UNIVERSITY OF
Southampton

Academic Thesis: Declaration Of Authorship

I, Lewys Glendon Brace,

declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

An Investigation Into The Influence Of Population Structures And Dynamics On The Emergence Of Linguistic Systems Through Iterated Learning

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. Either none of this work has been published before submission, or parts of this work have been published as: [please list references below]:

Signed:......................................

Date: 19/12/2017

UNIVERSITY OF SOUTHAMPTON

An Investigation Into The Influence Of Population Structures And Dynamics On The Emergence Of Linguistic Systems Through Iterated Learning

by

Lewys G. Brace

A thesis submitted in partial fulfillment for the
degree of Doctor of Philosophy

in the
Faculty of Physical Sciences and Engineering
School of Electronics and Computer Science

December 2017

UNIVERSITY OF SOUTHAMPTON

FACULTY OF PHYSICAL SCIENCES AND ENGINEERING
SCHOOL OF ELECTRONICS AND COMPUTER SCIENCE

Doctor of Philosophy

by Lewys G. Brace

ABSTRACT

Human language pervades in a complex and ever-changing social milieu, and although the tendency and ability to learn languages are clearly innate, given the rate at which lexical items change, it is clear that social-cultural factors and ontogenetic development play a significant role in the way in which languages change over time. This has resulted in research concerned with human language evolution being dominated by two, umbrella-like, research questions. First, to what extent is the human language faculty the result of genetic endowment, and to what extent might it result from non-evolutionary factors such as constraints imposed by the fundamental nature of observational learning and social interaction? Second, to what extent are the observed characteristics of human language the result of evolutionary selection on language users, and to what extent are they the result of individuals shaping languages during their usage? This thesis is concerned with both questions, and focuses specifically on the role of social learning in shaping language.

There is now a growing body of work which indicates that much of the contemporary linguistic form seen in languages around the world is the result of said languages being influenced by the population structure and social dynamics of their language communities. This, combined with emerging evidence that suggests a strong association between the origins of human language and a coincidental, and dramatic, shift in social structure, means that investigating the nature of the relationship between linguistic form and social structure has the potential to offer powerful insights into the nature of human language evolution.

This thesis explores this notion of a relationship between the structure of a language community and the linguistic structures that their language exhibits by modelling language changes as arising within the context of a social-coordination problem. In doing so, it utilises a specific form of expression/induction simulations known as iterated learning models. The key principle of these models is that the training data offered to a language learner is, itself, the result of training and learning on the part of another language user.

Four different models are presented here. The first introduces the concept of iterated learning, and explores how compositional languages emerge in a population of language users. The second adopts the principles of Roth-Erev reinforcement learning to look at the evolution of term-based languages; again, in a population of language users. The third, uses both the iterated learning framework and the principles of Roth-Erev reinforcement learning in order to explore the nature of linguistic change in a situation whereby agents create their own signals and syntactic rules while their population size is in a state of flux. The final model is adapted from the third, and explores the emergence of contact languages that tend to arise when independent language communities interact.

All four models demonstrate that the structure and make-up of a population influences the dynamics of language change over generational time. Specifically, it is shown that, by increasing the number of trainers from which an agent learns, the agent in question tends to learn a more expressive and stable language at a much faster rate, and with less training data. It is also shown that, so long as the number of mature agents is large enough, this finding holds, even if a learner's trainers include other agents that do not yet possess full linguistic competence.

Importantly, the findings presented here demonstrate that it is not population size *per se* that dictates how long, if at all, a fully expressive and stable linguistic system takes to emerge. Rather, it is how proportionally interconnected a given agent is to other agents in the social group that dictates the success of said population's language.

In addition, the final model, which looks at the nature of pidgin and creole language emergence, presents two key findings. First, and in contrast to the common claim within the pidgin and creole literature, social power need not play a key role in pidgin emergence. Here, the pidginisation process needed to be a bilateral process, with both parties contributing to the subsequent pidgin in order for a successful contact language to exist between the two different populations. Secondly, this model looked at the concept of tertiary hybridisation; the belief that a pidgin will have to be used as the *lingua franca* between two groups who do not possess a common language, and whose speakers are not native speakers of the original target language. The data from these model runs indicated that, when two groups without any common language come together, tertiary hybridisation is necessary in order for a creole to emerge; otherwise, the resulting language is an entirely new linguistic system.

In summary, the results of these models demonstrate that the evolution of language does indeed have an intimate relationship with population structure and social dynamics. In that linguistic variations and systems become more stable in situations where language users have a higher level of interconnectivity with the rest of the population. The reason for this is shown to be due to the way in which languages themselves evolve in response to individual learner biases so as to become easier to learn. In other words, as language users learn the linguistic system of their particular social group, the language is essentially exposed to a refinement process as it is past on from one generation to the next. Furthermore, although it has been argued that, in order for a language to be

learnable, its structure has to adhere to certain constraints placed upon its structure, and that any language that violates such a ‘linguistic blueprint’ would not exist because it would be unlearnable, the findings presented here demonstrate that this refinement process is highly efficient at producing similar results; even when input is highly variable and inconsistent.

Contents

List of Figures

List of Tables

Acknowledgements

It has been said that a man is as good as the company he keeps, and this is probably why I'm so great. That said, there are three people in particular, to whom I would like to extend a special thanks. First, there is Lorenzo, without whom this thesis would have been completed a lot sooner. Secondly, I would like to thank Alice; whose colourful use of threats encouraged me to actually finish this body of work. Finally, and by no means least, I would like to thank Aby for being herself; this would not have been possible without you.

I would also like to express my gratitude to my supervisor Seth Bullock, whose thoughtful contributions and practical aid has proven to be invaluable in the production of this thesis. I would also like to thank both Jason Noble and Glyn Hicks for agreeing to supervise me to begin with, and placing me on the right tracks.

To Kat.

Chapter 1

Introduction

> Enki, the Lord of abundance, of trustworthy commands,
> The Lord of wisdom, who understands the land,
> The leader of the gods,
> Endowed with wisdom, the Lord of Eridu
> Changed the speech in their mouths, [brought] contention into it,
> Into the speech of man that (until then) had been one.
>
> Enmerkar in 'Enmerkar and the Lord of Aratta'

The natural world is rife with organisms that communicate with each other, using a range of mechanisms that vary wildly between species, including colourful visual displays, electric organ discharges, pheromones, signature whistles, alarm calls, and fully syntactic human language. Although fascinating in their own right, nearly all of the communication systems that we see in the natural world, including our own non-verbal ones, such as smiling and crying, lack an important feature. Namely, the ability to combine signals syntactically in order to communicate about remote matters that exist in spatial, temporal, even fictitious, worlds beyond the here-and-now.

Indeed, it is this aspect of human language that enables us to communicate any thought in our minds, from which berries are edible, through to explaining to an eccentric business CEO why his dinosaur-filled theme park will ultimately lead to chaos. This arguably makes language one of the most interesting evolutionary developments to emerge within the last 5-10 million years; with many deeming it to be the most distinguishing feature of our species, as well as the primary impetus behind our expansion from a minor species in sub-Saharan Africa to the most dominant species on the planet (Fitch 2010).

A large part of the uniqueness of human language arguably stems from the way in which it gains its expressive power by exploiting a few formal principles which operate over various sub-systems and different levels of organisation (Tallerman & Gibson 2013*a*). Firstly, language recombines elements at a series of successively more complex stages;

starting with the phonemes of the sound system (MacNeilage 2013). These phonemes then combine in language-specific ways to form syllables, which are combined to form morphemes and words; which are then combined to form sentences. Secondly, languages have grammars; rules governing which sound sequences are and are not permissible, as well as how the meaning of a valid utterance arises from the meaning of its component parts. Finally, human language is in a constant state of flux, in that each language tends to exhibit a high level of variation and tends to change over time.

These characteristics of language, and whether they should be considered part of nature or human culture, is a debate that can be dated back to the classical era. Examples include the study of rhetoric that flourished with the formation of a system of grammatical categories based on Greek and Latin, and how Panini developed a highly diverse grammatical system in fourth century India. The study of language then boomed during the Renaissance period, with the help of both the application of classic philosophical thinking to various disciplines, and the discovery of new languages from around the world (Gardner 1985).

However, the study of the evolution of language was arguably formalised by Charles Darwin, and his interest in courtship signals. Darwin's belief was that the diversity seen in human language was a product of a process akin to speciation. Although Darwin was primarily concerned with visual expressions of emotion, which he believed came about in order to convey information about the signaller's emotional and motivational state, he also applied the same thinking to vocal expressions. Darwin was adamant that both human gestural expressions and language were derived from a biological ancestor. In addition, he noted his strong belief that natural selection resulted in significant changes in the design features of the communication system:

> *"I cannot doubt that language owes its origin to the imitation and modification, aided by signs and gestures, of various natural sounds, the voices of other animals, and man's own distinctive cries. When we treat of sexual selection we shall see that primeval man, or rather some early progenitor of man, probably used his voice largely, as does one of the gibbon-apes at the present day, in producing true musical cadences, that in singing; we may conclude from a widely-spread analogy that this power would have been especially exerted during the courtship of the sexes, serving to express various emotions, as love, jealousy, triumph, and serving as a challenge to their rivals. The imitation by articulate sounds of musical cries might have given rise to words expressive of various complex emotions".*
> *- Darwin (1871:56)*

Essentially, Darwin made two points regarding the evolution of human language; that its structure and function is a product of natural selection, and that the strongest link between human and non-human communication systems lie in their ability to be able to convey an emotional state.

Since the time of Darwin, much work has been conducted into the origin and subsequent evolution of language[1], and it is now largely accepted that human language consists of a number of, what Hockett (1960) termed, *design features*. Smith (2003) collates the numerous design features proposed within the literature, selecting the following seven key design features of human language.

1. Discrete: An expression of a particular type can be a subset of a larger expression of the same type.

2. Cultural transmission: A language user learns their language from other users, in some sense and to a certain degree.

3. Arbitrariness: A signal is only arbitrarily related to its meaning.

4. Duality of patterning: A small number of meaningless elements, phonemes in spoken language, are combined in order to produce meaningful elements; words in human language.

5. Displacement: Languages can be used to communicate about things that may be spatially or temporally removed from the actual communicative act, or may even be fictional.

6. Stimulus freedom: Language users can produce any signal that they wish to at any given time, and not only when the appropriate stimulus is present.

7. Open-endedness: The set number of sentences that can be produced or interpreted by a language user is infinite.

This final point can be further broken down into *compositionality* and *recursiveness*. The former refers to the way in which the meaning of an expression is a function of the meaning of its parts, and the way in which they are combined. This makes the interpretation of previously unencountered signals or utterances possible. The latter concerns how an expression of a particular type can be a subset of a larger expression of the same type. Thus allowing the creation of an infinite number of utterances from a finite set. These are what endow human language with its aforementioned expressive power.

Today, when one says that they are exploring the evolution of language, they tend to mean one of two things. Either they are referring to the search for an 'evolutionary story' of how and why humans came to acquire language, or they are interested in the fundamental explanatory challenge for linguistics; namely, explaining how the universal

---

[1] The historical details of this body of work are too vast to go into in detail here, and the interested reader is advised to consult Hauser (1996) or Fitch (2010) for a detailed overview.

properties of language arise from the complex adaptive systems that underpin it (Kirby 2013).

It is quite easy to see how these two interests relate to one another. Exploring the nature of language transmission, and the results thereof, in modern humans may provide a window into the nature of protolanguage in pre-linguistic hominins. Particularly when looking at the emergence of contact languages that arise when two linguistic communities, without any prior common language need to communicate with one another. Indeed, recent years have seen an expansion in the body of work with that exact aim; which many believe can be achieved by understanding how phenomena, such as contact events between different linguistic groups, can impact the creation and transmission of modern languages.

This is perhaps why the vast majority of human language research is concerned with one of two questions (Tallerman & Gibson 2013*b*). First, how much of the human language faculty can be ascribed to a genetic endowment for language learning, and to what extent can it be accounted for by other mechanisms; such as the nature of observational learning and social interactions? Second, how much of linguistic structure can be attributed to natural selection, as in which linguistic features were adaptive to early hominins, and to what extent does structure arise from self-organisational processes; such as language itself adapting to be easier to learn?

In accordance with this observation, this thesis explores the role of social structure and social dynamics in shaping language change over time by framing the subject as a matter of social co-ordination. In doing so, it investigates as to how the make-up and interconnectivity of social groups impact linguistic structure. This is achieved by subscribing to the working assumption that human language evolution is the result of interactions between three dynamical systems; biology, ontogeny, and social-cultural factors. More specifically, there are emerging bodies of work which suggest that, not only does an association exist between the size and interconnectivity of social groups and linguistic structure (Johansson 1997, Wichmann & Holman 2009, Lupyan & Dale 2010, Milroy 2013, Atkinson et al. 2015), with a change in these sometimes claimed as a driver for the evolutionary origin of human language, but also that the study of contact languages could provide a 'window' through which to better understand language evolution. This thesis builds upon these works in exploring such notions by way of expression/induction (E/I) models.

Iterated Learning Models (ILM) are the specific type of E/I simulation that are the focal point of this thesis, and they were selected for two main reasons. Firstly, the *iterated learning* framework was specifically designed to explore the self-organisational properties of languages themselves, and how these properties can be influenced by learner bias; with linguistic structures themselves adapting in order to become easier to learn. Secondly, the developers of ILMs, and the vast majority of E/I models, have traditionally not been concerned with matters of social structures and social dynamics and the impact that these may have upon linguistic form. Instead they have opted for models that involve very small populations of a constant size, often consisting of just one mature agent and one immature agent per generation. This constitutes a research gap that this thesis aims to address. In short, this thesis has four main research aims:

1. Does population size have an impact upon the dynamics of iterated learning over generational time?

2. If so, in what manner does the make-up of a population of agents impact linguistic structure within iterated learning models?

3. Does a fluctuating population size impact the rate of linguistic change within an iterated learning model?

4. Can iterated learning models be adapted in order to explore the development of contact languages. If so, what inferences can be made from the data yielded by such simulations?

Given these aims, this thesis proceeds thusly. Chapter 2 provides a theoretical and historical overview of human language evolution research, with the aim of providing the reader with the background necessary in order to locate the research presented here within the wider literature. The second half of this chapter will then provide the reader with an understanding of the various research methodologies that have been adopted by those working in the field of language evolution, with an in-depth look at a specific methodology; the comparative method. By looking at the data available in regards to the cognitive capacities, linguistic abilities, and social behaviours of other species, this section highlights a number of points about these matters that justify the assumptions made by those developing computational models. Models that tend to take the form of a population of individual agents that initially have no language, but do have the abilities necessary to learn one; and are predisposed to do so.

Chapter 3 then takes an in-depth look at the methodology that is the primary focus of this thesis; the various computational methods that have been used to research human language and other naturally occurring communicative systems. This includes an overview of game theoretical models and the various grammar representations that have been developed and implemented in relevant computational models. There is then an overview of some of the most significant agent-based models that feature in the literature, with a focus on expression/induction (E/I) models. Finally, this chapter concludes with a brief discussion of how, given the evidence presented in the previous chapter and the fact that the vast majority of E/I models have very restricted population dynamics, that it would be beneficial to the research field to explore such models with more complex social dynamics which would not only aid in exploring a number of areas of theoretical interest in the field, but will also bridge a gap in the E/I literature.

Chapter 4 then takes a notable iterated learning model (ILM), originally introduced by Kirby & Hurford (2002), which explores how expressivity and stability change over generational time in a population consisting of one mature and one immature language user per generation. This model is then extended to include a population of agents with multiple mature and immature agents per generation. It is shown that being taught by multiple language trainers is beneficial to the learning process and enables a compositional language to emerge more quickly.

Chapter 5 is motivated by the possibility that the results reported in chapter 4 may have been idiosyncratic to the particular model employed. This chapter therefore presents a completely different model, based around cognitively simplistic agents and the principles of Roth-Erev reinforcement-learning. It is shown that this second model displays similar behaviour to the first, in that an expressive and stable communication system emerges much faster in a social milieu in which language users learn their language from multiple tutors. Indeed, it is demonstrated that each individual language learner has a more challenging learning task when exposed to multiple trainers who may not share exactly the same languages, especially at the start of the simulation, but the learnability of the language itself benefits across generations.

Chapter 6 presents a third E/I model; one that again involves a linguistic population giving rise to a language through their interactions. The results of this 'baseline' model demonstrates that, as with the first two models, a situation where immature language users learn their language from a higher number of mature language tutors is beneficial to the emergence of a stable linguistic system. Furthermore, this chapter explores what impact a contracting or expanding population size has upon the emergence of a linguistic system. In doing so, it is demonstrated that the communicative success rate, which acts as a proxy for the stability of the linguistic system, is not related to the overall size of the population *per se*, but to the degree of integration within the linguistic community.

Chapter 7 then begins with a brief overview of the contact language literature, including a discussion of the theories regarding their origin and development. The 'baseline' model from chapter 6 is then extended to allow different populations of language users, each of which have their own perfectly stable language in place, to interact with one another in order to give rise to a contact language. This pidgin and creole model is explored in a number of ways. Firstly, the impact of the interconnectivity of the populations upon linguistic development are explored, where similar results to those seen in the baseline model are reported. Secondly, the role of social power, whereby one population dominates more of the interactions than language users from the other population, is then investigated. Here it is shown that a change in social power on its own does not impact the development of a linguistic system in any meaningful way. Thirdly, there is an analysis of the role of tertiary hybridisation, where a third population is involved in the process of creole development. It is demonstrated that tertiary hybridisation is essential to the ability of the linguistic populations within the model to give rise to an actual creole and not just a whole new language. Additionally, it is shown that the kind of linguistic structures seen in creoles, which have traditionally been associated with social power, may actually be the result of tertiary hybridisation.

Finally, chapter 8 is the conclusion of the thesis. It summarises the findings of the research presented in the preceeding chapters, where it fits within the wider body of work in the literature, and offers a brief discussion of future work.

1.1 Publication notes

The reader should be aware that sections of this work have been published prior to the submission of the thesis, and that these specific sections are labelled to that effect. Specifically, a large part of the work featured in chapter 4 can be found in Brace et al. (2015); and both Seth Bullock and Jason Noble should be acknowledged for their time and thoughtful contributions to this work. The work presented in chapter 5 has also been published, and can be found in Brace & Bullock (2016). Again, a personal thanks goes to Seth Bullock for his thoughtful input.

Chapter 2

The Evolution of Language and Evidence Thereof

Human language, like all naturally occurring communication systems, is clearly the result of evolution. However, there is still contention surrounding the issue as to whether or not language itself is an adaptation, or merely the bi-product of the evolution of other aspects of our cognition. Bickerton (2013) breaks down the various views on the matter into four distinct arguments:

1. That language owes its origins to natural selection, and that every subsequent step in its development was guided by natural selection (Pinker & Bloom 1990, Pinker 1994).

2. Language began through the process of natural selection, and was subsequently adaptive as a whole, but particular steps in its development were not selected for *per se* (Jackendoff 2002).

