact, being within 0.5 of the correct value for each number in the vector is considered successful. This is a very rough value but, as it transpires, this does work out well: in the end, the interpretations are very close to the intended meanings, often exactly right or within about 0.1 of the correct value for each position in the vector.

5.3.3 The Extra Experiment

Perhaps the most exciting thing about Batali's work is the extra experiment he performed on his agents. For this, he ran the experiment as usual, but with certain meanings missing from the pool available, such that the agents never developed the ability to express those meanings directly. Batali then has one agent attempt to express one of the unseen meanings to another agent, who then tries to interpret it correctly. The impressive result is that this process was very successful.

What this suggests is that there is some real syntax in the agents' languages, rather than a rote-learning of forms that appear syntactic. The ability to apply the syntax to unseen forms of words means that the agents really do have a system for creating expressions of meanings. The tokens of the words they use really have those meanings. It is not a coincidence.

This is a valuable experiment because it reinforces the idea that a real syntax has developed. A similar development is actually possible, but perhaps less impressive, in Kirby's model. Here we could make sure that, say, the phrase "John loves Mary" was never presented to any agents. When the agents had been trained, this meaning could be presented to them, and because they would have encountered all the *parts* of that phrase, they *would* be able to generate an utterance to express it. This is not really necessary, however, because we can *see* that the agents have a genuine syntax by examining their grammars. One of the chief disadvantages of Batali's model, then, is our inability to actually see into the "minds" of the agents. This is a traditional difficulty with neural networks in general.

5.3.4 The Results

An important statistic that Batali examines in his paper is how well *shared* the communication system of the agents was. In his case, it was impressive, with all agents producing the same sequence for the large part of the meanings, and most agents sharing the same sequence for the other meanings. Batali is the *only* researcher to record this statistic, which is curious as the shared nature of language is crucial to its very being. Perhaps it is just not possible for only one agent to develop a syntactic grammar, though, which would suggest that syntax would spread through a population as soon as it appeared; or, rather, syntax develops in populations *as a whole*, rather than in individuals.

Batali notes that the systems of syntax developed by the agents are by no means *perfectly* regular. In fact, a fair amount of variation is involved, although there can be no doubt that there is systematicity there. Batali is quick with a disclaimer:

> The existence of partial regularities in the system can be used as a convenient way to group the sequences, but may be entirely artifactual, and certainly does not entail that any sort of analysis in terms of those regularities is involved in the agents' interpretation of the sequences.
> [1, p.19]

Batali goes on to give a more technical explanation of the phenomenon observed in the agents' utterances. A string of characters is a way to move the activation values of an agent's network towards the appropriate meaning vector. Clearly, there must be a unique string for each meaning, clearly. In order for the agents to communicate, the trajectory followed through the networks to get to the meaning vector must be similar, and thus the character strings that induce these trajectories must be similar. As Batali says, "[i]f they were to diverge substantially at some point in the sequence, there would be little chance that the sequence of remaining characters would bring them back together"[1, p.19]. This, then, is Batali's more precise explanation for the results of the experiment. In this case, it makes sense for there to be syntactic regularities, because syntactically similar utterances will be "designed" to push the activation values to compositionally similar meaning vectors. Batali also notes that this explanation also explains why the agents tend to have the predicate part first in their utterances: they account for more of the meaning vector and thus are better to choose first as they have more impact.

As noted in 5.3.3, one of the more interesting things about Batali's experiment is the absence of explicit grammars in the agents. This makes us more confident that Batali has not built in the move to compositional grammars by having it as part of the grammar representation. Perhaps Batali's achievement of the same results as Kirby but without grammars also suggests that Kirby did *not* build in his result with the use of grammar. It was achieved without them.

5.4 Stephen Chong

5.4.1 The Agents

Even more so than in Kirby's experiment, there are no real *agents* to speak of in Chong's model. A more accurate characterisation of this model is of sets of rules attempting to communicate with each other and reproducing asexually via mutation when they are successful. There is decidedly *no* learning at work, only mutation. It is interesting, then, that he obtains very similar results to both Kirby and Batali.