3. Some of the particular prerequisites for language were selected for non-linguistic reasons, and language was then a resulting emergent property; but language itself, was not directly selected for (Jenkins 2000).

4. A suite of cognitive characteristics was selected for, which were unrelated to language, but which made the cultural invention of language possible (Donald 1991, Fauconnier & Turner 2008).

Part of this difference of opinions is due to there being two prolific academic traditions in human language research. According to the first, language was not promoted by natural selection, but represented a new cognitive domain that resulted from a single biological mutation. This, in turn, enabled the unlimited control of complex structures, which was used primarily for private thought, and only derivatively for communication with conspecifics. In contrast, the second argues that an accumulation of biological mutations, which were promoted by natural selection, gradually allowed ever greater interaction between pre-existing cognitive domains; no new domain was created. This allowed increasing control over complex structures, which were primarily used for communicative actions, but also allowed for advanced private thought (Hurford 2014).

Those that support the latter are often referred to as *adaptionists*, and their account tends to emphasise the role of natural selection, and how organisms evolve in accordance with the selection pressures placed upon them by their environment. Specifically, the adaptionist view point argues that the language faculty, whatever one deems that to include, is not only the result of gradual evolutionary development, but that all aspects of it also presented an evolutionary advantage in, and of, themselves. Thus, the adaptionists argue that, although human language is primarily a communication mechanism that allows for social interactions, it does stem from specific mental activity.

In contrast, those that support the former are known as *non-adaptionists* or *exadaptionists*. This group tend to focus on concepts such as *exaption* when arguing that:

> *"Organisms must be analysed as integrated wholes, with baupläne so constrained by phylogenetic heritage, pathways of development, and general architecture that constrain themselves and become more interesting and more important in delimiting pathways of change than the selective force that may mediate change when it occurs."*
> *- Gould & Lewontin (1979:281).*

These architectural constraints can lead to *sprandels*, traits that serve functions that it did not originally evolve to do, but were eventually co-opted for; such as feathers. The non-adaptionist view is that natural selection played only a minor role in the evolution of the language faculty. Instead, the belief is that the language faculty appeared very suddenly, without any primitive precursors as a result of an evolutionary by-product, a sprandel, by some, as of yet poorly understood, laws of form or as the result of macro-evolutionary changes that were caused by a single point mutation (Tallerman & Gibson 2013*a*). This then enabled language, which in turn, enabled mental representations and other aspects of human cognition; such as future planning and constructing relationships between symbols.

2.0.1 Chomsky

The most notable non-adaptionist is probably Noam Chomsky, who argues that humans are the only species to have an autonomous module in the brain, a *language organ*, which evolved to carry out combinational calculations. This module is a by-product of selection for other cognitive abilities and provides algorithms for specifying the details of our communicative utterances.

Chomsky's (1957) argument was that all language users have an intuitive understanding of certain language properties. He advanced this idea by exploring the rules that govern three aspects of language use:

1. An individual's ability to generate grammatically correct sentences.

2. The ability to know that these sentences are correct, and what they mean.

3. The ability to work out if sentences violate these norms.

By exploring the nature of these three skills, Chomsky came to the conclusion that language users must possess a set of rules or procedures for understanding when different parts of speech can occur within and among sentences; rules and procedures that adhere to the intuitions of said language users.

Chomsky (1980) went on to study grammar in theoretical terms as the search for *universal grammar* (UG). UG is language-specific and genetically determined within the mind of individuals, and is then specified, sharpened, and refined in accordance with the personal experiences of individuals. This process then yields the particular grammars found within groups of individuals. Chomsky (1986) discusses this process in regards to the concepts of *I-Language* and *E-language*:

I-language: This is the pattern of neurons that implement an individual language user's grammar within their brain.

E-language: This is the set of utterances that make up the spoken language.

Chomsky (1986) goes on to assert how UG can be refined and decomposed into two specific parts. The first being a system that is associated with certain variable parameters, which creates the hypothesis place for I-Language. The second being a *Language Acquisition Device* (LAD), which determines how the external experiences of an individual impacts the selection of a particular I-Language from the range of those possible. Here, Chomsky (1980) asserts that language knowledge is a series of states within the individual's brain. States that can be described by a core grammar, consisting of certain principles of UG.

This has been used to explain the problematic nature of language acquisition, known as the *poverty of the stimulus*. This is the argument that, given how the linguistic training data presented to language learners is indigent, the only explanation for how we came to acquire our linguistic competence is that some knowledge of language must be prespecified. The UG then guides the learning process in such a way that children are able to learn their native language, despite the impoverished nature of the input.

2.0.2 Wilkins and Wakefield

Wilkins & Wakefield (1995) are non-adaptionists, who are heavily motivated by the way in which linguistics has traditionally viewed the question of *how* language first emerged as a somewhat unnatural one to ask. In looking at this question, they attempt to present a plausible evolutionary scenario for the emergence of the neurological preconditions within the hominin lineage that were necessary for the development of language. The authors focus on how there was a paired expansion of the frontal and parietal neocortex within Pleistocene primate lineages. This resulted in an incipient Broca's area and a unique junction of the parietal, occipital, and temporal lobes of the brain (POT); which had a different configuration in these ancient hominins.

They argue that the evolution of an identifiable Broca's area and POT was related to the evolution of motor programs; behaviours such as manual manipulation and throwing. Their evidence for this stems from the observation that the modifications to the motor systems that were related to the use of the hands and thumbs of early hominins had strong interactive relationships with forearm development, brain organisation, and lateralisation; all of which have a strong association with language. Their belief is that it was selection pressures acting on the improvement of motor skills, and the feedback between the somatosensory cortex and the motor cortex that these new skills required, that led to the creation of the connection between the sensory and motor cortex; including Broca's area.

The essential idea here is that auditory, visual, and somatosensory inputs were processed in unimodal association areas of the cortex, to create unimodal representations. These then come together to form multi-modal representations of the original sensory input, which then converge at the POT, and in turn enables us to form amodal representations is crucial to linguistic abilities.

In other words Wilkins & Wakefield (1995) argue that an abstract semantic representation is created in Broca's area, by structuring the amodal representations of the external world. A similar process then links visual and auditory sensory input to these amodal concepts; thus allowing arbitrary symbolic reference to become possible.

2.0.3 Pinker

The seminal paper for the adaptionist side is Pinker & Bloom's (1990) *Natural Language and Natural Selection*, which revived interest in human language evolution research. This paper, and Pinker's (1994) subsequent book *The Language Instinct*, were an attempt to fit theories of language structure and function within a Darwinian framework.

Pinker & Bloom (1990) argue that the belief that language evolved as a by-product of selection for other abilities, or that grammar shows no genetic variation and therefore offers no selective advantage, are built upon inaccurate assumptions. Contrary to such beliefs, Pinker (1994) argues that natural selection is the only possible mechanism that can account for a trait that possesses the kind of complex design features seen in human language; meaning that it could only have evolved through gradual changes. Furthermore, Pinker (1994) observes three aspects of human language, which he believes demonstrates a degree of genetic variation that are sufficient for selection to act upon:

1. Variation between normal humans in grammatical competence.

2. In those families with preponderance of left-handers, right-handed individuals rely more on lexical than syntactical analysis.

3. Potential evidence of inherited deficit in grammatical ability within a family.

Pinker (1994) uses these observations in arguing that all humans are born with a *language instinct*. This is a form of universal grammar that constrains the range of potential variation in language, allowing for both linguistic diversity and universals. He goes on to argue that this language instinct is located within the brain, lateralised to the left hemisphere; and although a modular view of the neural substrate for language is favoured, evidence of a non-modular, multi-purpose system would not cause problems for the Darwinian account since selection can favour a structure with more than one function.

2.0.4 Bickerton

Bickerton is very much a non-adaptionist who becomes more adaptionist in later years. His work is largely concerned with a central paradox that he believes to underpin the evolution of language. Namely, that human language is very different to any other natural communicative system:

> *"On the one hand, language could not have evolved out of animal systems of communication. On the other hand, language must have evolved, since, for all its complexity, it is simply one of the countless adaptive mechanisms that have developed in species in the course of evolution"*
> *Bickerton (1990:75)*

Bickerton (1990) believes that one can overcome this paradox by accepting that language is first and foremost a system of representation, and is therefore like other communication system. In that it was a mechanism that, to a large extent, created its own output, and did not merely replicate, or reproduce in a similar form, its input. Bickerton (1990) also accredits this fact for enabling language to greatly increase, not just the number of things that individuals were able to communicate about, but also the *kinds* of things that could be discussed. Thus, he concluded that one would not be able to investigate the origins of language by looking at the methods by which other animals communicate with one another. Moreover, one would have to look at how systems of representation evolved, which refers to the way in which animals tend to use *primary representations* of whole situations, such as 'predator' and 'food', whereas human language is a *secondary representational system*; with verbs denoting actions, etc.

As such, the focal point of Bickerton's (1990) work concerns, what he termed, *protolanguage*. He believes that this represents the main transition between no language and fully syntactic language. This is a system that uses *secondary* representation, much like full language, even though it lacks many formal properties that are characteristic of full language; such as grammatical morphemes. He argues that this protolanguage first evolved as a result of group foraging activities (Bickerton 2002), and that full language then evolved as a result of a catastrophic processes. This came in the form of a single macro-mutation, which was responsible for considerable functional changes, such as the capacity for syntax and the modern human vocal tract.

In exploring the nature of protolanguage, Bickerton became particularly interested in cases where cultures came together. Specifically how, when two communities without a common language need to communicate, a form of protolanguage tends to emerge; this is referred to as a *pidgin*. Although the structure of a pidgin is relatively simple, we often see a refinement in its structure and usage as time progresses and subsequent generations are born. This refined version is known as a *creole*. It has been argued that the transition from pidgin to creole demonstrates that innate mechanisms guide the original rules for language production, and subsequently constrain language structure and usage (Hauser 1996).

In his later work, Bickerton (1998) retires the idea of language being the result of a macro-mutation, perhaps due to increasing evidence suggesting that the chances of a single mutation resulting in a system as complex and as ideally suited to a task such as language is staggeringly high (Pinker & Bloom 1990, Wolpoff et al. 2004).

However, Bickerton (1998) does maintain his belief that language emerged suddenly, within a single generation. The argument here is that protolanguage users already possessed the anatomical and neural apparatus, such as motor control and a primitive vocal tract; and that, through ontogenetic chance, these parts of the brain became connected and enabled the emergence of deep conceptual structure, which in turn allowed for the emergence of language as we know it today. This line of inquiry has, in more recent years, led Bickerton (2003) to argue that human language, as commonly seen today, involves the coming together of three things; modality, symbolism, and structure.

The first refers to speech and language, where he believes language to have begun as a mixture of both speech and signs. Information transfer was what selected for improved vocal capabilities, which evolved from the original mixture of grunts and gestures due to it better enabling individuals to communicate at night, over longer distances, or through dense vegetation. Thus, Bickerton (2003) concludes that speech modality was not the driving force behind language evolution. Moreover, it was contingent on the other two components of symbolism and structure; “being able to speak more clearly does not, in and of itself, give you more to say” (ibid:81).

For Bickerton (2003), the most important thing to bear in mind about the emergence of symbolic representation is that it must have been primarily a cultural, rather than a biological, evolutionary event. The widespread nature of potential for symbolic representation within the animal kingdom is indicative of analogous, rather than homologous development, akin to the convergent evolutionary process that gave fins to sharks, cetaceans and Ichthyosaurs; and could, therefore, exist within any animal with a brain of sufficient complexity. Thus, the emergence of symbolism arose culturally because the minimal biological equipment necessary was already in place, and it evolved because the exploitation of said symbolism benefited both individuals and groups.

In relation to structure, Bickerton (2003) supports the recently popular notion that holistic utterances, which were semantically equivalent to one-clause sentences in modern language, formed the earliest linguistic utterances. He believes that once the number of symbols began to increase, they started to be strung together in an ad-hoc fashion; and that these one-clause sentences, with a fixed word order, constituted a proto-syntax. To Bickerton (2003), this process does not merely constitute a step towards real syntax, but that once one achieves such a level of structure, real syntax will automatically follow.

2.0.5 Jackendoff

Jackendoff (2002) is another adaptionist, who focuses on various evolutionary stages of the human language faculty, while arguing that any increase in precision of communication or expressive power would offer a clear selective advantage. In presenting the evolution of the human language faculty in terms of a number of sequential evolutionary steps, Jackendoff (2002) lays out an evolutionary road map (figure 2.1).

For Jackendoff (2002) the most important preconditions for human language can be seen in the conceptual abilities of non-human primates and their communicative/alarm call system behaviours. He cites the many studies that have been done over the years (Kohler 1927, Cheney & Seyfarth 1990, de Waal 1996) in arguing that it is clear that primates have a combinational system of conceptual structure in place. A system that is not only able to deal with such matters as navigation and problem-solving in the physical

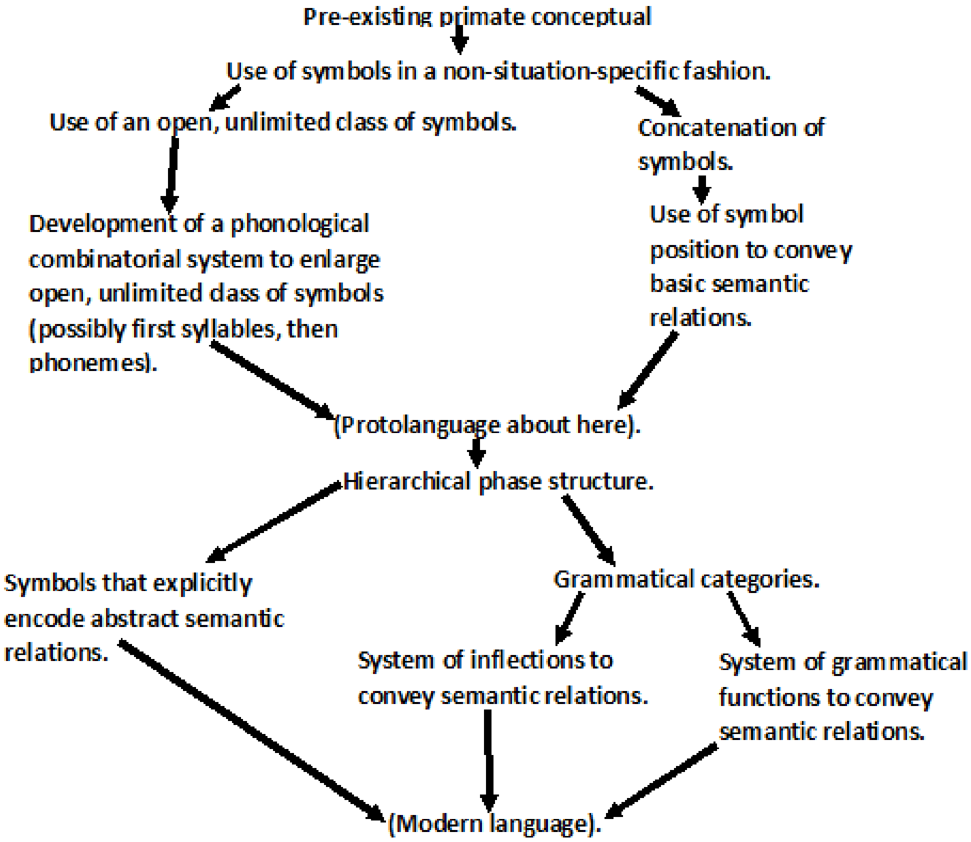


Figure 2.1: Summary of Jackendoff's (2002) incremental evolutionary steps of language. Logically sequential steps are presented from top to bottom, while logically dependent steps are presented side-by-side. Replicated from Jackendoff (2002).

environment, but which is also able to process other individuals as being intentional agents.

These primate systems also demonstrate the use of symbolism, which was a major stepping stone on the evolutionary road to human language. For Jackendoff (2002), the most important point to be made here is the observation that young human children, unlike non-human primates, use symbols in a non-situation-specific fashion. He observes that a child can use the word 'dog' to point out, summon, or enquire the whereabouts of the house pet, while non-human primates will give predator-specific alarm calls when they see a predator, but not to enquire as to when one was last seen.

According to Jackendoff (2002), the next evolutionary step required two things; an unlimitedly large class of symbols and the ability to concatenate them into larger utterances. He argues that these can be independent of one another, in that one could have a communication system with just one or the other. Jackendoff (2002) goes on to argue that the former stems from the capacity to imitate and, in agreement with others (Povinelli et al. 2000, Hurford 2007), behaviours such as pointing.

Jackendoff (2002) then goes on to discuss, what he terms, *proto-phonology*. The idea here is that, as the number of symbols increases, remembering and discriminating between utterances becomes a problem of increasing perceptual difficultly. Human language gets around this by building its words up by combining a few dozen, smaller and meaningless speech sounds. Jackendoff (2002) then proposes that an intermediate stage in the evolution of phoneme-based vocabulary was possibly based around the syllable as the generative unit. This is an argument that largely draws upon Lieberman's (1984) observations that the shape of the late Neanderthal's vocal tract would not have allowed for the multitude of perceptually different speech sounds evident in human language, but an open vocabulary would have still been possible with a less differentiated phonological system.

Jackendoff (2002) then goes on to argue that part of this proto-phonological development was related to a very important linguistic evolutionary step. Namely, the concatenating of two or more symbols into a single utterance, with a connection that was purely contextual. It should be noted here that Jackendoff (2002) draws a distinction between phonological combinations and this symbolic kind of combination. The former being a way of analysing and producing meaningful symbols from a repertoire of smaller, meaningless, units. While the latter is concerned with the putting together of meaningful symbols to form larger utterances; the meaning of which is a function of the meanings of the constituent symbols. This, for Jackendoff (2002), opens up the opportunity for enhancing the expressivity and precision of the communicative system through the use of linear ordering of concatenated symbols in order to express relations between said symbols, combined with the introduction of new kinds of vocabulary items that convey relations explicitly.

A Bickerton-esque notion of protolanguage then follows on from the previous evolutionary steps. However, for Jackendoff (2002), such a protolanguage demonstrates the earliest forms of syntax; in that semantic relationships and their positions come into play prior to the emergence of protolanguage.

All of the above steps use word ordering to signal the semantic relations between speech units. However, human language requires the use of headed phrases in grammar, which "allows principles of word order to be elaborated into principles of phrase order" (ibid:252). There are then, according to Jackendoff (2002), two more steps to be taken before modern human language can come about; steps that are dependent upon one another. The first of which is the ability to encode semantic relations among words and phrases by way of inventing words with which to express them; something that only becomes necessary once multiple-symbol utterances are possible. In human language, these tend to be in the form of morphological affixes and realised variants of word order.

The second concerns grammatical categories, and can, itself, be broken down into two dependent steps. Firstly, there is the arrival of the grammatical distinctions between nouns and verbs, along with the further differentiation of adjectives, prepositions, and the like. For Jackendoff (2002), it is the development of phrasal morphology and morphosyntax that then finally allows for the emergence of modern human language, which he argues can be seen as a series of independently evolved systems, each of which was built upon the protolanguage system, and which refined communication systems through their own expressive techniques.

2.1 Human Language Research Methodologies

From the brief overview presented thus far, it is clear that there are vast differences in opinion concerning the evolution of human language. Differences that can be attributed to a lack of evolutionary precursors, which in turn leads to a distinct lack of direct evidence, and uncontroversial, indirect evidence for any aspect of human language (Tallerman & Gibson 2013*a*). This has led to the issue of the origin and evolution of human language being described as "the hardest problem in science" (Christiansen & Kirby 2003). Although research conducted over the last couple of decades has given rise to a plethora of data, any potential insights are somewhat blurred by the way in which different disciplines vary in terms of what is considered acceptable evidence.

2.1.1 Archaeology

It has been argued that artefacts left behind by our hominin ancestors can offer insights into their symbolic and communicative abilities. This is certainly true, but there are two problems with this approach.

First, the archaeological record requires continual updating and re-evaluation. It was traditionally thought, for example, that stone-flaking first appeared in the Upper Palaeolithic, but it has since been shown that this skill actually appeared long before this point (D'Errico & Vanhaeran 2013).

Second, much of the archaeological evidence is largely open to interpretation. This led Balari et al. (2013) to conclude that the nature of the archaeological record made it difficult to make inferences about the nature of symbolic or cultural fossils, which are deemed to be more closely associated with language. Despite such evidence always being subject to inference, as the archaeological record increases, our conjecture about the phylogenetic history and the evolutionary pressures involved in shaping our species, will hopefully become more guided.

This is particularly the case with archaeological evidence of symbolic thought and cultural practices. Donald (1991) argues for a three-stage development of our capacity for symbolic representation:

1. Mimetic culture: This was the first step on the road to our current symbolic capacity. It was characterised by a significant improvement in motor control, and enabled our ancestors to voluntarily rehearse and refine their body movements. Donald (1991) goes on to posit that *Homo erectus* then utilised these skills in order to produce certain pre-linguistic symbolisms, such as tool making and cultural practices.

2. Mythic cultures: This second stage involved the development of speech and symbol use.

3. Technological culture: Our cognitive capacities then became dominated by one-to-one communication, which continually changes in form. This change first emerged in the form of Upper Palaeolithic visual symbols, which acted as a form of external symbolic storage; that allowed external memory to develop through the use of symbolism. Donald (1991) goes on to argue how this external storage capacity is only likely to increase with advancements in technology.

Indeed, there is much evidence concerning the use of symbolism within early hominins. Following this three-stage process, Balari et al. (2013) cite the work of others in arguing that this third stage should include objects that are made from antler, bone and ivory; and which are engraved with marks that appear to be some form of abstract notational system (Marshack 1991, Errico et al. 2003).

Furthermore, Errico et al. (2003) argued that pieces of ochre found in South Africa's Blombos cave with similar engravings are the earliest example of the kind of external memory storage discussed by Donald (1991). The authors then go on to argue how it appears as though true notational systems began to emerge at the start of the Upper Palaeolithic, and become incredibly complex by the end of that period; which they also argue demonstrates the use of an hierarchical organisation of information. As Balari et al. (2013) observe, this argument fits well with the lack of evidence for the use of symbolism in earlier hominins, particularly for the lack of notational systems prior to the arrival of anatomically modern humans (AMH).

2.1.2 Paleoanthropology

Paleoanthropology primarily uses the fossil record and skull endocasts in order to examine anatomical evidence of brain structures in ancient hominin species. This body of work largely involves finding homologies of the neuronal substrate of language, which in turn, is largely concerned with attributing equivalences of Broca's and Wernicke's area to other hominid species. For example, some researchers have posited that Australopithecines had a homologues Broca's area, and that the parieto-temporal area where Wernicke's area is localised underwent a reorganisation in *Homo habilis* (Balari et al. 2013).

However, there are various issues with this approach, such as the homologues areas in non-human primates having no linguistic specialisation, for example. Thus, Striedter (2005) argues that a much more informative measure of the relevance of these homologues areas would involve studying their level of interconnection, which is impossible to ascertain due to the way in which these connections, which consist of superficial cortical fibres, leave no fossil traces. Furthermore, much of the research on the homologous structures of the neuronal substrate of language is based upon a very strict application of the homology concept (Balari et al. 2013). In other words, it largely just looks at our phyletic group, and our relatively close ancestors.

This body of work does, however, support many of the arguments presented in the *social brain hypothesis*. The argument here being that the brains of primates are larger than those seen in other species due to the way in which the cognitive demands of living in large social groups creates a selection pressure for more cognitively powerful brains (Dunbar 1993, 2003*b*, Reader & Laland 2002). Those that subscribe to this notion argue that, while it is clear that language evolved in order to convey information, the focus of research should be on *what* information it evolved to share in the first place.

The traditional answer to this question is that the information encoded by language was primarily concerned with hunting and the use and manufacturing of tools. However, Dunbar (2003*a*) argues that such an explanation is unsatisfactory on two grounds. Firstly, matters pertaining to technological innovations take up a relatively small proportion of our time. Secondly, when we do partake in them, we rarely use language while doing so; with hunting being best done in silence and tool manufacture being past on more efficiently by demonstration rather than instruction. Thus, the argument here is that the reason as to why language is unique to the hominin line is because it essentially evolved for a completely social function. Namely, to allow the bonding of large social groups (Dunbar 1993, 1996*b*, 2003*b*, Deacon 1997).

Dunbar (1996*b*) argues that, in primate societies, grooming acts as a kind of ‘social glue’, and that, in larger groups there is not enough time to groom a good number of individuals. As such, vocal communication evolved as a way to bond larger populations of individuals. Thus we see that neocortex size co-varies with social network size (Dunbar 1996*a*). Indeed, Dunbar’s findings (Dunbar 1996*a*,*b*, Kudo & Dunbar 2001), coupled with a notable correlation between both neocortex size and rate of tactical deception in primates (Byrne & Corp 2004), and diet and social structure within anthropoid primates and their relative neocortex size (Sawaguchi 1992) make it arguable that an increase in brain size served this function in cases of increased social complexity (Hurford 2007). Further support for an association between the complexity of vocalisations within primates and the size, organisation, and distribution of social groups comes from data on group size and vocalisation size in forty-two primate species; which concluded that there is a “clear and strong relationship between social group size, grooming time, and vocal repertoire size" (McComb & Semple 2005:3). Thusly, Aiello & Dunbar (1993), posited that language could be seen as a multi-stage evolutionary process:

1. An initial form, much like conventional non-human primate contact calls, that served to keep track of other group members and warn of near-by predators.

2. A more developed form of this, which uses chorusing, and is designed to overcome the physical constraints on grooming that limit group sizes to around thirty to sixty individuals.

3. A more fully fledged language that uses grammatical structures to convey social information.

4. A fully developed, modern language, capable of abstract symbolic representation of concepts. At this stage changes may be more at the 'software' level than the 'hardware' level.

Each of these stages came about during the course of hominin evolution, as ecological changes brought around the need for change in social dynamics.

2.1.3 Morphological Data

The articulatory and perceptual systems of early hominins has also been studied in-depth. However, inferences concerning a creature's mental or behavioural nature made from such data are always questionable due to such characteristics being opaque in relation to the underlying structures involved. This is known as as the *form-function problem*, and refers to the way in which two species can have the same morphological structure, but use it for two different functions (Balari et al. 2013).

Many of the discussions concerning the evolution of our articulatory system stem from the work of Philip Lieberman; who studies the hominin, particularly Neanderthal, larynx. Lieberman (1984) argues that the linguistic ability of humans is based upon general neural mechanisms that structure our cognitive behaviour, as well as a limited set of language-specific mechanisms that differentiate the manner in which we transfer information. This leads Lieberman (1984) to argue that the neural mechanisms that structure other aspects of human cognition give rise to many aspects of human syntactic structure and semantic representation; homologous mechanisms that structure the behaviour of other species. The language-specific mechanisms are deemed to be the neural input-output mechanisms for the production of speech, combined with some of the neural mechanisms involved with the structure of rule-governed syntax. Thus, Lieberman (1984) posits that human language, in its present form, is the product of how these language-specific mechanisms work along with the more general neural substrates for cognition.

Human speech, as well as that of many nonhuman species, is largely considered to be the result of a source of sound energy, such as the larynx, being modulated by a filter function, which is determined by the shape of the supralaryngeal vocal tract. Thus determining the phonetic quality of the sound produced. Lieberman (1968) found that the vocal mechanisms of captive rhesus monkeys, chimpanzees and gorillas did not appear to be able to produce human speech, and that these creatures did not alter the shape of their supralaryngeal vocal tracts by means of tongue manoeuvres during vocalisations. This, along with a number of observations concerning the key differential features between the vocal tracts of humans and nonhuman primates (table 2.1), led Lieberman (1968) to the conclusion that “speech cannot be viewed as an overlaid function that makes use of a vocal tract that has evolved solely for respiratory and deglutition purposes” (ibid:1574). He further concluded that the hominid fossil record displays a series of changes from the primate vocal tract, that may have been partly for the purpose of speech. Furthermore, Lieberman (1968) argued how nonhuman primates are unable to produce the point vowels that are seen in all human languages (/a/, /i/, /u/), and that they use a limited set of articulatory gestures in order to modify the resonances of the supralaryngeal vocal tract. Lieberman (1984) thus argued for the idea that a process of gradual anatomical change can, at certain points, lead to functional advantages that will ultimately lead to different patterns of behaviour within a species.

Feature	Humans	Nonhumans
Position of larynx in throat	Low	High
Length of mandible	Long	Short
Size and shape of tongue	Large and rounded	Short and flat
Angle between pharyngeal and oral cavities	Approximately 90 degree end	No bend or very slight
Number of resonating cavities	Two-plus	one

Table 2.1: Featural comparison between the supralaryngeal vocal tracts of humans and nonhuman primates. Replicated from Hauser (1996).

This logic led to Lieberman (1968) making three important points regarding the vocal tract anatomy of humans. First, the supralaryngeal vocal tract of a neonate is more like that of a nonhuman primate, with it dropping at around three months of age before becoming adult-like at the ‘babbling’ stage at six months. Secondly, the vocal tract of modern humans would not fit into a Neanderthal. The belief here was that the larynx was placed significantly higher in Neanderthals, and moved lower down the throat with the arrival of anatomically modern humans (Negus 1949, Lieberman 2006). Therefore, the Neanderthals lacked an adult-human supralaryngeal vocal tract, with theirs being more akin to that of human infants and non-human primates. As such, they were unable to produce the quantal vowels /a/, /i/, and /u/. Thus indicating that human speech emerged with the evolution of the modern human vocal tract post-Neanderthals.

Whether or not the Neanderthal larynx was high or low in the throat is unclear due to the fact that the soft tissue that makes up the vocal tract does not fossilise, forcing all reconstructions to rely upon indirect inferences (Balari et al. 2013). However, Boe et al. (1999) demonstrated that, by calculating the position of the larynx and the length ratio of the pharyngeal cavity to the oral cavity, that the ratios are within range of those of a modern human female or child. Using an articulatory model, the authors also argued that the vowel space of Neanderthals was not much smaller than that of modern humans; and that they were, therefore, not handicapped for speech. In addition, by attempting to reconstruct the auditory system of *Homo heidelbergensis*, Martinez et al. (2008) demonstrated that the skeletal anatomy of the outer and middle ear, which supports the perception of human speech were present in this species; which can be dated back to around 5,000 years ago. As Balari et al. (2013) observe, although what sounds *Homo heidelbergensis* or the Neanderthals made is unkown from the audiogram produced by Martinez et al. (2008), it can be seen that they were able to discriminate between /i/ and /u/, as well as being able to distinguish relevant acoustic properties of sibilant sounds like /s/.

2.1.4 Genetics and Molecular Biology

The results from the sequencing of the human genome revealed that humans only have around 20,000-25,000 genes, far fewer than anticipated (International Human Genome Sequencing Consortium, 2004). This, combined with how humans share 99% of their genome with chimpanzees and bonobos, suggests that relatively few genes are responsible for the differences between ourselves and the Paninas (Gibson 2002). Two factors account for this; most genes are pleiotropic, meaning that they control more than one trait, and many genes are regulatory in nature.

These regulatory genes serve as binary switches that turn multiple downstream genes on or off, and those that are active during the early stages of development can have a profound impact upon phenotypic traits of individuals. Thus, it is believed that the small number of genetic differences between ourselves and the Paninas is largely the result of regulatory genes (Tallerman & Gibson 2013*a*). These observations support the argument against the idea that the distinctively human aspects of neurology, behaviour, and language are controlled by distinct genes (Gibson 2002, Diller & Cann 2013).

An area of genetic research that has been particularly interesting over recent years, for those studying language and the communicative systems of other species, concerns the discovery of *FOXP2*. This is a regulatory gene that dictates when and where other genes are transcribed into proteins. While it is a very conserved gene among species from worm to humans, meaning it has not changed much throughout evolutionary time, the number of genes regulated by it has increased gradually throughout evolutionary time (Okanoya 2013). More specifically, three point mutations have occurred in *FOXP2* throughout the course of mammalian evolution; two of these within the six million years or so since we shared a common ancestor with chimpanzees. This has led to the belief that the mutation of this gene was accelerated and reserved in the human lineage (Lai et al. 2001, Okanoya 2013).

Following the discovery that FOXP2 was relevant for human speech, many experiments were conducted on songbirds; with much of this data supporting link between this particular gene and vocal learning (Scharff & Haesler 2005, Fisher & Scharff 2009, Bolhuis et al. 2010). Indeed, such findings are perhaps the reason why one often encounters the assumption that *FOXP2* is directly responsible for human language. Thus earning it the title of the *language-gene*.

However, the human version of FOXP2, while being no more than 200,000 years old and very much relevant to human language, is certainly not the single gene responsible for all of the *Homo sapiens* language abilities (Enard et al. 2002, Hurford 2007). This is largely due to language being a trait that is too polygenic to be controlled by this gene alone, meaning that *FOXP2* could not constitute the kind of macro-mutation that Bickerton (1981) referred to. This is demonstrated in work, such as that of Krause et al. (2007), who made the particularly interesting discovery that the two mutations of *FOXP2* that were believed to be unique to modern humans were actually also present in late Neanderthals. However, the problem here is that, while this finding is somewhat enlightening, it doesn't actually tell us anything about the linguistic capabilities of Neanderthals. It is also quite likely that these amino acids were selected for in Neanderthals, but for causes quite different from that which operated upon our own species; causes that were perhaps unrelated to complex language (Balari et al. 2013).

Such an idea is supported by various bodies of work, which have demonstrated how this gene has undergone very few evolutionary changes within vertebrates, as well as displaying different orthologues in different species; such as bird song and eco-location in bats. Indeed, Konopka et al. (2009) demonstrated how the two aminoacids that differentiate the human and chimpanzee versions of *FOXP2* change the function of the gene by conferring differential transcriptional regulation in vitro. It is therefore likely that "the ability with which each *FOXP2* orthologue relates in each species is a function of the molecular context in which the protein coded by the gene integrates in each particular case, and not of minor structural modifications experienced by its different variants" (Balari et al. 2013:499).

Thus, the key to understanding the significance of *FOXP2* presence within Neanderthals would be to know whether the targets were the same; information that is still unavailable to us (Balari et al. 2013).

In short, the surest thing that can be said about *FOXP2* and its relation to human language, as research currently stands, is that it appears to have a consistent role in the early brain development of avian and mammalian species (Bolhuis et al. 2010). The role of *FOXP2* in this process does, however, lend credence to the adaptionist argument, due to the way in which mutations tend to be tiny copying errors, which typically affect a single nucleotide base, and the kind of massive macro-mutation suggested by the non-adaptionists, which change the whole DNA sequence, seldom happen all at once (Pinker 2003).

2.1.5 The Comparative Method

Modern humans are incredibly sophisticated creatures, of that there is no doubt. We are at once complicated individual psychological beings and highly social creatures that partake in various social-cultural behaviours. However, we are not the only species on the planet to display such sophistication. Recent work has started to show that many human-like cognitive processes occur in other species, such as elephants (Payne 2003, Plotnik et al. 2006), dolphins (Tyack 2003, Reiss & Marino 2001), killer whales (Whitehead 2003), and certain avian species (Bolhuis et al. 2010). Social transmission of learned behaviour, which allows for the existence of simple cultural evolution has also been documented in such animals (Whiten 2000, Rendell & Whitehead 2001, Sadedin & Paperin 2011, Whiten & van Schaik 2007).

However, despite the readiness of the scientific community to accept that a large amount of human behaviour can be explained in terms of complex cultural processes, it is largely unaccepted that cultural transmission within non-human animals is as influential on behaviour and evolution. Moreover, such processes are thought to not only be more simple and rare within other species, but also, with the exception of birdsong, deemed to not possess the stability that is necessary to make any substantial impact upon genetic evolution (Feldman & Laland 1996, Rendell & Whitehead 2001).

Although there is considerable debate in regards to what the analysis of the communicative and cognitive abilities of other species can tell us about human language, evolution is not prone to creating radically new systems with no evolutionary precursors. It is thus good scientific practice to look for homologous features that are the result of common ancestry in closely-related species, as well as analogous mechanisms that could be the result of convergent evolutionary processes. Furthermore, much of the literature on human language evolution tends to focus on the evolution of syntax, and although this is understandable given that syntactic structures are its most salient feature (Oliphant 2002), it is arguably worthwhile for a researcher to focus on the more basic features that separate human language from other communication systems seen within the animal kingdom.

This idea may be controversial for linguists, whose field has a tendency to assert that human language is so much more complex than that of any other animal communication system that none of them could be used to model the emergence and evolution of language. After all, despite certain characteristics being present in other species, it is true that human language, as a system, is exceptional and largely unique in most respects, with no ‘simpler’ analogues or homologues in other species (Bickerton 1990, Tallerman & Gibson 2013*a*).

However, it is important to remember that any system cannot be a model of any other one without some level of abstraction (Okanoya 2013). Furthermore, we are not looking for traces of, language as we know it, in these species, but mechanisms that would allow for proto-linguistic abilities in early hominins.

Indeed, despite this lack of analogues/homologues issue, the comparative method has proven to be rather fruitful in providing insights in to language evolution and its associated cognitive architecture (Fitch 2010). For example, it was traditionally believed that, unlike human language, most animal species had a closed system of vocal development, whereby they have a relatively fixed repertoire of species- and sex-specific vocalisations, that were very much situation specific. It was deemed that there was little evidence that non-human species of mammal, even non-human primates, possessed the kind of vocal plasticity that is characteristic of an open system of vocal communication. However, work done over the last few decades has built up a body of evidence which appears to indicate that such vocal characteristics are present in a number of species that are phylogenetically quite different to us *Homo sapiens*. These include the bottlenose dolphin (*Tursiops truncatus*), the harbour seal (*Phoca vitulina*), the humpback whale (*Megaptera novaeangliae*), the beluga whale (*Delphinapterus leucas*), and the horseshoe bat (*Thinolophus ferrumequinum*) (Janik & Slater 1997).

The comparative method has traditionally been employed in two ways. The first involves comparing similar traits within a clad. The wide spread use of tools within great ape species, for example, leads one to believe that such behaviour is a homologue that stems from common ancestry. It therefore follows that the last common ancestor (LCA) between the *Pan* and *Homo* genera also used tools in some form.