One of the most curious issues in Chong's model is that the agents do not necessarily understand *themselves*. That is to say that an agent might fail in a conversation with itself. This seems bizarre and too far from the folk psychology concept of using one's own response to gauge the response of others to be useful in thinking about the real world. Clearly, this ability to misunderstand oneself is an artifact of the way Chong has set up his model, but it is hard to see why it did not concern him more.

Chong's agents reproduce via asexual mutations. It is initially difficult to think of how this is relevant to modelling language change in the real world. When we address the idea of language as a complex adaptive system, however, it may become more relevant. For now we will note that, so far as this model is used to establish assumptions that could lead to syntax, it does not seem exceedingly useful. Mutation does not look like a good explanation for language arising in humans. That said, perhaps mutation *could* be viewed as a crude generalisation of the learning process: the child obtains something *like* the language of the parents.

5.4.2 The Fitness Landscape

The most important feature of Chong's model comes in his interpretation of fitness for language use. In his case rewards and penalties are based on the brevity of utterances and understanding, both positives. These fitness functions generally seem plausible, except perhaps for the situation where a speaker creates an utterance that the listener does not understand. In this situation the speaker is rewarded based on how brief the utterance is, while the listener is punished based on how much they would have gained if they had understood the utterance. It seems that a more realistic reward function would involve penalising the *speaker* also for a communications failure. While the speaker is not *rewarded* for communication, there is no penalty. The reward for brevity is in place to keep utterance lengths under control.

Another thing to notice about fitness in this model is that it is not *solely* dependent on the performance of the agent. One agent's fitness effectively depends on the performance of everyone that communication is attempted with. This creates the broad interdependence between agents that encourages a shared communication system: they are all in the same boat.

5.4.3 Genetic Transmission

Perhaps the key question to ask about Chong model is just whether it is a good way to model language evolution. Given that we are discussing language *evolution*, explicitly modelling it with genetic algorithms seems an intelligent choice, but how much does it tell us about the actual causes of this evolution? The idea of treating language specifically as a complex adaptive system will be addressed in section 5.7 and we will address the pros and cons there.

5.4.4 The Results

Most interesting in the results of Chong's experiments is that they are effectively the same as Kirby and Batali's. This gives us a *third* way to elicit the emergence of syntax in a population and it is this multitude of ways that is so interesting. The fact that the same results are obtained from three methods of investigation surely says something about the emergence of syntax. This line will be taken further in section 5.6

5.5 Zuidema and Hogeweg

The work of Zuidema and Hogeweg is extremely similar to Stephen Chong's (or vice versa) and so the discussion in this section will be brief. Their agent set-up is more or less the same as Chong's with context-free grammars representing the rules rather than directed graphs, allowing recursion. Once again mutation is the only operator and so asexual reproduction is how new agents are introduced into the population. The fitness function is based on "success in speaking, understanding and (not) being understood."[22] Unfortunately, no further detail is offered in their paper, so we cannot easily assess how appropriate their fitness function is.

5.5.1 The Results

Zuidema and Hogeweg discovered three possible results in their runs of the model. Most common was a quick growth in the number of meaning expressible, ultimately leading to composition and then recursive grammars. The second kind of result was a much slower growth of expressiveness eventually ending in largely compositional grammars, with little recursion. The final result was very little growth in expressiveness at all, and involved almost entirely lexical grammars.

What is most interesting in that under some parameter sets, Zuidema and Hogeweg observed all three results depending only on the seed to the random number generator. This suggests that in some situations the development of syntactic language is extremely sensitive to the initial conditions, which is an interesting discovery. An additional finding was that if models were restarted later in development (with a new random number) a change in the type of grammar developing became increasingly unlike, suggesting that the process becomes more and more set as the generation progress. This lends some sort of support to a catastrophist hypothesis that claims that something is needed to push language over the edge for it to develop. Had this thing not been present, language might never have arisen.

A final result that was interesting is that Zuidema and Hogeweg observed that the larger lexical grammars got, the *less* likely syntactic structure was to emerge. This contradicts the standard idea that grammar developed when there were too many meanings to express via simply lexical means.