The second manner concerns the way in which different evolutionary lineages have evolved surprisingly similar behavioural and genetic solutions for particular functions; such as vision. This is known as evolutionary convergence, and is more formally defined as referring to functional traits that occur in species as an adaptation to similar environmental pressures; and not because of phylogenetic closeness. The archetypal example of functional convergence without structural convergence is the evolution of flight; something that has evolved in birds, bees, bats and pterosaurs.

2.1.5.1 Non-human Primates

Whether or not evolutionary precursors, or intermediates, to human language can be found in the communicative systems of non-human primates has been a subject of considerable interest; likely due to their phylogenetic closeness to ourselves. Indeed, many non-human primates share a suite of complex cognitive capabilities, which include motor control, memories, transitive reasoning, basic numerical abilities, and hierarchical understanding of social structure; all of which are relevant to language acquisition (Fitch 2010).

Research into the existence of evolutionary precursors to human language in non-human primate species tend to be dividable into two schools of thought. The first focuses on vocal modality, and believes that human language stems from the evolution of the vocal systems in our ancestors (Lieberman 1984, Seyfarth & Cheney 1984, Mithen 2000), while the second posits that human language owes its roots to the use of gestural communication (Rizzolatti & Arbib 1998, Corballis 2003, Arbib 2005, Meguerditchian et al. 2011). However, as will be seen, the data suggests that both are true, to a degree.

2.1.5.2 Non-human Primate Vocalisations

The vocalisations of non-human primates tend to be the main form of communication for these creatures. It was traditionally thought that they were largely species-specific, which was indicative of them being mostly innate. This is probably what led to many dichotomies being drawn between humans and non-human primates during the early years of primate vocal research; voluntary versus involuntary, semantic versus affective, graded versus discrete, learned versus unmodifiable, etc. However, such views are being increasingly challenged with the emergence of data that appear to contradict these initial beliefs (Seyfarth 1987, Slocombe 2013); this is largely the result of work done in three inter-related research areas.

First, there are the ‘ape-language’ projects, which have achieved some success in teaching elements of human language to chimpanzees, gorillas, and orangutans. Although these creatures are unable to control their lips and tongues in the manner necessary to produce the spoken sounds that we humans do (Slocombe 2013), teaching them to communicate with American sign language (ASL) demonstrated that apes can use a large number of signs in a manner akin to humans (Gibson 2013). This line of inquiry thus became primarily concerned with searching for grammar within the test subjects (Seyfarth 1987).

The contemporary data for this line of enquiry indicates that, while captive apes can use signs to refer to objects, which suggests that they are able to understand the relation between a sign and referent, they are highly unlikely to be able to combine such signs into anything resembling human sentences (Cartmill & Byrne 2010, Hurford 2007).

Secondly, studies of primate vocalisations have indicated parallels with the way in which human words are used. Despite initial beliefs, it has been shown that primate calls are not just reflexive responses to external stimuli. The vocalisations of chimpanzees, for example, are the result of a number of social factors, such as the composition of the audiences and the presence of alphas (Townsend et al. 2008, Laporte & Zuberbuhler 2010, Slocombe & Zuberbuhler 2007, Slocombe 2013); with orangutans even being known to partake in deceptive vocalisations (Hardus, Lameira, Van Schaik & Wich 2009). Indeed, great apes appear to be aware that others have thoughts and intentional states, and are able to use this knowledge in a number of social ways (Call & Tomasello 2008).

Furthermore, it is now known that the vocal repertoires of monkeys are both much larger and more specific than initial studies had indicated. They appear to be under volitional control, as well as effectively representing objects in their environment.

Additionally, there is now a large body of work that suggests that acoustical structures of vocalisations can be modified, likely through a process of vocal learning; such as the way in which the acoustic parameters of the 'pant-hoot' call appears to vary between different chimpanzee communities (Crockford et al. 2004). Such evidence suggests that, although the vocal repertoires of great ape species are largely fixed, some degree of novel vocalisation can be produced; an argument that is supported by a number of findings. These include the way in which Herbinger et al. (2009) used playback experiments in order to demonstrate how, when numerous groups of chimpanzees inhabit the same geographical region, individuals are able to distinguish between the 'pant-hoot' calls of members of different groups. The way in which orangutans have exhibited group-specific vocalisations (Hardus, Lameira, Singleton, Morrogh-Bernard, Knott, Ancrenaz & Wich 2009), with reports of captive orangutans replicating human whistling (Wich et al. 2009). As well as reports of captive chimpanzees producing raspberry sounds, which serve as attention-getting vocalisations; something that does not appear to occur in the wild (Hopkins et al. 2007)[1]. Additionally, it has been demonstrated that Diana and Campbell's monkeys often form mixed-species associations; with species responding to one another's alarm calls (Zuberbuhler 2000).

Indeed, non-human primates also appear to make subtle acoustic discriminations when distinguishing between calls (Seyfarth 1987, Cheney & Seyfarth 1990), and it has been shown that chorusing causes convergence. Both of which are indicative of differences at the group and individual levels having their origin in social influences (Marshall et al. 1999, McGrew 2004).

The third and final area of non-human primate vocal research that is of interest to us here is a body of work that indicates that monkeys perceive and classify objects in the external world, including conspecifics (Cheney & Seyfarth 1990, Tomasello & Call 1997). There are two main, and contradictory, stances concerning the call comprehension of these species. The first posits that they possess only the knowledge of how to best respond to a particular signal, without really understanding the associated referential context of such signals (Deacon 1997). For example, the vervet hearing a leopard call and climbing a tree because it knows that is the best thing to do when it hears such a signal. The second argues that the processing of such signals are potentially more complex,

---

[1] Two things should be noted here. First, identification of group-specific calls do depend upon their absences in other groups, something that is hard to ascertain with certainty (Slocombe 2013). Second, unlike human speech, none of these vocal behaviours engage the larynx (Fitch 2000).

perhaps more akin to the cognitive processes that underpin human language perception. Here, the signal refers an associated mental state to conspecifics. In other words, upon hearing a conspecific giving a call, an individual generates a mental representation of the current situation, and then uses this representation to guide its behaviour (Cheney & Seyfarth 1990, Zuberbuhler 2003).

Current empirical evidence adds more credibility to the latter stance. Specifically, the evidence suggests that even members of the same group will produce distinct vocalisations and that those within a social group will respond differently to calls from distinct individuals; depending on variables, such as social rank, etc (Hansen 1976, Tenaza 1976, Smith et al. 1982, Seyfarth 1987, Cheney & Seyfarth 1990). Additionally, the meanings of vocal signals vary in accordance with how signallers use them, and listeners are often unable to rely on just simple stimulus-response associations, but need to retrieve meaning through inferential processes (Zuberbuhler 2013). All of which indicates that these species are able to hold mental representations of other individuals and relate them to events.

In short, the conclusions that can be drawn from the contemporary data on vocalisations of non-human primates is perhaps best summed up by McGrew (2004):

> *"The upshot of vocal communication in chimpanzees is that it is influenced socially in ways that, if we saw similar variation in other areas of chimpanzee life, we would term it 'cultural'. On the other hand, we are reminded that no behavioural pattern exhibited by any organism develops in an environmental vacuum."*
> *- ibid:133.*

2.1.5.3 Non-vocal Communication in Non-human Primates

Communication in non-human primates is not just vocal, it occurs in all major modalities, and a prominent one is the use of gestural communication. The literature on gestural communication is incredibly thin compared to that of vocal communication. Perhaps due to the increased difficulty in exploring this phenomenon in the wild (Slocombe 2013).

It is generally accepted that primates use a range of facial and body movements as signals, such as the chest beating of gorillas. Indeed, gestural communication is present in all great ape species, leading a number of human language researchers to argue that early hominins used manual gestures in a linguistic capacity prior to the emergence of speech (Arbib et al. 2008, Corballis 2013, de Waal & Pollick 2013).

The study of gesture use has predominantly been carried out on great apes, where it appears to function primarily to facilitate ongoing social interactions, or to manipulate behaviour in a desired way (Zuberbuhler 2012). Although there is very little evidence of such gestural communication carrying any form of symbolic meaning (Cartmill & Byrne 2010), given that apes can take each other's attentional state into account so as to better deploy signals (Tomasello & Call 1997), it is possible that behaviours such as pointing were the origin of referential communication (Hurford 2007). This would make intuitive sense, given how this ability arises quickly during ontogenetic growth in humans (Povinelli et al. 2000), and how it has been shown that both wild and captive apes are able to mimic gestures made by humans and conspecifics (de Waal & Pollick 2013).

Corballis (2003) takes such data as supporting the *gestural-origins-of-language hypothesis*, and argues that that there is sufficient evidence to support the notion that the evolution of human speech, and even language itself, is unrelated to animal communication systems, and that it instead emerged from manual gestures. This argument stems from the way in which gesture usage in apes tends to be very much culture-specific (de Waal & Pollick 2013); in a manner similar to that seen in our own species. Corballis (2002, 2003) suggests that the flexibility seen in gestural communication indicates that this was perhaps the communicative channel through which our hominin ancestors developed symbolic meaning, perhaps along with referential vocalisations; an ability that arose with bipedalism.

However, Corballis (2003) is eager to point out that it might not have been language *per se* that characterised the emergence of our species. Moreover, it was autonomous speech, in the form of a combination of speech and gestures; not unlike Bickerton's protolanguage. This is supported by the fact that apes possess greater control over their gestural communication than they do the production of other signals (de Waal & Pollick 2013). Corballis (2003) goes on to argue that facial gestures were also crucial to this process, and that gestures were punctuated with vocalisations and other sounds, such as the teeth chattering seen in chimpanzees; which became increasingly associated with facial gestures, which then led to considerable modifications to the tongue and vocal tract.

Further support for the manual-gesture-origin-of-language hypothesis came from Rizzolatti and Arbib's evolutionary framework for 'language readiness', which concerns the brain mechanisms for action and perception; mechanisms that are shared by both human and non-human primates.

Referred to as the *mirror system hypothesis* (MSH) (Rizzolatti et al. 1995, Rizzolatti & Arbib 1998, Arbib 2010, 2005), this line of work stems from the discovery that the rostral part of the ventral premotor cortex (area F5) in monkeys contains neurons that discharge when a monkey either grasps an object or witnesses a human experimenter grabbing an object. This is of particular interest, as it is largely believed that this part of the brain is the monkey homolog of Broca's area in humans; an area that is closely related to speech production.

Thus, it has been concluded that the specific function offered by the human brain's Broca's area developed from ancient mechanisms that were related to the ability to perform and perceive motor actions (Rizzolatti et al. 1995, Wilkins & Wakefield 1995). Therefore leading Rizzolatti & Arbib (1998) to argue that language stems from communication methods as an integrated whole, including manual gestures and facial expressions. It is important to note, however, that the authors are not positing that the mirror system evolved for communicative purposes. Moreover, they argue that it provided the ability to recognise and generate actions, which was crucial for *language parity*; meaning that what counts for a speaker would also count for the listener (Arbib 2010, 2005).

In short, gestural communication of the kind seen in other ape species is a modality that has the strong potential to have acquired symbolic meaning in early hominins (de Waal & Pollick 2013), and the discontinuity between humans and other primate species regarding the gestural modality is indicative of a fairly recent swift towards more intentional and flexible communicative behaviours.

2.1.5.4 Evidence Of Primate Syntax

We have seen that non-human primates do produce vocalisations that consist of a combination of calls, that are meaningful to others in context-specific ways. This is seen in chimpanzees (Crockford & Boesch 2005), bonobos (Clay & Zuberbuhler 2009), gibbons (Clarke et al. 2006), Campbell's (Zuberbuhler 2002), and Diana monkeys (Stephan & Zuberbuhler 2008). Although, it is unclear in the case of apes as to whether or not different sequences carry any specific meaning (Zuberbuhler 2013), a study has recently indicated that the chimpanzees of the Tai forest produce many call sequences that are a combination of other calls (Crockford & Boesch 2005). Indeed, combining calls appears to be an important part of chimpanzee vocal behaviour (Slocombe 2013). However, there does not appear to be any form of re-ordering in order to convey different meanings, as was demonstrated by the aforementioned 'ape-language' experiments.

Despite this, Seyfarth & Cheney (2013) argue that, although these aspects of non-human primate social knowledge are by no means human, they do resemble the meanings that are expressed in human language; which are built by combining discrete-valued entities in a structured, hierarchical, rule-governed, and open-ended manner. This leads one to the belief that our internal representations, that we utilise for language meaning were built upon the knowledge of social relations possessed by our pre-linguistic ancestors (Worden 1998, Bickerton 2000, Seyfarth & Cheney 2013).

The idea here is that a precursor to the hominin mind evolved in a social milieu that was characterised by various social challenges, and that these, in turn, created selection pressures that favoured structured, rule-governed intelligence (Seyfarth & Cheney 2013). Indeed, as Studdert-Kennedy (2013) observes, this need for understanding social life requires discrete, compositional thinking; making such a cognitive area a prime candidate for the origin of compositional structure in spoken language.

2.1.5.5 Avian Species

We refer to a particular class of avian vocalisations as ‘birdsong’, not just due to the way in which the sound patterns are organised, but because our musical aesthetics naturally create such an analogy (Hartshorne 1973). Indeed, despite diverging from the lineage that eventually led to modern humans roughly 280 million years ago, avian species can provide interesting models for the evolution of vocal communication (Pepperberg 2013); and, like humans, birds have a huge repertoire of sounds (Slater 2013).

These ideas were largely pioneered by Peter Marler, who investigated the nature of structure-function relationships within avian alarm calls, and found that they were non-arbitrary (Marler 1957). In a seminal paper, Marler (1970) went on to note a number of parallels between birdsong and human speech:

1. Juveniles learn the species-typical repertoire from adult models.

2. Dialects are formed as a result of learning.

3. Experimentally guided learning is most significant during a critical period.

4. To develop a normal vocal repertoire, the young must be able to hear sounds from their species-typical repertoire and to hear themselves reproduce such sounds.

5. Like human infants, young birds also go through a series of developmental stages, including a sub-song phase that resembles babbling.

6. Vocal imitation, in and of itself, may be self-reinforcing.

7. The left hemisphere is dominant for the control of sound production.

Since then, many studies have now been conducted into the development of songs in avian species, all of which tend to paint a similar overall picture. Young males being isolated from other members of their species will develop more simplistic songs that lack the detailed structure and phrasing seen in normal adult song. However, if they are raised in an environment where adult males are present, or where there are audio recordings of adult male songs, then they will reproduce the songs of the adult males recordings. Sometimes, however, slight errors are made in song copy; and this tends to lead to the emergence of new song types, which can lead to regional differences in songs (Catchpole & Slater 1995, Slater 2013).

Birdsong is of particular interest to human language researchers, given how it conveys the motivational, cultural, and ontogenetic conditions of the signallers, and is often referred to as having syntax, due to the way in which the notes of the song are ordered according to rules. However, these notes do not have any referential meanings; hence, birdsong does not have syntax, merely the ordering of the elements (Hauser 1996, Tallerman & Gibson 2013*a*). Thus, given that birdsong is a graded or holistic system of communication and human language is compositional (Hauser 1996, Okanoya 2013), one cannot directly compare the two. Despite this, the study of avian vocalisations can still provide a potentially important biological model for the study of human language; a model that looks at the interactions between culture and heredity, due to the way in which it is learned and has social functions, as well as depending upon a certain level of innateness to acquire it (Bolhuis et al. 2010, Okanoya 2013).

Indeed, it is now commonly accepted that avian vocalisations are learned, to differing degrees, in three avian species; parrots, hummingbirds, and passerines, which make up around 5,000 species (Okanoya 2013, Slater 2013). Initial data supporting this belief came from the way in which, without exposure to a tutor, a juvenile produces an abnormal so-called *isolate song*; which demonstrates an amount of predisposition, in the way in which they retain some species-specific features (Bolhuis et al. 2010).

Feher et al. (2009), for example, found that upon tutoring zebra finches (*Taeniopygia guttata*) with such an isolate song, they produced a fairly accurate representation of the song; but also 'improved' the song to match the species-specific features. Then, using each generation's newly improved song to tutor next generation, it was found that, with each subsequent generation, the song more closely resembled, and eventually became indistinguishable from, the typical zebra finch song.

It has been observed that these findings demonstrate that the neural substrates for learned vocalisations can apparently carry quite exact pre-specifications for particular acoustic features and an overall acoustic 'gestalt', yet allow a high degree of plasticity for what can be learned within those constraints; in a manner similar to the way in which Bickerton (1981, 2000) describes the emergence of pidgins and creoles.

This is particularly interesting, given how the architecture and connectivity of avian and mammalian brains are much more similar than originally thought. For instance, "avian pallial 'song' regions bear functional similarities with human auditory and motor cortices and the importance of the basal ganglia for both speech and birdsong is starting to be understood mechanistically" (Bolhuis et al., 2010:747). Suggesting that language evolved within a biological substrate which is largely shared with other animals.

As mentioned above, another similarity between human and avian vocal learning is the way in which both humans and songbirds exhibit a *sensitivity period* during early ontogenetic development; the point in life history at which vocal learning is best achieved. In addition, the production phase of vocal learning is preceded by a listening phase in both groups (Bolhuis et al. 2010). Indeed, there are striking parallels between the way infants learn to speak and birds learn to sing. In that both involve the imitation of species-specific communicative sounds (Doupe & Kuhl 1999, Bolhuis et al. 2010); much like humans learning to speak.

2.1.5.6 Cetaceans

Out at sea, one can find many examples of social, cognitive, and cultural co-evolution taking place; in ways that parallel those seen in the above terrestrial creatures. Despite the fact that they diverged from primates over 65 million years ago, cetaceans display sophisticated communicative abilities, complex fission/fusion social structures, and advanced abilities in cognition and social learning (Rendell & Whitehead 2001, Whitehead 2003) All of which suggest that complex communication is not limited to just primate species (Janik 2009, Zuberbuhler 2012). This is fascinating given how these creatures inhabit an environment where there is “little potential for technology, no refuges against predators, few barriers and cheap movement, excellent acoustic propagation, and high biological variability over medium to large temporal and spatial scales” (Whitehead, 2003:461-462); in stark contrast to life on land.

Of all cetaceans studied, dolphins are perhaps the most interesting to those studying human language, due to the way in which they not only develop distinctive calls that function in many ways like the contact calls of parrots (Bradbury 2003), but because the process of learning to match calls appears to be very strongly linked to the development of social bonds (Tyack 2003).

Furthermore, it appears as though bottlenose dolphins are able to create abstract mental representations of objects and actions in order to guide their own behaviour (Herman et al. 1994); much like we see in primate societies. Indeed, Kako (1999) argued that the comprehension abilities of dolphins were akin to that seen in language-trained apes, in that they were able to understand the referential nature of artificial signals (Herman & Forestell 1985, Janik 2013), and pointing by human trainers (Herman et al. 1999).

This indicates that they are able to form concepts (Janik 2013), as well as being capable of other forms of social learning other than just that of a vocal nature; including imitation (Herman 2002), self-recognition (Reiss & Marino 2001), and emulation (Tayler & Saayman 1973). However, as Gibson (2013) observes, there is no evidence that suggests that wild dolphins communicate about absent animals, food, or events; i.e. exhibit displacement.

2.2 Conclusion

The fact that evolution does not make leaps and bounds, but only develops what already exists, makes it beneficial to look for insights from other species when investigating a natural phenomenon such as human language.

From the above discussions, it can be seen that the existing data indicates that strategic learning behaviours have evolved in a range of taxa, with strikingly similar context-specific patterns of copying to those observed in humans. This suggests that the evolution of such copying behaviour is best regarded as a convergent response to specific selection pressures, and might not be well predicted by the relatedness of a specific species to our own (Rendell et al. 2011). Furthermore, the presence of such abilities in non-human primates suggests that mental representations are phylogenetically older than language, which allows us to argue against the idea that language evolved to allow abstract thought. Although, it may well have been a driving force behind this capacity becoming increasingly abstract and more influential as an evolutionary selection pressure.

Possession of mental representations, vocal learning, social learning/feedback, and a high level of self-awareness and understanding of conspecifics are arguably the basic prerequisites for any advanced communication system to emerge. We have seen that mental representations are present in a wide range of animals, which appear to use them to guide their own behaviour (Herman et al. 1994), and this is likely to have created a selection pressure for high-levels of self-awareness to co-evolve with these internal representations. This, in turn, perhaps led to these skills being co-opted to allow an individual to interact with others. The existence of these abilities in apes also suggests that these abilities were in place prior to the *Pan*/*Homo* genera split, and therefore in the last common ancestor (LCA) (Calvin 2004).

Indeed, all species of great ape appear to possess sufficient mental and communicative capacities to enable them to use a form of protolanguage; and yet, none do in the wild (Gibson 2013, Knight & Power 2013, Cann 2013). Further, the fact that all great ape species appear to be able to master the essential components of protolanguage indicates that the required cognitive capacities were also present in the LCA (Steiper & Young 2006, Gibson 2013). Thus, although we can safely conclude that language is a hominin innovation, and that australopithecines and early *homo* had taken strides towards human-like adaptations (Tallerman & Gibson, 2013*a*:27), such as bipedalism (Bickerton 1981, MacLarnon 2013), it is unknown whether any of the known hominin species in the fossil record possessed any form of proto-linguistic speech abilities approaching our own.

The fact that there appears to be little in terms of evolutionary precursors to human language does tend to cause issues for those studying it, a problem that is exacerbated by the fact that human language is not a monolithic entity, but a complex bundle of traits that have emerged and developed over evolutionary time.

It therefore makes sense that no single research methodology would be able to explain the nature of its existence to a satisfactory standard. Moreover, it is perhaps more beneficial to treat each of these different types of data as pieces in a puzzle; one that is slowly being filled in, but has a long way to go before it is completed (Fitch 2010). Nevertheless, as a result of this, the literature on human language is vast, often contradictory, and potentially bewildering to the casual observer. This chapter has aimed to give the reader a brief overview of the parts of the literature that are most relevant for the central tenets of this thesis.

The first half of the chapter was very much an historical overview, which demonstrated how certain ideas to be explored later have a long pedigree in human language research. Most importantly, later chapters will explore the role of individual learner bias in the evolution of linguistic systems, meaning that languages themselves evolve to be easier to learn by language users; an idea that other researchers have touched upon before. Indeed, we have seen that Chomsky (1980) came to the conclusion that language users must possess a set of rules or procedures for comprehending when various parts of speech can occur within and among sentences. He argued that these rules and procedures adhere to the intuitions of said language users; a process that can be seen to allude to such an idea. It has also been seen that Bickerton (1990) believed language to be a phenomenon that produced output that was the result of the input given to the language users; an idea he extended to the evolution of pidgins and creoles.

The second half of this chapter took a brief look at the different research methodologies that have been utilised in human language research. The aim here was to provide the reader with the necessary background knowledge to understand the work to be presented, as well as to enable them to see as to where this work should be placed within the wider literature. More specifically, we have seen that many different methodologies have been utilised in human language research, and these have offered many valuable insights.

However, there is a methodology that was not discussed here. Namely, modelling methodologies, which use mathematical equations and computational simulations in order to model the dynamics of either linguistic systems or individuals that attempt to communicate with one another. In doing so, they aim to understand the dynamics that underpin language evolution and change. This will be the focus of the next chapter.

Chapter 3

Modelling Language Evolution

3.1 Introduction

In recent decades, with the arrival of more powerful desktop computers, research into naturally occurring communication systems has benefitted from computer-based simulations. Indeed, exploring the origin and nature of language, and communication systems in general, involves understanding the nature, behaviour, and complex interactions between large numbers of variables within a dynamical system. It is here that computer-based models can aid one in overcoming problems provided by using either verbal theorising or mathematical models in isolation; largely due to the way in which they provide one with a middle ground between abstract theorising and rigorous mathematical approaches. The motivation behind such computational-based work is the belief that language is like a traffic jam, in that it exhibits the properties that it does as a result of the interaction between many individual elements (Kirby 2013).

Given the growth of computational research into human language over the last two decades, it is unsurprising that there have been numerous attempts to classify the different models seen in the literature into various categories, so as to allow comparisons and contrasts to be drawn (Perfors 2002, Kaplan et al. 2008, Vogt 2009, Jaeger et al. 2009, Grifoni et al. 2015). Indeed, a number of different approaches have been used in attempting to do so, such as the theoretical stance or the methodology adopted by the authors. The scheme presented here results from a merge of those presented by Kaplan et al. (2008) and Grifoni et al. (2015). This overview of models will begin brief and grow in detail the closer we get to the focus of this thesis.

3.2 Grammatical Representations

The literature on human language modelling contains a wealth of different modelling paradigms that are concerned with representing linguistic features, with a particular focus on how grammars are used to represent the symbolic units, semantics, and syntax seen in human language. These modelling grammars come in various flavours, which have been implemented in a number of different models[1].

3.2.1 Universal Grammars

First and foremost, Universal Grammars (UGs) are theoretical grammatical systems that have been used in various language evolution models. As we have seen, this idea was first proposed by Chomsky (1980), as part of his investigations into the *poverty of the stimulus*, as a search space upon which natural selection allows for the emergence of syntactic structures. As Grifoni et al. (2015) observes, while UG helps the learner to generalise rules and allow for creative language use, it does not take the different social-environmental factors of learners into account; just those of ideal learners.

3.2.2 Context-free Grammars

Context-free grammars (CFGs) are perhaps the oldest, and certainly most common, type of grammar formalism found within the computational literature. They are essentially a set of recursive re-writing rules that can be used in order to generate patterns of strings; where the expansion of a symbol does not depend on its context.

Their theoretical foundation stems from the generative stance that dominated linguistics during the latter half of the 20th century, which deemed the learning of syntax to be the central problem in the acquisition of language. Thus, CFGs are often employed for work that involves learning a mapping between the space of a sentence and the binary space code for syntactic correctness/incorrectness (Kaplan et al. 2008).

In being used in such a manner, this formalism lends weight to the aforementioned *poverty of the stimulus* hypothesis, which claims that a child's linguistic input underdetermines what they learn about language. Thus the need for some form of genetic pre-programming of specific linguistic neural circuits (Pinker & Bloom 1990, Pinker 1994).

CFGs have a number of advantages in that they are computationally tractable, making it easy to determine whether or not a produced string is grammatically correct, they are easily interpretable by both human beings and computers. Most importantly, however,

---

[1] The interested reader is advised to consult Jaeger et al. (2009) or Grifoni et al. (2015) for a more detailed discussion.

they are able to capture all linguistic phenomena of human languages while assuring a low parsing complexity. However, CFGs do not provide any semantic information while carrying out derivation; meaning that they cannot express sentences which are context-dependent (Grifoni et al. 2015).

De Pauw (2003) presents an agent-based evolutionary computing technique as a machine-learning method for data-driven grammar development. Referred to as GRAEL, this system is geared towards the induction and optimisation of natural language grammars. Here, a population of agents hold a set of idiosyncratic linguistic structures, and use a CFG formalism in order to both compose their own sentences and to analyse those of other agents. The evolution of the agent's grammatical knowledge then occurs as a result of a series of interactions between agents, with the aid of error-driven learning, and it is shown that such a framework is able to optimise any type of grammar, regardless of its initial starting quality.

3.2.3 Attribute Grammars

Attribute Grammars (AG) were developed with the aim of addressing the aforementioned lack of semantic information and increase the context-dependency seen in CFGs (Boyland 1996). In short, AGs are essentially CFGs, but with an added method that allows for context-sensitive derivation. Thus allowing them to encode the constraints on language evolution as part of the production rule specification, as well as allowing it to encode semantic information into attributes (Grifoni et al. 2015). Further advantages come from the way in which AGs allow new production rules to be added without altering other rules (often referred to as modularity) and the way in which the production rules do not have to be applied in any specific order (Boyland 1996).

Juergen & Pizka (2006) use AGs as a specification formalism for both syntax and semantics in presenting a tool for exploring the nature of the evolution of domain specific languages. This is done by looking at how making evolutionary changes to domain specific languages cheaper, by way of automating the adaptation of its parser, impacts the domain specific languages.

3.2.4 Christiansen Grammars

Christiansen Grammars (CG) are in turn an extension of AGs, and are adaptive grammars that are capable of altering their own production rules (Christiansen 1990). This is achieved by having the first attribute of every symbol take the form of a CG, which defines the production rules that can be applied when expanding or deleting rules. While CGs allow for the creation of new linguistic constructs, as well as allowing the grammar to be modified while its being used, it does only possess a small set of primitive constructs that are combined in a small number of ways, so as to build the linguistic structures (Grifoni et al. 2015).

Ortega et al. (2007) present an evolutionary algorithm, which they refer to as *Christiansen grammar evolution* (CGE), that extends standard grammar evolution by replacing CFGs with CGs. In doing so, they added semantics so as to ensure that individuals are created that are semantically and syntactically valid. Their belief here is that, given how grammar evolution is a general purpose stochastic search technique that utilises CFGs so as to improve syntactic performance, adding semantics should improve performance even further.

The results presented by Ortega et al. (2007) not only support this notion, but also demonstrates how this improvement does not actually depend upon the context-dependent nature of constraints, but on the ease of the formalism used to express them.

3.2.5 Fluid Construction Grammar

Inspired by Construction Grammars, Fluid Construction Grammar (FCG) was primarily designed to allow researchers to formally write down the inventory of lexical and grammatical constructions necessary in parsing and producing utterances. This was done with the aim of capturing the cognitive and interactional foundations of language and present it within a grammatical model (Steels 2011*a*,*b*).

In short, FCGs are a form of construction grammar that offer a way to structure and represent semantic meaning into patterns and named constructions. FCGs have two main advantages over those above, in that they are able to use the same construction for both parsing and producing without compromising efficiency, as well as offering flexibility in usage; meaning that sentences can be understood even if they are not completely grammatical. They do, however, demonstrate a lack of matching between syntactic structures and semantic categories (Grifoni et al. 2015).

FCG was used by Steels & Garcia-Casademont (2015) in order to explore phrase structure and its relation to the origins of syntax. In doing so, they present a minimal language game whereby grammar is needed in order to avoid semantic and syntactic uncertainty. This model demonstrated how agents were able to give rise to phrase structure grammar during the course of their interactions as a result of collective invention, adoption, and alignment of culturally established conventions.

3.3 Game Theoretic Modelling

Not all models of language evolution involve exploring specific linguistic representations. Game theoretic models explore the nature of language evolution in regards to the search for optimal behaviour within populations of interacting individuals. Game theory is a method of modelling rational decision making during interactions between individuals with the use of mathematical equations. Here, all individuals involved in an interaction, or game, receive some kind of payoff; a payoff that is the direct result of the actions taken by said individuals during the course of the game.

When exploring language evolution, such a game tends to take the form of a communicative interaction between a speaker and hearer; whose actions take the form of the production and interpretation of an utterance, respectively. The maximum payoff for the players being the communicative strategy that maximises communicative accuracy.

More specifically, such a game theoretic model involves meanings being presented to a speaker in accordance with some probability distribution, P, which represents cognitive and communicative tendencies, and not the peculiarities of a given language. It is then assumed that there are a fixed set of meanings, M, and utterances, U. A speaker, S, is then any strategy or function that produces a mapping from M to U; what is often referred to as a production grammar. Thus, in a single game, a speaker is presented with a single meaning m (where $m \in M$), which it has to convey to the hearer using an utterance, u (where $u \in U$). The interaction is then deemed successful if the hearer recovers the intended m from the presented u.

3.3.1 Jager

Jager (2007) uses just this approach in order to explore the patterns, in terms of accusative and ergative systems, within the typology of case marking across the languages of the world. In other words, he explores why certain features of language are typologically common and stable across various languages.

Here, a given grammar is considered to be better than another grammar if it increases communicative success while minimising speaker effort. Jager (2007) makes the simplifying assumption that the complexity of such utterances can be measured numerically as its cost, which is a function from U to non-negative real numbers. Thus, the speaker wants to achieve the highest level of communicative accuracy while simultaneously minimising the complexity of u; which is captured by the utility function. The hearer then wants to recover the intended m as accurately as possible.

This model demonstrates that, while the majority of languages are in a stochastically stable state, whereby stability is a gradient notion and the empirical predictions of an evolutionary approach are necessarily probabilistic instead of categorical, there are some exceptions. In short, Jager (2007) is able to use a game theoretic model to predict what type of harmonic alignment patterns seen in languages are typologically dominant.

3.3.2 Niyogi & Berwick

A different flavour of such work comes from Niyogi & Berwick (1997), who present a model of language change that is derived from a model of language acquisition; whereby language users acquire a particular grammar from a class of grammars. This model explores how a population of such language-users give rise to global linguistic characteristics that develop over multiple generations. More specifically, it is shown that any triple of a family of target grammars, G, acquisition algorithm used by language-learners to hypothesise a grammar based on linguistic input, A, and a distribution that dictates the sentences presented to the language-learner, P, can be transformed into a dynamical system that can capture the evolving linguistic composition of a population that uses memoryless learning algorithms and parameterised grammatical theories.

The key point of this work is that such models are logical consequences of grammatical and learning theories; meaning that whenever one is proposed, a specific evolutionary framework is also being proposed.

Specifically, Niyogi & Berwick (1997) use this approach to propose a diachronic criterion for grammatical theories. Building upon other work, this enables them to argue that G and A can be reduced to a dynamical system, the evolution of which should match the real-life evolution of human language. This is demonstrated by looking at the historical loss of ‘verb second’ during the transition from old to modern French.

3.3.3 Nowak

In a similar vein, Nowak (2000) presents a model whereby agents are represented as a n x m association matrix, A; where n is the number of objects and m the number of signals. The entries on the matrix, a_{ij}, are non-negative real-numbers and represent the strength of the association between an object, i, and signal, j. Here, it is demonstrated how errors made during language learning lead to evolutionary change and improved information transfer. It was also found that maximum fitness is achieved in a system where agents use a small number of signals to refer to a small number of objects. This constitutes an error limit that can be overcome with word formation.

In a later work, Nowak et al. (2001) proposed a similar model, but with the aim of exploring the evolution of UG. Here, each agent has an idiosyncratic grammar. Sometimes, however, a grammar produced by the speaker's grammar, G_i, may not be interpretable by the hearer's grammar, G_j. The probability of this being the case, A_{ij}, being $0 \leq a_{ij} \leq 1$ and a_{ii}=1. Pay-off utilities for an agent using a specific grammar is then gained from the reproductive success of those using said grammar. Children in this model acquire the grammars of their respective parents.

The authors go on to show how, in this model, the adoption of a single dominant grammar by the entirety of the population is an incredibly rare occurrence, and will only occur if children learn the grammars of their parents perfectly.

3.3.4 Mitchener

Similarly, Mitchener (2007) explores the case of two genetic variants of UGs, where it is assumed that each allows for two possible grammars. It is shown by way of a communication game that, if the pay-off matrix obeys certain constraints, then genetically homogeneous populations are evolutionary stable. More specifically, it is shown that, due to substantial differences in the languages admitted by the two UGs, they are stable against invasion from one another.

Thus, the benefits of communicating with the rest of the population limit linguistic evolution to innovations that are fairly compatible with the current UG, meaning that some beneficial mutations will die out before their advantage is realised.

Additionally, it is shown that, for very similar UGs, evolutionary stability cannot be assured by the payoff matrix constraints alone, and that the various UGs can determine the outcome. It is also shown that it is hard to define the fitness of one UG in regards to another, and such a definition would have to be parameterised by the linguistic environment. UGs are meta-structures, meaning that they are one level removed from the payoff, indicating that the particular mixture of languages present in a population may dictate whether a mutation spreads or dies out.

In other words, the results presented by Mitchener (2007) indicate that the genetic history of languages are constrained by the need for compatibility with the current UG, and that mutations in the human language faculty may have spread or died out in accordance with historical accident rather than as a result of any simple notion of relative fitness.

3.3.5 Oliphant & Batali

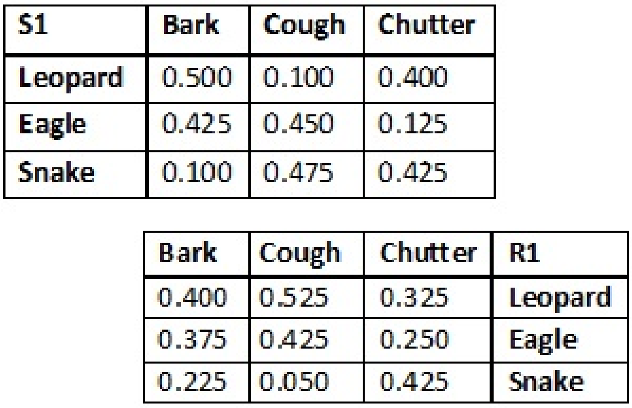
Another model that will prove to be of interest in the next chapter was originally presented by Oliphant & Batali (1997). Here, each speaker assumes that the hearer's internal mapping between signals and meanings is similar to its own. Consequently, when choosing which signal to make for a particular meaning, the speaker will choose the signal that, if presented as input to themselves, would cause them to infer the meaning currently under consideration; the authors termed this the *obverter procedure*.

Oliphant & Batali (1997) prove that individuals using the obverter will tend to improve their communicative accuracy over time until an optimal communication system is achieved. Since the space of signals is finite and relatively small, this type of mechanism is feasible in the model.

Oliphant & Batali (1997) use Cheney & Seyfarth's (1990) work on the vervet monkeys (*Chlorocebus pygerythrus*) as a backdrop in exploring how a coordinated communication system can emerge from learning within a population. Here, it is assumed that individuals within a population are able to recognise different types of predators, specifically leopards, eagles, and pythons, in such a way that there is a distinct and correct response to each; the authors refer to these as 'meanings'. Oliphant & Batali (1997) also assume that individuals have a distinct set of actions that they are able to perform, at no cost to themselves, which are recognisable to conspecifics; these are referred to as 'signals'.