5.6 Summary

We have now seen the results of four models of syntax emergence in populations. Two concerned populations of learners, while two focused on a genetic algorithms approach, yet all four got exceedingly similar results. This leads us to question: do these results demonstrate the interesting idea that we can get syntax emergence in many different ways, or are these researchers subtly performing the same experiment in different guises?

Certainly, a case could be made that Batali and Kirby's experiments are fairly similar. Both are using agents whose main feature is *pattern recognition*. That being the case, is it surprising they obtained the same results? Perhaps not.

As regards the genetic algorithm approach to the problem, it does seem too far removed from the social transmission idea to really be the same thing. It is hard to believe that Kirby and Chong are performing the same experiment.

The other big question that we should ask of all these models is, given their set-up, were they just building in the emergence of syntax from the beginning? If so, then their results are fairly uninteresting. We must, however, always be very careful about saying that results are inevitable. It is far easier to say this when the results are already out. Could we have predicted what would happen before the syntax had emerged? In the case of Kirby it seems we could have. What *else* would the inductive procedure on grammars *do*? As discussed in the analysis of Kirby, the simplification rules in the inductive procedure seem obviously geared towards generating a compositional syntax. The behaviours in Batali's model, on the other hand, would be a little more difficult to predict. Nevertheless, as neural networks are traditionally built for pattern recognition, there does seem to be some degree of inevitability in his model also. Perhaps what these models really show, then, is that some form of pattern recognition is necessary for the emergence of syntax. While that does seem an obvious result, it is nonetheless useful to have it demonstrated more conclusively.

That syntax emerges under the genetic transmission models is possibly more interesting. This type of model emphasises the adaptive nature of language specifically and gets the same results as the social transmission models. This suggests that we really can treat language as an adaptive system, and that syntax is an adaptive response to an environment involving patterned meanings.

One other way we could think of these models is as a search through grammar space. We could envisage a landscape of grammars such that agents are moving across it searching for a way to communicate with each other. The agents start either at random positions on the landscape, as with Batali's randomly initialised weights, or at the zero-rule position. The learning mechanism combined with utterances from other agents is what allows the agents to traverse the landscape. A key factor is that in order for there to be genuine communication, the agents must keep roughly together, so that their grammars are similar enough for them to understand each other. We thus get the image of something like a *pack* of agents moving across the landscape.

5.7 Language as Agent - Evolution

5.7.1 Introduction

Probably the most important observation to come out of all this analysis is that we can apparently think of language itself as a complex adaptive system. While that obviously would not explain the absolute beginning of language, it could well explain how language has changed and developed over time.

Looking at language in this way would not immediately seem to fit in well with Chomsky's idea of a language acquisition device, but that is not necessarily so. It is perfectly possible that language as an abstract concept, not in terms of individuals but in terms of populations, could be a complex adaptive system. In that case, the language acquisition device would be an integral part of the adaptive *environment*, perhaps. Or perhaps we could even think of it, in some sense, as an individual language-agent's *body*, the genetic side of language. While Kirby's model suggests languages can change completely through social transmission, that does not rule out some sort of biological device playing a part.

One good reason for thinking language is adaptive has been the way we have seen the language changing based on the types of concepts needed to be expressed. For compositional semantics, we have seen compositional grammars arrive; for recursive semantics we have seen recursive grammars. This is evidence that a language system can change and develop in order to express the kinds of concepts a population needs to express. In this section we will closely look at how we might think of language as an agent in an environment, and develop these ideas further.

5.7.2 Agent

There seem to be two main things that we could cast as the agent: either individual rules in a grammar, or I-languages as a whole. Both have their merits, and choosing between them is difficult.

It is fairly easy to see how individual rules could be replicators. In fact, Simon Kirby discusses this possibility in [11]. Kirby notes that for one exposure of a meaning to a learner only *one* rule that can express the meaning has a chance to be learned. In that case we have a competitive situation between rules that can express that meaning. The key to success, Kirby thinks, is rules that can express more than one possible meaning, because these rules have a higher chance of being presented *multiple times* to a learner over the course of their training. This must be balanced, of course, by the difficulty of learning this more general rule. The balance, then, is between the number of possible meanings expressible by the rule, dictating the likely number of times it will be presented to a learner, and the number of times the rule must be presented in order to be learned. If we can find this balance, then we can easily calculate the fitness of an individual rule or replicator.