Oliphant & Batali (1997) characterise the communicative dispositions of the individuals with two probability functions; a 'send' and a 'receive' function. The former gives the average probability that an individual will send the indicated signal for the meaning in question; giving the 'leopard' alarm call when it spots a leopard, for example. The latter, gives the probability that another individual will interpret a given signal as the intended meaning.

Within a send and receive function, the probability that a signal, σ, is sent for a meaning, μ, by a send function, s, is represented as $s(\mu, \sigma)$. Conversely, the probability that the corresponding receive function, r, interprets a signal, σ, as meaning, μ, is represented as $r(\sigma, \mu)$.



S1	Bark	Cough	Chutter
Leopard	0.500	0.100	0.400
Eagle	0.425	0.450	0.125
Snake	0.100	0.475	0.425

Bark	Cough	Chutter	R1
0.400	0.525	0.325	Leopard
0.375	0.425	0.250	Eagle
0.225	0.050	0.425	Snake

Table 3.1: An example send and receive function replicated from Oliphant & Batali (1997). In send function S1, each entry gives the average probability that an individual will send the indicated signal for that meaning. While the receive function, R1, gives the average probability that the member of the population will interpret a given signal as a given meaning.

For the send and receive function depicted in table 3.1, while the cough signal is most likely to be selected to indicate the presence of a snake, with a probability of 0.475, the chances of this signal being interpreted correctly is only 0.050.

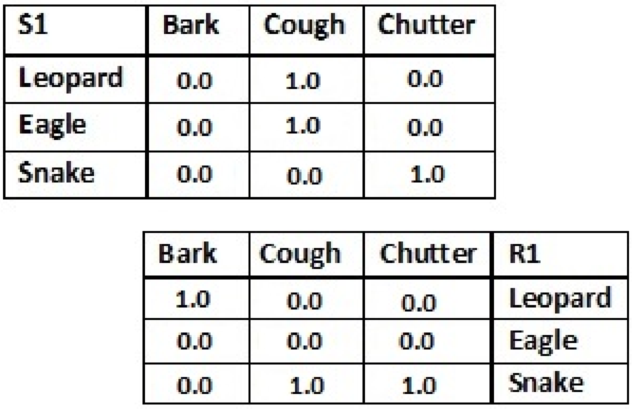
This led Oliphant & Batali (1997) to the observation that, if one's signal is to be interpreted correctly, then the signaller should send, for each meaning, the signal that it would itself most likely interpret as that meaning. Thus, in order to maximise the probability that one will correctly interpret the signals sent by conspecifics, one should interpret a given signal as the meaning it most often encodes. In this example therefore, the speaker would achieve greater communicative success if it sent a chutter signal in order to draw attention to a snake.

This *obverter procedure* can be expressed more formally in the following manner:

For each meaning μ:
s.1: Find the signal k_μ for which $r(k_\mu, \mu)$ is maximum.
s.2: Set $s_{ob}(\mu, k_\mu) = 1.0$, and set $s_{ob}(\mu, \sigma) = 0$ for all $\sigma \neq k_\mu$

For each signal σ:
r.1: Find the meaning η_σ for which $s(\eta, \sigma)$ is maximum.
r.2: Set $r_{ob}(\sigma, \eta_\sigma) = 1.0$, and set $r_{ob}(\sigma, \mu) = 0$ for all $\mu \neq \eta_\sigma$

Oliphant & Batali (1997) demonstrated how the obverter procedure yields the highest possible level of communicative accuracy within a given population when applied to its send and receive functions. Take, for example, the send and receive functions in table 3.1, if one were to apply obverter-based learning to these, it would yield the set of send and receive functions in table 3.2:



S1	Bark	Cough	Chutter
Leopard	0.0	1.0	0.0
Eagle	0.0	1.0	0.0
Snake	0.0	0.0	1.0

Bark	Cough	Chutter	R1
1.0	0.0	0.0	Leopard
0.0	0.0	0.0	Eagle
0.0	1.0	1.0	Snake

Table 3.2: The send and receive functions for a population, whose average communicative accuracy was depicted in figure 3.1, after application of the obverter learning procedure. Replicated from (Oliphant & Batali 1997).

Oliphant & Batali (1997) use mathematical proofs to show that, unless the population in question already has an optimal communication system, individuals using obverter will acquire a send and receive function with a communicative accuracy higher than the average accuracy. Indeed, Oliphant & Batali (1997) go on to demonstrate that, with obverter-based learning, the average communicative accuracy of the population will steadily increase until an optimally coordinated communication system emerges.

3.4 Evolutionary-based Computational Models

Evolutionary computation stems from the field of artificial intelligence, and tends to focus on simplifying the representation of the four evolutionary mechanisms of selection, reproduction, mutation, and crossover, in order to solve problems of optimisation.

These kind of computational models arose as a result of the need to simplify mathematical models that grew in complexity as more complex systems were being investigated (Grifoni et al. 2015). An example of such an evolutionary algorithm can be seen in figure 3.1.

3.4.1 Genetic Algorithms

Genetic algorithms involve a population of potential solutions to an optimisation problem, and tend to function in the manner depicted in figure 3.1. Here, individuals take the form of fixed-length bit strings, which in the case of language evolution models, represent linguistic structures which then evolve.

In the aforementioned work of De Pauw (2003), for example, a genetic algorithm was implemented whereby individuals were constrained by a fitness function in order to explore the content and quality of grammars that evolved over multiple generations.

3.4.2 Evolving Grammar Models

Evolutionary models of grammar add some form of semantics into the grammatical process, allowing for only syntactically and semantically correct languages to be generated. Thus avoiding a generational time step of syntactic mistakes (Ortega et al. 2007, Grifoni et al. 2015).

Here, the grammar evolves itself, which allows for the automatic incorporation of new features into the language. Such models tend to follow an algorithm similar to that depicted in figure 3.1, with an additional key feature in the form of a grammatical representation that constrains and biases the search space, and which encodes domain knowledge (Grifoni et al. 2015). An example of grammatical evolution can be seen in Juergen & Pizka (2006), who use a three-step evolutionary process in order to evolve domain specific languages.

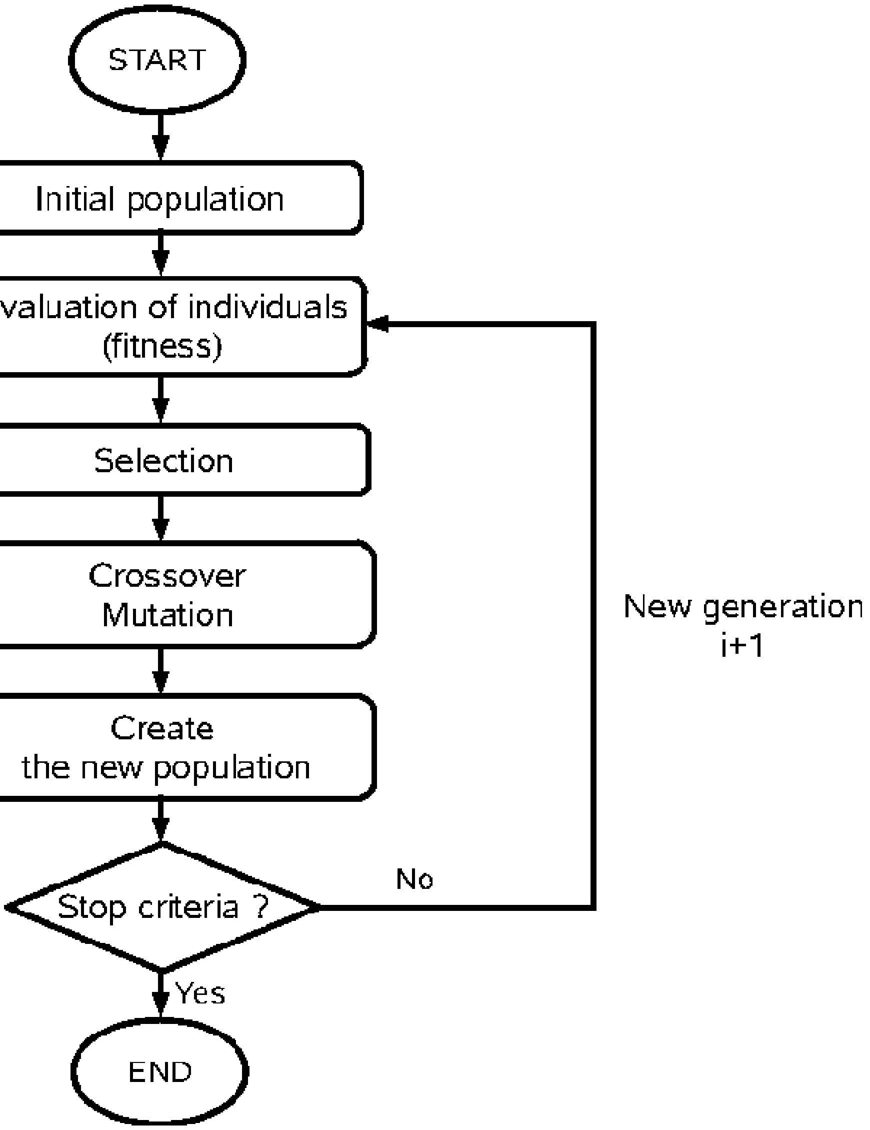


FIGURE 3.1: A basic evolutionary algorithm.

3.5 Agent-based Models

Agent-based models (ABMs) involve creating a population of agents, who are equipped with various cognitive, linguistic, and social behaviours, and allowing them to interact with one another in accordance with particular principles of a model. This then allows language evolution researchers to explore the origins and evolution of language from a 'bottom-up' perspective; as an emergent phenomenon of the interactions between the agents. ABMs will be the focus of this thesis, and given how the literature on them is already quite vast and is still growing, we will focus the next few pages on a number of notable models.

3.5.1 Evolution Of Communication

Some of the earliest ABM models focused on the evolution of natural, non-linguistic, communication; with a particular focus on the role of cooperation.

3.5.2 Werner & Dyer

One such model is that of Werner & Dyer (1992). The premise here was to create an environment where there was a selection pressure placed upon the agents within it to communicate. Then, as the environment became increasingly complex, progressively more interesting and complex communication systems would arise.

Werner & Dyer (1992) presented a grid-world environment, consisting of roughly 40,000 possible locations, containing both male and female agents, which are represented as an 8-bit genome of integer values that correspond to the connection strengths of each unit of a recurrent neural network that is capable of feedback. Females can see the males and utter sounds in order to gain their attention, while the males are blind, but can hear the utterances of the females. The males must then attempt to make their way to the females in order to reproduce; creating two offspring, one male and one female. Thus, only those that are good at communicating will reproduce into the next generation.

During simulation runs, the agents initially behave randomly, but after a time, the males begin to develop strategies. Those who stand still were not selected for, while those that continuously moved in a straight line, thereby maximising the area covered, were selected for. After a certain number of males incorporated this strategy, the males began listening to the females due to the random walks becoming the non-optimal strategy. Thus creating a situation where communicating offered an evolutionary advantage. Furthermore, the more agents that adopted this strategy, the more stable the communicative system became.

3.5.2.1 MacLennan & Burghardt

Another example of such work that is also widely cited within the literature is that of MacLennan & Burghardt (1994), who attempted to make as few assumptions about their agents as possible by simplifying their model in every way.

They present a population of agents, whose level of cooperation is used as a fitness measure. This, as well as the organisation of the signals used by the agents, are then compared and contrasted under three conditions; with suppressed communication, permitted communication, and permitted learning.

Agents take the form of finite state machines, consisting of transition tables in the form of genetic strings, where each state is represented by a finite number of integers. In order for these agents to be successful at communicating, they need to both transmit a signal and take action at some point. The former transfer information about the local environment to the global environment, so that it is accessible to other agents. Whereas the latter are necessary to base a behaviour on the agent's local environment, so as to allow cooperation.

Fitness is then calculated from the number of times an organism has cooperated with another agent. In other words, how many times said agent acted upon the local environment of others. Due to the way in which information about the local environment is unavailable, except through communication via the global environment, measures of cooperation are a direct measure of communication. Thus, if they coordinate more than they should by chance, it can be seen that communication is having an impact upon the system.

3.5.3 Evolving A Symbolic System

Another flavour of agent-based models that explore aspects of language evolution are those that focus on the establishment and evolution of symbolic communicative systems.

3.5.3.1 Steels

One of the most widely cited, and perhaps influential, models in language modelling was originally presented by Steels (1995), and has since been worked on by a number of others. This has become colloquially become known as the *Naming Game*.

Here, a population of agents take turns in being a speaker and a hearer. They, play a series of language games, involving the speaker being presented with a subset of all possible meanings and selecting a topic from these at random, and looking up the name of the selected object from its memory. This name is then presented to the hearer, who is only given the name selected by the speaker, and has to infer the topic from it. A game is then considered a success if the hearer selects the correct topic. The speaker can invent a new word if one does not yet exist for the topic, and the hearer is able to adopt an unknown word because the speaker corrects a wrong choice if the game fails. The speaker and the hearer can change their opinion about which word is most common for the naming of an object. The agents have bi-directional memories, which consist of weights between topics and their names; these are bounded between 0.0 and 1.0.

This work is largely based upon two hypothesis (Steels 1995). First, that language, like many other phenomena in nature, is a self-organisational process. This belief is based on the observation that no single language user has a complete view of their language, in that there is no single linguistic master that controls the form of the language for all those that use it, and that language users shape and reshape language through their localised conversations. Steels (1995) also argues, perhaps more contentiously, that no separate mechanisms for language acquisition are necessary because the same mechanisms that are behind language formation also allow new language users to acquire the language of their given social group. Second, Language is driven towards ever greater complexity in light of the constraints placed upon language use; limited time to communicate/interpret an utterance, limited feedback about success, etc.

The Naming Game, as presented in its original form, then explores how a group of agent, within a single generation develop a vocabulary to name themselves and to identify each other using spatial relations. Specifically, Steels (1995) explores how the amplification and self-enforcement of fluctuations, which here take the form of random changes to word-meaning associations, are influenced by communicative success. This ultimately leads to coherence in the form of an emerging vocabulary, even when new agents and meanings are introduced.

The agents in this model exist in a two-dimensional plane, and each agent has a directional orientation; north, east, south, and west. Agents are able to perceive spatial relationships between themselves and other agents by way of a visual field. This enables them to both determine the direction in which an object is located, so as to identify it to another agent, and to determine which objects satisfy a given spatial description so as to deduce which object an agent is referring to. When communicating to one another, both the speaker agent and the hearer agent share the same context.

Linguistic behaviour then takes the form of a language game, the basic premise of which is similar to the game theoretic models above. The speaker starts the game by selecting another agent to talk about, then uses its own language to try to produce a word to identify them. The hearer then receives said word and attempts to deduce which agent the speaker was referring to. The speaker then indicates whether or not the hearer was correct in its interpretation.

The communicative success of such a word-meaning association is recorded over multiple conversations, and the average success rate is then used to consider a change in the association. If the average success is 1.0, no change is required; and if it is 0.0, the opposite is true. Whether or not a change is required for intermediate values between 0.0 and 1.0 is then calculated by a sigmoid function. The words used in the last context then influence any subsequent change in the agent's language. Such a change takes one of two forms. If no word was used in the last conversation, which would happen if no word-meaning pair yet existed for the meaning the speaker wanted to convey, a new word is created with a probability of 0.5. Alternatively, if a word was used in the last context then it is probabilistically adopted in the new word-meaning association.

The model presented by Steels (1995) demonstrated that, under these circumstances, agents will give rise to a conventionalised language. More interestingly, it shows how these same self-organisational processes can support language change. This is due to the way in which, although new language users will create new words and associations during their initial conversations, these will die-out and be replaced by the linguistic conventions of the group due to them not achieving any communicative success.

A more recent incarnation of the Naming Game has been put forward by Steels (2007). Here, the ideas of *recruitment theory* are adopted in arguing that language users experiment with different strategies for solving the task of communication, and opt for those that optimise communicative success and cognitive economy; much like the aforementioned model by Jager (2007).

Each strategy requires specific cognitive mechanisms, which may not have evolved for purely linguistic purposes. The adoption of any particular strategy has an impact on the emergent language and fixates the strategy within a population of language users. In other words, the idea here is that autonomous language users discover strategies that help them to establish and negotiate a shared lexicon through interactions with their peers; meaning that the resulting lexicon is a culturally transmitted temporal consensus.

Steels (2007) demonstrated how this model behaves with the agents using various strategies:

1. **Adoption:** There is no lateral inhibition and no decrease on failure.

2. **Enforcement:** There is an increase in cases of success but no decrease of competitors and no decrease on failure.

3. **Lateral inhibition:** There is both enforcement and lateral inhibition but no decrease on failure.

4. **Damping:** There is enforcement, lateral inhibition and also a decrease on failure.

All of these strategies lead to the emergence of a successful communication system, but only strategies three and four lead to an optimal lexicon size.

Thus, although a number of different strategies may be successful, some are more successful than others due to such strategies resulting in a smaller lexicon. This allows a population to arrive at a communicative convention, even when the population is in flux. It is also demonstrated in this model that the key mechanisms that were required for these agents to be successful was a bi-directional associative memory.

3.5.3.2 Cangelosi

In keeping with the theme of spatial navigation, Cangelosi (2001; Cangelosi et al., 2002) presents a model that explores how symbols and syntax can emerge from relatively simple communication signals. This model is based on the argument that there is a need for models of language evolution to draw a distinction between signals, symbols, and words, as well as the need for simulating the grounding of symbols and words.

These models contain a population of agents, which take the form of three-layered, feed-forward neural networks with randomly assigned weights. These agents live in a grid world, where there are both edible and poisonous mushrooms. Agents reproduce on the basis of their ability to eat edible mushrooms and avoid poisonous ones, which obviously involves categorising them. In order to categorise a mushroom within an agent's visual field, the input into the agent's neural network consists of the angle (in a clockwise direction) of the mushroom from the agent's current facing direction, a 10-bit binary representation of the mushroom's perceptual features, and a 3-bit binary representation of one of the eight possible signals. The output of the neural network then includes the movement of an agent and one of the eight possible emitted signals in the same manner as the signal encoding input units. Agents then wander around the world, and in each iteration, perceive and encode the details of the nearest mushroom in this manner.

Three different population types are experimented with; one where there is no language, one where the language is provided by the researcher, and one where the language evolves autonomously. In the third type, the speaker agent selects the closest mushroom and the hearer is randomly selected from the population and placed next to the speaker so that they have the same objects in their visual field. The hearer processes the chosen mushroom's perceptual features and the signal part of its output is used as the signal-part for the input for the speaker; who then has to decide whether or not to eat the mushroom.

This model demonstrated that the agents were not only able to evolve an efficient and informative language, but by comparing the fitness performance of the three different population types, were also able to demonstrate that language is a useful adaptation for such agents.