On the other hand, perhaps we ought to consider whole I-languages as agents, as they actually represent the thing that affects the fitness of the creature using it. An creature's fitness is not affected by individual rules in its grammar, but by its use of language in general. In that case, perhaps whole I-languages must compete to be used by creatures and transmitted between generations.

Finally, we might consider a hybrid picture of language's agency. Perhaps we could think of the individual rules of a language as being its genes, which fits in with the replicator story. The expression of the genes becomes the language's phenotype, or E-language. Its genotype is clearly the I-language as a whole, made up of its genes, or rules.

A key problem with casting language in this light comes when we ask how to carry out reproduction. In this case there is no obvious way to achieve this. Rules are expressed in the phenotype by being part of the creation of utterances. They are successfully transmitted to another being if that being learns them. A new I-language, or genotype is not really created as a whole in one effort of reproduction, but is rather built up over a series of encounters with the phenotypes of other I-languages. The process never really ends, in fact, so this is a highly dynamic system indeed.

5.7.3 Environment

The environment in which language exists is one of creatures attempting to communicate with each other. Language that is not useful for this task will not survive. Language that is will be passed down the generations, effectively reproducing.

There are several features of the environment to be considered here. The key aspect of the environment is probably the learning mechanism used by the hosts of language. This is what language must be directly adapted to in order to survive. The easier a piece of language is for a creature to learn, the more likely it is that that piece of language will make a copy of itself in that creature's head. An additional consideration is how useful a piece of language is. This ties in directly with how frequently it will be expressed as part of the language's phenotype. It is no good being easy to learn if one is never actually expressed.

We could think of the transmission of rules as a process analogous to that of distributing seeds. Rules that are easy to learn are in seeds that are easy for creatures to digest. Rules that are frequently expressed are abundant in the environment for creatures to absorb. Rules that are extremely useful will be more desirable to creatures, and thus will be eaten more frequently. Rules that are easy to apply are more likely to be used than those that are not, perhaps like seeds that are easy or awkward to place in the environment. Finally, if a rule is seldom or never expressed, it is a seed that is almost never found in the real world. This rule has little or no chance of ever being digested by a creature and thus passed on through the generations. It will probably die with its present container.

The key points to note here are that individual rules must be:

- Easy to learn.
- Easy to apply.
- Useful to the creatures that use language, which leads to them being:
- Frequently expressed.

5.7.4 Language Change

A final possibility that must be accounted for is the birth of *new* rules in I-languages. If the entire process consisted of the transmission of rules from one creature to another then clearly languages would just consist of a set from the pre-existing rules in the world. What really happens is that new rules come into being via the learning mechanism in the creatures as they interpret the structure of the utterances they hear.

Thus, if the creatures learn via induction, new rules are created or born when a new type of rule is derived from an utterance. The learning mechanism, then, is the battleground of rules. Fitness is determined by the ability to make the transmission from E-language, through the learning mechanism, and into the I-language of a new creature. As we have seen, syntactic structure is an excellent way of achieving a high fitness.

5.7.5 Thoughts

A key factor that is missing in the models discussed previously, which we can now see, is the interaction between language and creature. It is an important part of the process that language *aids* the creature in some task, as that is the most positive way to be expressed. In the models we have seen the creatures have no task except to speak to each other and learn language. It would therefore be extremely interesting to see what difference an external task would make. Angelo Cangelosi has done some work on this particular aspect of the problem in [4].

Looking at language in this way also serves to illustrate why it is that syntax seems to inevitably emerge. Any time we create an adaptive situation for language it will behave in the manner just described. In the case of the models we have examined, this is what has happened. Syntactic grammar is the most efficient way to encode the information needed to be expressed and so it is not surprising that this is what eventually evolved. The best replicators are those that express the underlying *structure* of the language, because those are the ones that are likely to be most frequently expressed by the creatures. Rules that only address one particular meaning from the language are simply not versatile enough to survive in the competitive environment.

5.7.6 A Model

This discussion leads us to a potential model of language development in creatures that emphasises the interaction between creature and language, but also the phenotype/genotype distinction we have created for language. An important aspect of this model would be the interaction between *creatures*. This takes place when they communicate, which is the breeding ground for language, but also when they attempt to complete tasks and so forth. A tough challenge will be to find some useful task that the creatures can perform which requires the use of language. The obvious example would be something like a foraging task, or possible a predation task. Hunting, especially would be an excellent candidate for a task involving communication needs. In that case, the language would involve semantics concerning the location and movement of a given prey.

The key to the model is that the creatures' actions be in some way influenced by their *interpretation* of the utterances of other creatures. Thus, not only would their internal grammar change through learning, but they would also *do* things differently.

We can get some idea of how interaction between two creatures might work by looking at figure 4. Here we can see the various links that would exist between creatures: a link between the side of the creature that deals with the world and the language used to express itself; a link between two creatures' languages, indicating learning of language; a link between a creature's language and another creature, dealing with the influence of others using language; and a link between the real-world aspects of creatures, indicating physical interaction.

Clearly, this model is complicated, and its implementation would be difficult. Nevertheless, an attempt at it could yield some fascinating thinking of the entire language problem, rather than just specific aspects of it. Additionally, the model could be limited in various fashions, perhaps by eliminating semantic grounding

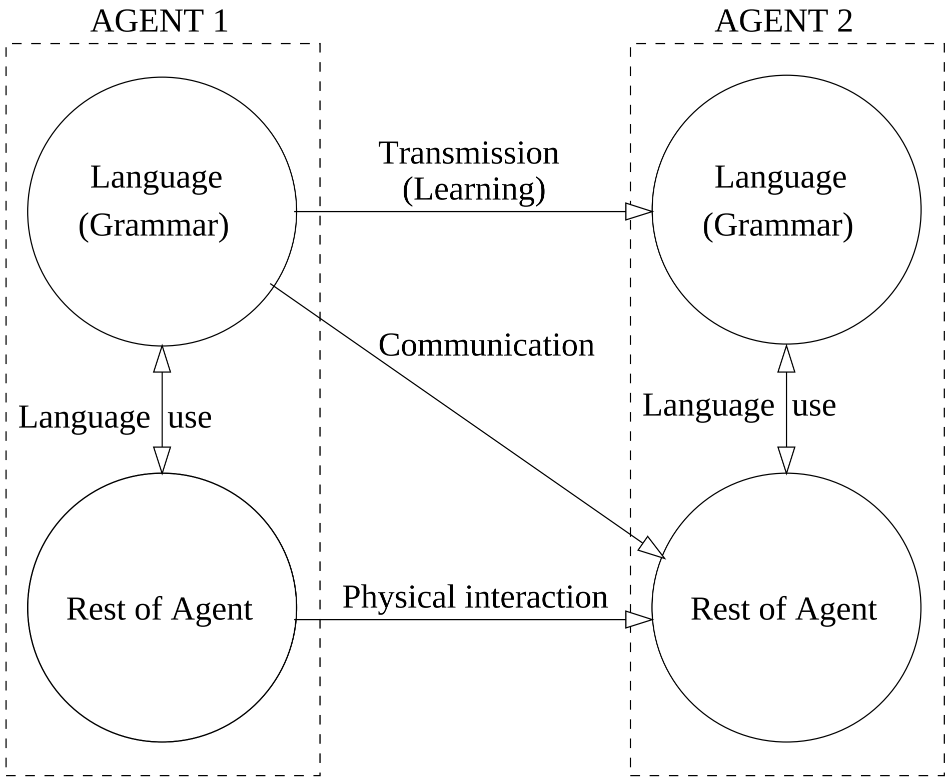

Figure 4: Agent Diagram

to begin with. The benefit of the model is that it at least *allows* the possibility of semantic grounding. The creatures exist in a real world, with other creatures, and thus their concepts and expression of those concepts can have meaning directly relevant to their existence.