Cangelosi (2001; Cangelosi et al., 2002) then explore the emergence of symbolism with a similar model set up. The input into the neural network consists of the location of the closest mushroom, its perceptual features, and the communicative symbols. While the output layer consists of the mushroom type, the movements required to approach it, and the mushroom names. The name being divided into a 2-bit binary 'verbs' which translates to 'eat' or 'don't eat', while the rest of the representation refers to the mushroom's perceptual features. During the first 300 generations, agents do not communicate and therefore do not use their symbol nodes. Generations 301-400 then involve only the newly created agents foraging for mushrooms, while the older agents act as teachers of the mushroom names.

This model demonstrated that, in the vast majority of cases, agents learned a language that was based on symbolic association between the mushroom names and the two verbs.

This indicated that the agents in this model adopted a symbolic strategy when learning linguistic symbols and the syntactic rules for combining them.

3.5.3.3 Hashimoto

In exploring the emergence of a symbolic communication system, Hashimoto (1997) built upon previous work that demonstrated how syntactic structure was useful in a community (Hashimoto & Ikegami 1996). This later model incorporated a word-meaning feature, which allowed for the analysis of the relationships between words; and in doing so, Hashimoto (1997) was able to investigate the development of linguistic structures.

Underpinning this work is the belief that the meanings of words can be best understood in regards to their inter-relationships. This is achieved by adopting an algorithm that measures the graded similarity between words, based on the logic that similar words are used in similar sentences and similar sentences are composed of similar words[2].

Agents in the model are defined as a grammatical system consisting of a set of non-terminal symbols, a set of terminable symbols, a start symbol, and a list of rewriting rules. Agents can have either a context-free or regular grammar. The model then explores the development of structure in terms of word similarity within these agents.

It was found that the structures that developed formed clusters of word types. These clusters constitute a categorisation of words due to the way in which words within a cluster have stronger relationships with one another than they do with words in other clusters.

3.5.3.4 Ginzburg & Macura

In a similar vein, Ginzburg & Macura (2007) look at the emergence of a class of utterances that they refer to as *meta-communicative interaction utterances* (MCI), in order to illustrate that semantic expressiveness is not correlated with syntactic complexity. MCIs are utterances that require either an acknowledgement in order to ensure that the utterance has been correctly understood, and clarification if there are unclear aspects of the utterance or it there are indications of erroneous assumptions concerning naming. These allow concepts associated with a state of equilibrium or lack of divergence to be maintained within a linguistic population.

This gives rise to three questions. First, under what circumstances does a non-MCI linguistic system evolve into an MCI system? Second, what mechanisms are involved in such a shift? Third, why is the resulting interaction system maintained?

---

[2] An algorithm originally developed by Karov & Edelman (1996).

The simulation starts off with a set of agents that possess distinct lexicons, and ends in a state whereby the agents associate meanings with each word in their lexicon. This is achieved due to the way in which the convergence rate of a population that relies solely on introspection is intrinsically bounded.

The utterances made by the agents consist of a single word and are stored in a tuple along with the meaning for that particular object (a plant in this model); i.e. `(plant-type,plant-word)`. The speaker always has a word to use for the selected plant. However, the hearer may not be able to interpret said word, in which case, one of two things happen, depending on the type of agent. A *clarification request agent* (CRA) will ask for a clarification, which involves the speaker presenting the hearer with the tuple to store in its lexicon. An *introspective agent* (IA) will try to guess the meaning by looking at each of the plants in its field of vision, increase the association of the plant-word, and then store it in a temporary lexicon until it has enough information to pick a meaning for the unknown word. At which point, the plant-word becomes part of its permanent lexicon.

Agents have a vision capacity, and randomly wander around their world in search of food and other agents. When two agents meet, they play a naming game whereby the speaker chooses a food resource in its field of vision, and sends its name for it to the hearer, who then tries to interpret it.

Ginzburg & Macura (2007) ran the model with various values for several parameters; meaning space, population size, and acquisition threshold. The last of these being the number of times an agent has to be exposed to an unknown word before it can acquire it. In doing so, they demonstrated that the convergence rate of a population of purely introspective agents is bounded; a bound that decreases as the population size increases. However, this bound disappears completely once CRAs are introduced.

3.5.3.5 De Boer

In a model with quite different research interests, de Boer (1997) puts forward a functional explanation for certain universals that are present in vowel systems. Other work had demonstrated, by way of computational simulations, that such universals within a vowel system can be explained, while simultaneously minimising the necessary articulatory gestures (Lilijencrants & Lindbolm 1972, Carre & Mrayati 1995). Building on this, de Boer (1997) explores by which actions of the individual language-users such a minimisation is caused.

All agents possess a list of vowels that are empty upon creation. Vowels are represented by three parameters, tongue position, tongue height, and lip rounding, all of which have a value between 0 and 1. The simulation is then concerned with the emergence of a coherent and useful phonology within a population of initially empty agents. Thus, agents need to invent speech sounds in order to get the simulation running, as well as to introduce new sounds into the simulation run; both based upon a probability function.

Once created, the agents participate in language games. Here, the speaker agent selects one of its phonemes and conveys it to the hearer agent, who then interprets it in relation to its own phonemes and then produces the phenome it thinks it has recognised. The speaker then interprets the hearer's phenome in relation to its own list to see if it matches the one it originally produced. If it does, the game is considered to have been a success.

The agents keep track of how many times each phoneme is successfully used. The quality of the phoneme is the success score divided by the number of times it was used. If successful, the hearer attempts to shift the phoneme used a little closer to the one given by the speaker. If it is unsuccessful, and if the quality of the phoneme is low, the phoneme is shifted. It is not shifted if the quality is high, due to the probability of it being a good representation of another possibility. Otherwise, a new phenome is created, again using a hill climber heuristic, that is similar to the one presented by the speaker. In addition, phonemes with a persistently low quality have the possibility of being removed, while phonemes that are similar will be merged.

The results presented by de Boer (1997) demonstrated that a population of language users can, under a specific set of constraints, give rise to a vowel system without the need for innate mechanisms or the ability to mind-read. Furthermore, this vowel system is in a constant state of flux, much like natural language.

3.5.4 Evolving Grammar And Syntax

So far, we have looked at models that have been primarily concerned with either the evolution of communication or the evolution of symbolic representation. There is a third category to be discussed here. Namely, models that are concerned with the evolution of syntax.

3.5.4.1 Batali

Batali (1998) presents a model where a population of agents can not only encode structured meanings as a sequences of tokens, but also interpret sequences as tokens. Agents in this model take the form of real-numbered 'meaning vectors' and a simple recurrent neural network is used to receive sequences of characters from a set.

During each communicative interaction, all of the values in the speaker's meaning vector are set to either 0.0 or 1.0, depending on which of the set of meanings is to be conveyed. The meaning vector is then used by the network to determine the sequence of characters to send. This sequence is then presented to the hearer, who attempts to infer the meaning intended by the speaker. Communicative accuracy of the interaction is then assessed by comparing the values of both the speaker's and the hearer's vectors after the hearer has processed the sequences.

Here, agents eventually develop a communication system that is highly successful, even when presented with novel meaning combinations. This inter-agent coordination is achieved as a result of a distributed process in which individuals learn by observing the behaviour of others, with no external guidance.

3.5.4.2 Briscoe

The work of Briscoe (1998, 1999) was motivated by four aims. Firstly, to demonstrate that specific triggers can lead to the creation of individuals who can effectively acquire a grammar. Secondly, to explore to what extent certain environmental variables could impact the grammar that is learned. Thirdly, to analyse the extent to which such a grammar is encoded 'biologically'. Finally, to explore how selection pressures, such as learnability and interpretability, interact with one another to mould and constrain the evolution of language.

In his models, Briscoe (1998, 1999) implements his agents with a *language acquisition device* (LAD). This consists of twenty parameter settings, which define seventy languages and roughly three hundred grammars. Whether or not an agent has default parameter settings dictates whether or not said agent has some form of innate language ability. The agents are also equipped with an algorithm consisting of three steps; shift, reduce, and halt. These modify the stack that contains the categories that correspond to the input sentence. A secondary algorithm, that is attached to the first, ranks the ability to parse sentences. Finally, the agents also possess a parameter setting algorithm, which will alter the parameter when the input cannot be understood. When a communicative act fails, a parameter is chosen, in accordance with the location within the partial ordering of the inheritance hierarchy of parameters, and is reset. Due to the way in which each parameter can only be reset once, the most general are reset first, and the most specific reset last.

These models identified three selection pressures that drive the evolution of language; learnability, expressivity, and interpretability. The first is seen in the number of parameters that need to be set to acquire a target grammar. The second, is reflected by the number of trigger types that are necessary in order to converge on a language. Finally, interpretability is reflected by working memory load; the parsing cost in the context of this model.

3.5.4.3 Gong et al.

In a piece of work that fits in with the research themes seen in Briscoe (1998, 1999),Gong et al. (2014) presents a rule-based model exploring the nature of lexicon-syntax co-evolution.

The agents in the model are equipped with general learning mechanisms and clearly defined holistic and compositional linguistic rules. The aim here is to ascertain whether or not these mechanisms can bring around a transition from a holistic protolanguage to a compositional communal language. Here, languages take the form of meaning-signal pairings.

Agents share a semantic space, where the meanings can be encoded into signals (sentences). While signals are a string of syllables drawn from a signalling space. A signal encoding an integrated meaning can be segmented into subparts mapping semantic constituents, and sub-parts can also combine to encode an integrated meaning.

Lexical rules come in both, holistic and compositional form. The former take the form of `chase<wolf,bear>` ↔ `/abcde/`, while the latter map semantic constituents onto a subpart of a sentence; `‘wolf’` ↔ `/cd/`. In the latter case, ordering rules are used, such as `category 1 << category 2`, which denotes that category 1 comes before, but not necessarily immediately before, category 2.

It is shown that, when the agents start off by sharing a small number of holistic rules, there are at first many iterations that are holistic. However, given more linguistic experiences, recurrent patterns start to emerge and are processed as compositional rules. There is then competition between holistic and compositional rules, which the latter win due to the advantage they gain from being combinational, which allows for many meanings to be expressed involving encoded constituents.

3.5.4.4 Smith

Motivated by the above work of Oliphant & Batali (1997), Smith (2007) sought to explore the inferential nature of language transmission and its relationship between language change and the indeterminacy of meaning. In this sense, inferential communication refers to how information in a communicative act is transferred indirectly, by way of the hearer having to infer the meaning of a signal with the aid of pragmatic insights and the context in which it is used.

The motivation here is that the majority of models of cultural language transmission involve the explicit conjunction of meaning-signal pairs. Smith (2007) argues that this weakens the claims made by them due to the way in which such a set-up would lead to the emergence of a syntactic structure that is identical to the pre-defined semantic structure. Thus, this model starts off with agents that have neither conceptual nor lexical structures, but do possess the ability to create their own representations and infer meanings based on their experiences.

To create a meaning, the speaker selects a subset of meanings to act as the context, and then randomly selects a meaning from this subset to try to convey it to the hearer. The speaker then tries to find a distinctive feature of this object, which does not describe any other object within the context. If one cannot be found, then the speaker selects an existing category and splits it into two equal parts, thereby creating two new categories that are each a subset of the existing category. This, over time, results in a hierarchical meaning tree, with nodes near the tree root representing more generic meanings.

Once a distinctive feature has been choosen, a signal is selected for it from the speaker's own lexicon. This is done by way of the speaker using the aforementioned obverter method, which involves choosing the signal which it would be most likely to interpret correctly given the current context and its own existing semantic categories; resulting in signal choice being based upon the speaker's own interpretive behaviour. If there is not a suitable signal, one is created from a random string of letters. This signal is then presented to the hearer agent who can only observe said signal and the context. In turn, the hearer uses its own conceptual structures to try and discriminate each object in the context from all of the others, and thereby create a list of possible meanings.

By allowing these agents to interact in a generational setting, Smith (2007) showed that individual meaning creation, and the uncertainty inherent in meaning inference lead to different degrees of variation in both conceptual and lexical structure. With the conceptual variation and imperfect learning resulting in different bottlenecks on transmission, which in turn, result in rapid language change across generations. Despite this rapid change, however, the language itself remains sufficiently stable; with each subsequent generation able to re-establish itself and maintain the utility of its successful communication system.

3.6 Conclusion

Many models have been put forward consisting of a population of agents attempting to learn meaning-signal mappings. Models where the meanings are fixed and non-changing, while the associations between them and the signals are learned without the aid of any linguistic overlord dictating which signal should map to which meaning; and vice-versa (Steels 1995, Hurford 2000, Kirby 2002*c*, Kirby & Hurford 2002, Barr 2004, Belpaeme & Bleys 2007, Steels 2007, Smith 2007).

A disproportionate number of the models discussed here can be classified as *expression/induction* models (Henceforth E/I). One of the main characteristics of E/I models is that the representation of language pervades across generations of the model in two forms; the mental grammars of the individual agents and the utterances used in agent interactions. A distinction that Chomsky (1986) originally observed with his concepts of *I-Language* and *E-Language*, respectively. However, E/I have a number of other characteristics that are worth mentioning here[3]:

Agents: In E/I models, there are populations of individual agents that each have two features. First, they possess some form of expression/invention capacity. Whereby an agent, upon being presented with a meaning to convey, will either use its current mental grammar to produce an utterance or will create an entirely new utterance if the meaning is novel. Second, they possess an induction capacity, which enables them to acquire an internal representation of a language system from an impoverished sample of said language.

Non-evolving agents: During the course of simulation runs, all new agents are created equally, there is no biological evolution, and they all have identical linguistic abilities. As such, E/I models are always primarily concerned with the cultural evolutionary processes behind language evolution. Given that E/I models always start of with no language (mappings between meanings and utterances) in place, they thus seek to explore the emergence of linguistic systems through cultural transmission and not the process of historical change in relatively more complex languages[4].

Pre-defined meanings: E/I models always have a pre-defined meaning space that does not change during the course of the simulation, with only their mappings to utterances changing. One can interpret the kind of meanings seen in these models as being reminiscent of how the grammar seen in early linguistic communication systems may have resembled the inflectional morphology of verbs in modern language (Batali 1998).

Finite vocabulary: E/I models always have a finite set of atomic symbols from which utterances can be created. Indeed, Hurford (2002) observes how the size of this vocabulary relative to the meaning space is an important factor in the behaviour of the models.

3.6.1 Population Dynamics

The literature on E/I models features two types of population dynamics; *uni-generational* and *multi-generational*. In the former, the evolution of language is explored within a single generation of agents, meaning that the population comprises of the same individuals for the duration of the simulation run. In the latter, language is passed from

---

[3] The interested reader is advised to consult Hurford (2002) for an in-depth discussion on the nature of E/I models.

[4] Although, the feasibility of applying such methods to the analysis of such historical linguistic development has been explored (Hare & Elman 1995).

one generation of agents to the next, with agents being periodically removed from the population; thus remvoing their particular grammar. Their legacy lives on in the form of their observed behaviour, which acted as training input into new language-learners that were created without any internalised grammar.

The main focus of this thesis is on E/I models, with a particular focus on how population size and dynamics impact linguistic change. It has been seen that multi-generational E/I models tend to have two aspect to their population dynamics.

First, the size of populations tend to be rather small (Hurford 2000, 2002, Kirby 2002*c*, Kirby & Hurford 2002), with some only having two agents per generation; one mature speaker and one immature learner.

As Hurford (2002) observes, although such small population sizes may at first seem like a gross oversimplification, it does eliminate one factor from the evolutionary scenario. Namely, the behaviour seen in most E/I models is the result of social coordination between individual agents; an establishment of conventionalised usage of meaning-signal pairs between individual agents. Thus, as will be seen later, at the start of a simulation run with a mature population of more than one agent, there will be a number of different signals for the same meaning. However, as the authors of such models observe, such coordination is not actually the focus of these models. Moreover, they are interested in the evolutionary transition into syntax; which is a matter of how individuals organise their internal grammar, and not of coordination behaviour. Therefore, a simulation with such a minimal population size enables the researcher to focus on the development of such grammar, without having to be concerned with matters of coordinated behaviour.

This is true. Nevertheless, as was demonstrated by the aforementioned early A-life simulations of Werner & Dyer (1992) and MacLennan & Burghardt (1994), coordination has undoubtedly played a crucial role in the emergence of language and change in linguistic structure.

Second, in E/I models, the number of agents within a population tends to remain constant throughout the simulation run. This is perfectly fine when addressing the issues that the authors of these models were interested in. However, it would be interesting to explore how an expanding and contracting population size, with varying numbers of mature language users and immature language learners, could impact the emergence and form of syntax (Johansson 1997, Hurford 2002).

Indeed, such an E/I model geared towards these interests could provide valuable insights for a growing body of research that is interested in the nature of the relationship between language and population change; such as the impact of population size on linguistic structure (Johansson 1997, Nettle 1999*a*, Wichmann & Holman 2009, Lupyan & Dale 2010, Milroy 2013, Trudgill 2013). This will be the focus of the remainder of this thesis.

Chapter 4

Language Change and Transmission Through Iterated Learning

4.1 Introduction

As we have seen, although all language evolution researchers would agree that language is the result of both innateness and lifelong learning, there is much contention over which, and where, lines should be drawn. Notwithstanding some of the interesting findings of previous work, it would perhaps be beneficial to, instead of positing such work within strict dichotomies, analyse these findings in terms of gene-environment interactions (Tallerman & Gibson 2013*b*), and how such dynamics may be largely responsible for linguistic structure and behaviours. Kirby (2002, 2002*b*) presented a framework that could prove to be useful to researchers of this persuasion; a framework that views language as the result of the interaction between three different complex systems (fig. 4.1).

Biology: Here, the focus is on all of the phylogenetic adaptations that must have occurred in the hominin lineage in order for language to emerge. Such adaptations, which enable language comprehension and production, would have evolved in accordance with known biological processes. They must have therefore developed in accordance with selection pressures from the environment in which early hominins found themselves, with protolanguage and its accompanying neurological or anatomical changes having no a negative impact upon selection.

Culture: Although an instinct to learn languages may have been encoded into the human genome, no specific lexical items of syntactic constructions are. Thus, socio-cultural processes play a role in language evolution. Perhaps not just through

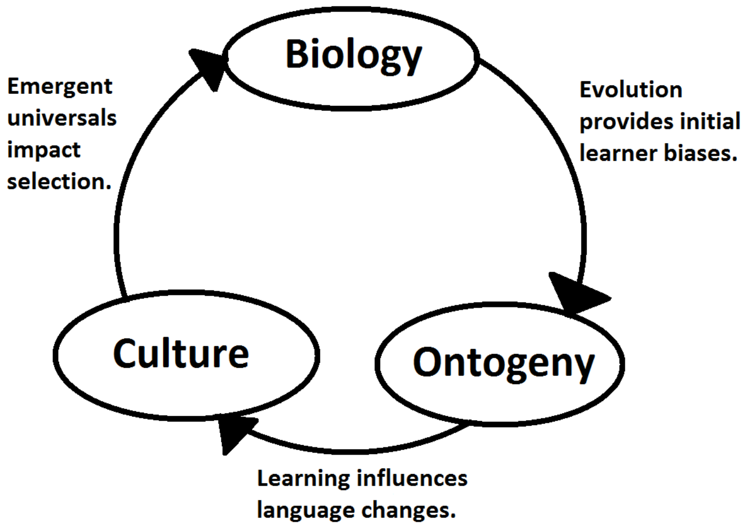


Figure 4.1: Diagram adapted from Kirby & Hurford (2002) showing the interactions between the three adaptive systems that give rise to language.

the continual flux we see in vocabularies and variations in the flexibility of use, but also through the way in which extensive contact between different linguistic communities is likely to have had an impact upon the evolutionary process of language.