Another aspect of the model which may prove interesting is that it can be viewed in at least four unique ways, essentially corresponding to the links in figure 4. First of all we can examine the whole picture: the interaction of creatures and their languages in an environment. This will naturally be the most difficult to understand, but should also be the most rewarding over time. The second view we can take is just of the creatures and their interactions, without looking at how language is taking place. Here we can observe how language helps the creatures complete their tasks, without worrying about the evolution of the language itself. Thirdly we can look solely at the evolution of the languages in the creatures. Here we will parallel Kirby and Batali's models, except that we know that the language evolution is based on several more factors, especially a real world task. The final angle we could take is to examine just a creature, and see how the interaction between the creature's "thought processes" and their language use works and evolves.

6 Semantics Evolution

6.1 Introduction

As we discussed in section 5.2.4 there is a real need to explore grounded semantics in models of language evolution. This was not achieved at all in the syntactic models we have examined, as they focussed solely on structure rather than meaning. Grounded semantics would be an excellent addition to the model concept introduced in section 5.7.6. This section will very briefly outline the work of Luc Steels in a particular article ([20]), as well as indicate other research by Steels.

6.2 Luc Steels - The Work

In [20], Luc Steels and Frederic Kaplan explore the idea of a visually grounded semantics for a group of robotic agents. The agents exist as programs on computers which have digital cameras attached so that the agent can "look" at a board in the room which has shapes of different colours on it. These are what the agents communicate about.

Communication takes place in the form of a game. The speaker says a word that indicates the object on the board that it is looking at, and then the listener looks at what it thinks the speaker meant. The speaker then tells the listener directly what object it really meant and both agents adjust themselves for future games based on whether the listener got it right or wrong. This adjustment involves increasing the score of a successfully used word and decreasing the score of unsuccessful words, including those that were passed over, but could have been used by the speaker.

For the agents a meaning is a category that distinguishes one thing from another. Thus a possible meaning of a word might be "HPOS < 0.5" which means "something that is in the left half of the board". If this successfully distinguishes just one object, then it *means* that object. The objective of the agents, then, is to develop a lexicon such that they can successfully indicate any object on the board to each other in the course of the game.

The agents play this game repeatedly until a shared lexicon develops. Steels and Kaplan examined the dynamics of this lexicon to draw some conclusions about how the robots are performing. Their conclusion is that the game ultimately leads toward incoherence, although there are some tendencies that combat this.

6.3 Luc Steels - Analysis

Steels and Kaplan claim that they are focusing specifically on what are known as the Gavagai and Grounding problems. The Gavagai problem is simply the issue that you can never know what somebody really *means* by something they say. This is the problem of not having mind-reading abilities. Thus, in the world, you only observe communicative success, not the underlying meanings. The Grounding problem is that different agents perceive the same thing differently because of orientation, and other perceptual differences. That being the case, languages need to abstract away from this sort of issue such that communication is possible between agents with different perceptual standpoints.

For the remainder of this section, rather than an analysis of the specifics of Steels' and Kaplan's experiment, we will note some of the interesting behaviours that emerged, and then finally analyse what progress was made on the Gavagai and Grounding problems.

6.3.1 "Just" Pointing

One issue about how meanings function in the model is that it doesn't seem much like what we mean by semantics in the real world. In the game the meaning of "lu" can be "HPOS < 0.5" or, in more real terms, "on the left half of the board". That does not seem to be what we mean when we say "the square", for example. This response, while reasonable for our highly developed semantics, does not seem relevant to what is being simulated here. If we agree with Calvin that pointing formed the origin of language [14, p.218], then the meanings in Steels' model simply pointing to objects makes sense. The purpose of meanings is that they indicate some set of perceptual properties of the thing being indicated, such that we can distinguish it from other things. The model only seems wrong because it uses just *one* perceptual property at a time, and thus seems overly simplified. In fact, this is likely to be a realistic representation of how language began.

6.3.2 No Mind Reading

> The agents cannot even know which meaning the other one uses because they have no access to each other's brain states.
> [20, p.13]

Perhaps the most interesting thing about this model is that there is no direct access to other agents' meanings, where "meaning" in this case refers to the specific concepts Steels has made available such as "HPOS < 0.5". This directly contrasts with the syntax modellers previously examined, who were able to tell *exactly* what the meaning of an utterance was more or less every time. Clearly, Steels' and Kaplan's approach is far closer to the way the world works: we cannot see exactly what people mean, but have to rely on observing the success or failure of our communications to estimate it as best as possible.

What this means is that there are situations where agents can successfully use a word to communicate, but have totally different meanings associated with the word. For example, "lu" might mean "left half of the board" to one agent, but "bright red" to another. So long as both those meanings pick out the same object, neither agents will be any the wiser, and communication will be just as successful as if they used the same meaning. While this is an extreme example, it does mirror the real world, where people do mean subtly different things by words, but not so much that communication breaks down.

A side effect of this ability to communicate without necessarily sharing meanings arises when the world changes. Steels and Kaplan observed that when they changed the world such that one meaning no long picked out a single object, the agent with that meaning would migrate to another meaning over time as it discovered another category to discriminate the object from others. This highly adaptive behaviour is one of the high-points of the study. Also interesting was that, the more times the environment was changed (by adding a new object), the *less* time it took for the agents to readjust and have a successful lexicon. Presumably this is because their discrimination categories became more sophisticated, and thus were able to handle new objects without having to change as much of the lexicon as previously.

6.3.3 The Gavagai Problem

Steels and Kaplan have certainly developed a system where the Gavagai problem exists. As in real language, though, the Gavagai problem is not so much a *problem* as a phenomenen that is dynamically handled by the nature of language. For the most part people do not agonise over the fact they do not know exactly what another person *means*. Successful communication is good enough. Similarly in the model, the agents can end up with different meanings for the same thing, as we saw, but this does not prevent successful communication either. Thus, while the Gavagai problem exists in the model, just like in the real world, it is not really a problem at all except in special circumstances.

6.3.4 The Grounding Problem

The grounding problem does appear to exist in the model also, and again seems to have been successfully addressed by the agents. Steels and Kaplan, however, do not seem to specifically raise the problem for the agents, or discuss it, so it is difficult to discuss. That said, a little speculation shows that the agents should be able to handle the problem much as people do.

This issue is mentioned briefly in the article where they say there might be a case where "'left' for the speaker may be 'right' for the hearer". This is a case of the same *meaning* denoting different things for different observers. That does not happen in the model, however, because meanings are not shared, only forms that express meanings. Therefore it is a little difficult to say that the grounding problem is really put to the test in this model. When the speaker identifies an object they do not say "the object on the left", they say "wu" which might *mean* the object on the left to *them*, but not necessarily to the listener. Therefore the grounding problem is not in play in this instance. Instead, what happens is that agents with different perspectives on things use the same word, but have different *meanings* behind the word.

6.3.5 Other Work

Luc Steels has performed an impressive amount of research on computational models of language, with a definite focus on the development of semantics. As we have seen from the previous analysis, his work is an interesting and careful examination of certain aspects of language. There is not enough time here to go into any further detail on Steels' other work, but we will conclude by indicating some other directions he has taken, along with paper references.

General Survey In [19] Steels surveys current work on language as a complex adaptive system and attempts to derive some of the key points that are being worked on. He feels that five principles in particular are important: reinforcement learning, self-organisation, selectionism, co-evolution through structural coupling, and level formation. A survey of work in the language evolution field can also be found in [17].

Syntax Steels has also done work on the development of syntax in visual robotic agents. This research can be found in [18] and deals with robots much as in the article discussed, observing a moving scene and communicating about it. Steels observed the evolution of distinctions, lexicon and eventually some syntactic structure in the language of the robots.

Perceptually Grounded Meaning Along with the work we have seen in [20], Steels has also looked into the area in other ways. In [16] Steels takes a more formal look at the process of semantic grounding in a discrimination task.

7 Conclusion

Over the course of this essay we have examined language from several angles. First, we looked at language from the standpoint of theorists who speculate as to its origins in section 2. This allowed us to understand some of the more established theories concerning language which were reflected on later in the essay when we discussed computational models. We also looked at a basic definition of language as syntax and semantics.

In section 3 we considered the utility of computational models in general and also noted some of the important caveats concerning them. We found that, as far as language modelling goes, the focus is on examining assumptions, rather than realistic simulation.

The descriptions of syntax models in section 4 gave us a good idea of the sort of work that is being performed in the world of computational modelling of language. Then, in section 5, we spent considerable time analysing that work. We found that there were some issues with the models, but largely spent time coming to a deeper understanding of what the modellers were trying to achieve. In section 5.7 we took our look at modelling possibilities deeper still by outlining how we might think of language itself as an agent and its interactions with the physical agents it lives in. This discussion lead us to roughly sketch a model that could be created which might shed yet more light on how language evolves.

Finally, in section 6, we very quickly outlined a particular piece of work by Luc Steels (and Frederic Kaplan) as well as pointed towards other research by Steels. This gave us the balance of having looked at both syntax *and* semantics. As we discussed earlier in the essay, the prospect of combining the two in one model is quite exciting, and, indeed, such work has been begun by Steels.

Overall, then, we have had a wide look at the work that has been performed on the emergence of language in populations as agents. We have also pointed towards further directions both by critiquing current models, and sketching a model of our own.

References

[1] John Batali. Computational simulations of the emergence of grammar. In James R. Hurford, Michael Studdert-Kennedy, and Chris Knight, editors, *Approaches to the Evolution of Language: Social and Cognitive Bases*, pages 405–426. Cambridge University Press, 1998.

[2] D. Byrne. Simulation: A way forward? *Sociological Research Online*, 2(2), 1997.

[3] Brett Calcott. Learning and evolution. Project for partial completion of the requirements of COMP473 at Victoria University of Wellington, 2001.

[4] Angelo Cangelosi. Evolution of communication and language using signals, symbols, and words. *IEEE Transactions on Evolutionary Computation*, 5(2), April 2001.

[5] Noam Chomsky. *Knowledge of Language Its Nature, Origin, and Use*. Praeger Publishers, 1986.

[6] Stephen Chong. Simulating the emergence of language via rule systems and genetic algorithms. Honours Project for the Mathematical and Computing Sciences Department of Victoria University, 1997.

[7] N. Gilbert. Emergence in social simulation. In N. Gilvert and R. Conte, editors, *Artificial Societies*. UCL Press, 1995.

[8] G. Johnson. *Fire in the Mind*. Viking, 1996.

[9] Steven Johnson. *Emergence*. Scribner, 2001.

[10] Simon Kirby. Syntax out of learning: The cultural evolution of structured communication in a population of induction algorithms. In *European Conference on Artificial Life*, pages 694–703, 1999.

[11] Simon Kirby. Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners. In Michael Studdert-Kennedy Chris Knight and James Hurford, editors, *The Evolutionary Emergence of Language: Social function and the origins of linguistic form*, pages 303–323. Cambridge University Press, 2000.

[12] Chris Knight, Michael Studdert-Kennedy, and James R. Hurford, editors. *The Evolutionary Emergence of Language: social function and the origins of linguistic form*. Cambridge University Press, 2000.

[13] David Lightfoot. *The Development of Language*. Blackwell Publishers Ltd, 1999.

[14] William Noble and Iain Davidson. *Human Evolution, Language and Mind.* Cambridge University Press, 1996.

[15] Brian Skyrms. *Evolution of the Social Contract.* Cambridge University Press, 1996.

[16] Luc Steels. Perceptually grounded meaning creation. In M. Tokoro, editor, *Proceedings of the International Conference on Multi-Agent Systems.* The MIT Press, 1996.

[17] Luc Steels. The synthetic modeling of language origins. *Evolution of Communication Journal*, 1(1):1–34, 1997.

[18] Luc Steels. The origins of syntax in visually grounded robotic agents. *Artificial Intelligence*, 103:1–24, 1998.

[19] Luc Steels. Language as a complex adaptive system. In M. Schoenauer, editor, *Proceedings of PPSN VI*, Lecture Notes in Computer Science. Springer-Verlag, September 2000.

[20] Luc Steels and Frederic Kaplan. Bootstrapping grounded word semantics. In T. Briscoe, editor, *Linguistic evolution through language acquisition: formal and computational models.* Cambridge University Press, 2001.

[21] John Maynard Smith & Eors Szathmary. *The Major Transitions in Evolution.* Oxford University Press, 1995.

[22] Willem Hero Zuidema and Paulien Hogeweg. Social patterns guide evolving grammars. In *3rd Conference The Evolution of Language*, 2000.