Ontogeny: Ontogenetic processes in this context are concerned with how an individual's language use develops and changes throughout the course of their lifetime, and with how they alter their knowledge in accordance with their environment. MacNeilage (2013) argues that ontogeny recapitulates phylogeny in the form of speech production, in that infants and early hominins must both share the same biomechanical constraints on mouth movements. This leads to initially simple syllable patterns in both cases; such as babbling.

The idea here is that we learn languages by observing others, but these languages change through a process of cultural change, which in turn alters the environment, and thus the selection pressures that guide human biological evolution (Kirby 2002*b*). Indeed, it is likely that linguistic structure has, at least in part, been shaped by external functional pressures. However, such pressures are by no means completely 'external', in that they are a result of the way in which humans process and learn about the world around them (Tallerman & Gibson 2013*b*); which in most cases are likely to be phylogenetically primitive.

Thus, it is arguable that the learning mechanisms possessed by early linguistic hominins dictated the linguistic structures that emerged through selection for the structures that were easiest to learn. In other words, these early languages evolved in accordance with learner bias. Assuming that such a language faculty is adaptive, one would expect that "early humans evolved in ways that would better enable them to learn ambient languages quickly, and making the most of a fragmented input" (Tallerman & Gibson 2013*b*:506).

The scheme laid out by Kirby & Hurford (2002) above was the motivation behind a particularly influential E/I model that sought to explore such notions; the iterated learning model. Following the conclusions of chapter 3, in regards to future research using populations of agents, this chapter will take a model originally presented by Kirby & Hurford (2002). This model had a population size of one mature speaking agent and one immature language learning agent per generation. Here, it will be extended to incorporate a population with multiple mature and immature agents; with language learners learning their language from multiple trainers.

4.2 Compositionality and Iterated Learning[1]

As stated above, human language has a number of notable design features. One that is of particular interest is the way in which utterances are constructed from sub-parts, such as words and parts of words, which are re-used and recombined in systematic ways. Thus, the meaning of an expression is related to the meanings of its constituent parts and the way in which they are combined. This trait enables language to be expressively open-ended, and is known as *compositionality* (Brighton & Kirby 2001, Kirby 2002*b*, Smith et al. 2003).

Compositionality endows human languages with an obvious adaptive advantage in terms of their ability to communicate novel meanings (Kirby 2007). Indeed, given the utility associated with the ability to construct a wide range of messages from just a few learned basic units (Kirby 2013), it is remarkable that we do not see compositionality being used as part of a learned mapping between meanings and signals in the communication systems of other species[2].

This, combined with the view that language is culturally-transmitted, and that this may have a crucial role in shaping the way in which it is formed (Smith 2002, Brighton et al. 2005, Christiansen & Chater 2008) has led to a body of work arguing that compositionality may have arisen, not due to its utility to us, but because it better ensures the continued existence of the language itself (Kirby 2007). This work sees the self-preservational development of language occurring as a result of a cultural-evolutionary process termed *iterated learning* (Brighton & Kirby 2001, Kirby & Hurford 2002, Smith et al. 2003, Kirby et al. 2008, 2014). This is a process whereby an individual learns

---

[1] The following section draws from Brace et al. (2015).

[2] It has been claimed that bee dances display limited compositionality (Kirby 2012).

their cultural behaviour from other individuals, who have themselves acquired their cultural behaviour in the same way. In other words, the input into an individual's learning process is, itself, the output of prior learning in other individuals.

Here, this involves an individual being presented with a set of meanings that it wishes to convey, choosing signals for each of these meanings, and then transmitting these meaning-signal pairs, or utterances, to another individual who then learns from them. This process is repeated generation after generation, and can be seen to represent how language competence and understanding can develop through observational learning (Brighton 2002).

Iterated learning is arguably one of the two major categories of E/I models, along with the aforementioned naming game (Steels 1995, 2007, Steels & Garcia-Casademont 2015), and captures an important aspect of language. Namely, the *linguistic bottleneck*, which refers to the way in which no single language user will ever have a complete picture of a given language. Thus, resulting in language users always learning their language through a limited subset of the full, and potentially large set, of grammars. The belief here is that languages exists in two forms; the internal version, consisting of mappings between meanings and signals within the minds of language users, and the external form consisting of the actual spoken language. Thus, in order to pervade from one generation to the next, a language needs to be consistently mapped from the internal form to the external form, and back again. This is reminiscent of Chomsky's (1986) aforementioned concepts of *I-language* and *E-language*.

4.3 The Iterated Learning Model[3]

There have been several published variants of the Iterated Learning Model (ILM). Here, the focus will be on the ILM originally presented by Kirby & Hurford (2002). The central contribution of iterated learning models is first to have successfully idealised this process in a simplified setting that is amenable to study. Then to have demonstrated that the character of the linguistic bottleneck is crucially important to both whether or not language can be successfully passed from one generation to the next and, in the situations where this transmission can be achieved successfully, show that it is also crucial to the linguistic structure that arises. Kirby and Hurford's (2002) model has four key components:

1. A finite meaning space, $\mathcal{M}$.
2. A finite signal space, $\mathcal{S}$.
3. One speaker.
4. One learner.

---

[3] The following section draws from Brace et al. (2015).

Here, a language is defined as a mapping between a finite space of signals and a finite space of meanings. Each meaning and each signal are represented as an 8-bit binary string. Thus:

$$\mathcal{M} = \{m_1, m_2, \ldots, m_{256}\} \tag{4.1}$$

$$\mathcal{S} = \{s_1, s_2, \ldots, s_{256}\} \tag{4.2}$$

Each agent's personal mapping from signals to meanings is implemented in the form of a three-layer feed-forward artificial neural network with eight nodes in each layer (figure. 4.2). Each of the eight nodes in the input layer are influenced by one of the eight bits in an uttered signal. The degree of activation of each node in the input and hidden layers influences every node in the immediately downstream layer via a weighted connection. Each node's activation is determined by the weighted input it receives from all of the upstream nodes, squashed by a standard logistic activation function:

$$y_i = \frac{1.0}{1.0 + e^{-x_i}} + \theta_i + I_i \tag{4.3}$$

$$x_i = \sum_j \omega_{ji} y_j \tag{4.4}$$

Where y_i is the activation level of a particular neuron, i, and x_i is incoming stimulation received by i, calculated as the weighted sum of all of the immediately upstream activation values. Each neuron in the hidden layer also receives a constant bias input, $\theta_i = 1.0$, and may receive an external input $I_i \in \{0, 1\}$. The activation values of the output layer are then translated into an 8-bit binary meaning by thresholding each node's activation value. This string represents an agent's best guess as to the meaning of the utterance that was input into the network. During learning, an agent updates the weights of its network using back propagation with a learning rate of 0.1 and no momentum term (Rumelhart et al. 1986).

Initially two agents are created, a mature language user (or 'speaker") and an immature language user (or 'learner"). At the outset of the simulation there is no established language in place, so the mature language user is assigned a language comprising of a random mapping from each meaning to a randomly chosen signal. The immature language user is assigned a random neural network, i.e, each network weight is drawn from a normal distribution with a mean of 0 and a standard deviation of 0.1; with each node's bias input being 1.0.

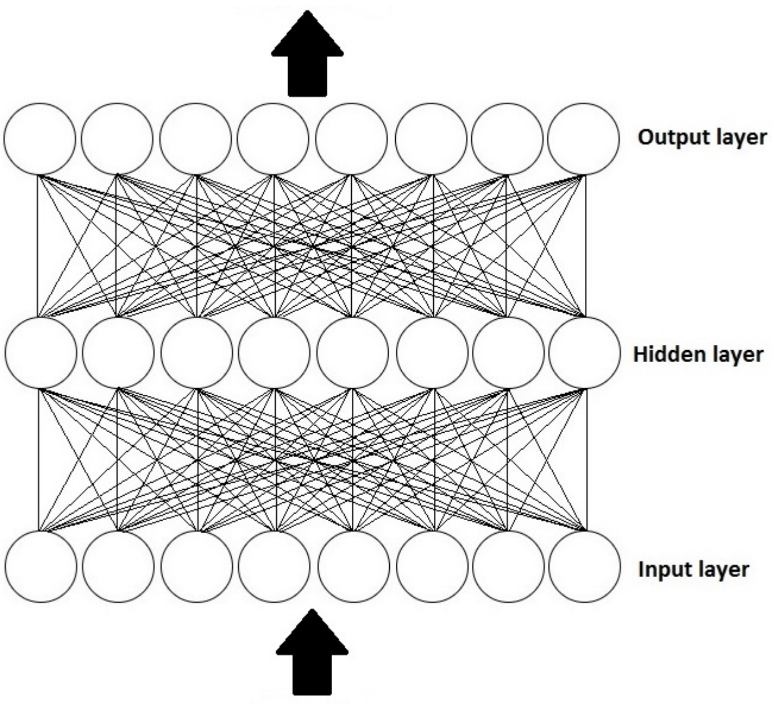


Figure 4.2: The agent's neural network architecture.

The mature language user, M, then trains the immature language user, I, for a number of training episodes, T. Each episode involves M being assigned a meaning to express and generating an associated signal, and I using their neural network to infer a meaning associated with that utterance. Any difference between the true meaning that M attempted to express and the meaning that I infers then results in back propagation making changes to I's neural network in an effort to minimise this comprehension error. Note that in order for this supervised learning to take place, ILMs assume that I is able to make use of knowledge of the true meaning that M intended to convey.

The full set of training episodes that an immature language user experiences often involve multiple exposures to the same fixed set of unique meanings. An agent might experience E epochs of training with each epoch comprising the same set of B randomly chosen unique meanings experienced in an order that is randomised for each epoch; i.e., $T = E \times B$. The number of different meanings communicated to a language learner, B, represents the aforementioned bottleneck.

After all training episodes are complete, the mature language user is discarded, the immature language user is promoted to become the new mature language user, and a new randomly configured immature user is created to be trained. This process repeats for some fixed number of generations. Note that at the start of every generation the immature language user is assigned an entirely random neural network; there is no inheritance of language other than through experience of language learning episodes. Note also that the population structure is $1 + 1$; at any moment in time one mature speaker is training one immature learner.

Since these agents have a neural network that maps *unidirectionally* from signals to meanings, they require an additional mechanism in order to generate signals for particular meanings. To this end, Kirby & Hurford (2002) adopt the aforementioned *obverter* learning procedure by Oliphant & Batali (1997). Here, each speaker assumes that the hearer's internal mapping between signals and meanings is similar to its own. Consequently, when choosing which signal to make for a particular meaning, the speaker will choose the signal that, if presented as input to their own neural network, would most strongly cause them to infer this meaning themselves.

In order to apply the obverter procedure to the ILM, Kirby & Hurford (2002) employ a confidence measure to determine which signal to produce for a given meaning. A speaker aiming to express a particular meaning, m, identifies their favoured signal, s^*, in the following manner:

For each signal, $s \in \mathcal{S}$, the speaker calculates an associated confidence value[4] :

$$V_s = \prod_i (1 - |m[i] - o[i]|) \tag{4.5}$$

where $m[i]$ is the i^{th} bit of the target meaning and $o[i]$ is the i^{th} real valued output of the signaller's neural network. The signaller then picks s^* as the signal with the largest confidence. Here, the obverter procedure does most of the work. Lets take, for example, a situation whereby an agent is trained on a simple language where the input signal is replicated in the training meaning (for simplicity, they are each two bits in length):

Signals	**Meanings**
0, 0	0, 0
0, 1	0, 1
1, 0	1, 0
1, 1	1, 1

Now, imagine that, after a number of training sessions, the agent produced the following output activation patterns for each of the training input signals:

Signals	**Meanings**
0, 0	0.1, 0.1
0, 1	0.05, 0.15
1, 0	0.4, 0.15
1, 1	0.45, 0.4

At first glance, although it has learned the first pairing, the agent has performed rather poorly at learning the other three pairs. However, at this point, obverter is used, and

---

[4] Note: the product is used here over any other method, such as Euclidean distance, due to this model being a replication of that presented by Kirby & Hurford (2002), who themselves used the product measure.

this algorithm produces the following output (note that the meanings are now in the first column and the signals in the second):

Signals	**Meanings**
0, 0	0, 1
0, 1	0, 1
1, 0	1, 0
1, 1	1, 1

Thus, despite the raw output of the agent's network producing poor activation values, the obverter procedure 'cleans up' the output to provide us with adequate mappings. It can be seen that the obverter procedure can cause ambiguity in the language, with the signal $0, 1$ being the best signal for both of the first two meanings. This is due to the output activation of $0.05, 0.15$ being, in accordance with the above closeness measure, the best match to both $0, 1$ and $0, 0$. Therefore, it actually produces the wrong, and more ambiguous signal, for the one meaning it seemed to have learned best by looking at the raw output pattern.

4.4 ILM Results

Three metrics were employed to evaluate language development; expressivity, stability and compositionality. A language's *expressivity*, X, is the proportion of possible meanings that are generated by the full set of possible signals. A language with maximal expressivity is said to be complete.

A language's *stability*, S, is a relational property involving two agents, and is measured as the proportion of the meaning space that can be recovered accurately when one agent signals to another. When a language is maximally stable, any meaning expressed by one agent can be inferred correctly by the other. In the results presented in this chapter, S is calculated by subtracting the number of identical meaning-signal pairs produced by to randomly selected immature agents from 256; this latter number being the total number of meanings and signals featured in the model.

The *compositionality*, C, of an agent's language is the extent to which utterance parts convey distinct meanings. A language with zero compositionality is one in which every utterance is paired with a meaning in an uncorrelated fashion. Knowing part of the utterance provides no knowledge of part of the meaning. In contrast, a fully compositional language is one in which every part of an utterance conveys an associated part of the meaning perfectly.

The degree of compositionality in an agent's language is evaluated by first employing the obverter procedure to generate a signal for each of the meanings in the meaning space. We then calculate the values of each of the 8×8 correlations, C_{ij}, between the

256 values at the i^{th} bit of the set of signals and the 256 values at the j^{th} bit of the set of meanings. For each row, i, of this matrix we then calculate $C_{i*} = \max_i C_{ij}$, the maximum correlation between the values at index i of the signal set and the values at each of the indices of the meaning set[5]. Finally, C is calculated as the average of these eight maximal correlation values; $C = \frac{1}{8}\sum_i C_{i*}$. For a random language mapping meanings to signals, $C = 0.5$. Where a complete language is fully compositional, $C = 1$, each bit in an utterance conveys the value of one bit in the associated meaning.

It should be noted that this measure of compositionality does not completely reflect the linguistic concept of compositionality upon which it is based. Specifically, linguists define compositionality as being the way in which the meaning of a sentence is made up of the meanings of each of the individual words contained within it, and the way in which they are ordered. However, the meanings in this ILM are represented as 8-bit binary bit-strings. As a result, this measure of compositionality would equate to a linguist trying to measure the compositionality of a single word by looking at each of the individual letters that make up said word in the real world. Despite this, the compositionality measure defined here will sufficiently allow us to explore the behaviour of this ILM due to the binary nature of the representation of language itself.

The model by Kirby & Hurford (2002) displays three different types of behaviour, depending upon the size of the bottleneck. If the bottleneck is too small, then the agents do not learn; this results in a language that is both inexpressive and unstable. If, however, the bottleneck is too big, then an expressive and stable system is eventually reached; although, only after a prolonged period of time. However, agents quickly achieve a language that is expressive and stable (fig. 4.3) with a bottleneck of size 50.

The original paper by Kirby & Hurford (2002) presented results of a single typical run. Thus, figure 4.3 depicts the results of a typical single run of the replicated ILM. As such, it was deemed beneficial to also depict the behaviour of this standard ILM over multiple runs. Accordingly, figure 4.4 demonstrates how the agents still achieve an expressive and stable language, but measuring this over multiple runs ‘smooths out’ the lines. Additionally, the error bars demonstrate how the amount of success during the early generations fluctuates between different runs. The same is also true for figure 4.5, which depicts the amount of compositionality during these runs.

---

[5] It should be noted that the average of each row could have just as easily been calculated and used. However, due to the nature of how the overter procedure operates, the maximum correlation felt like the more natural design choice to the authors.

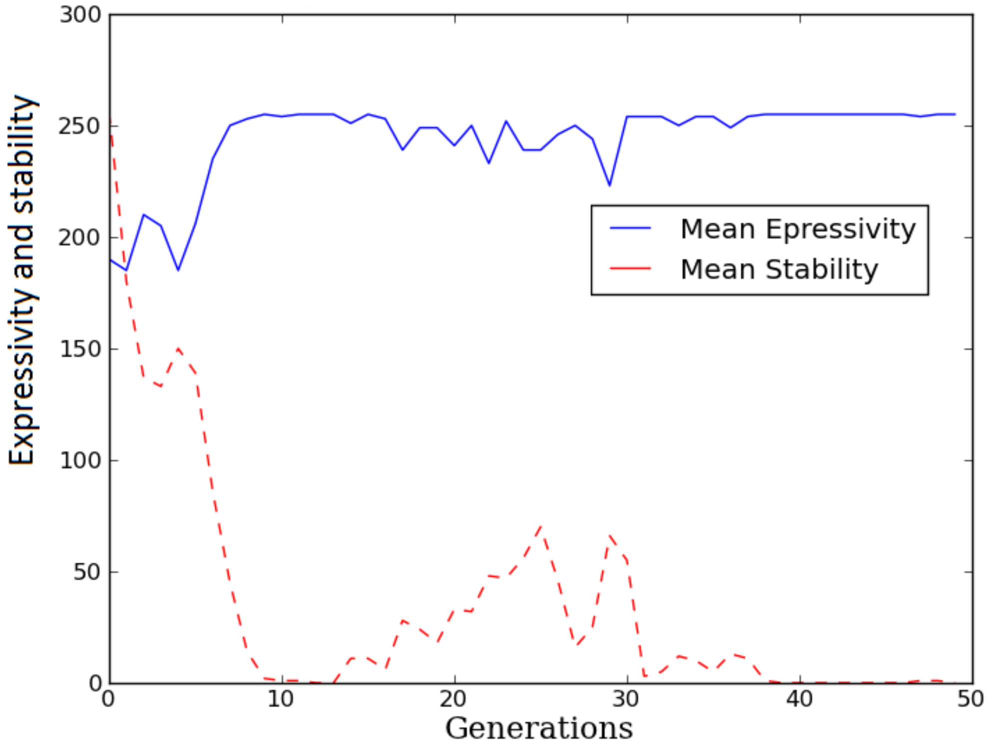


FIGURE 4.3: ILM replication behaviour. The blue line represents language expressivity and the red line represents the language stability, $256-S$. Where, N_M=1, N_I=1, B=50, E=100, M_T=1, I_T=0.

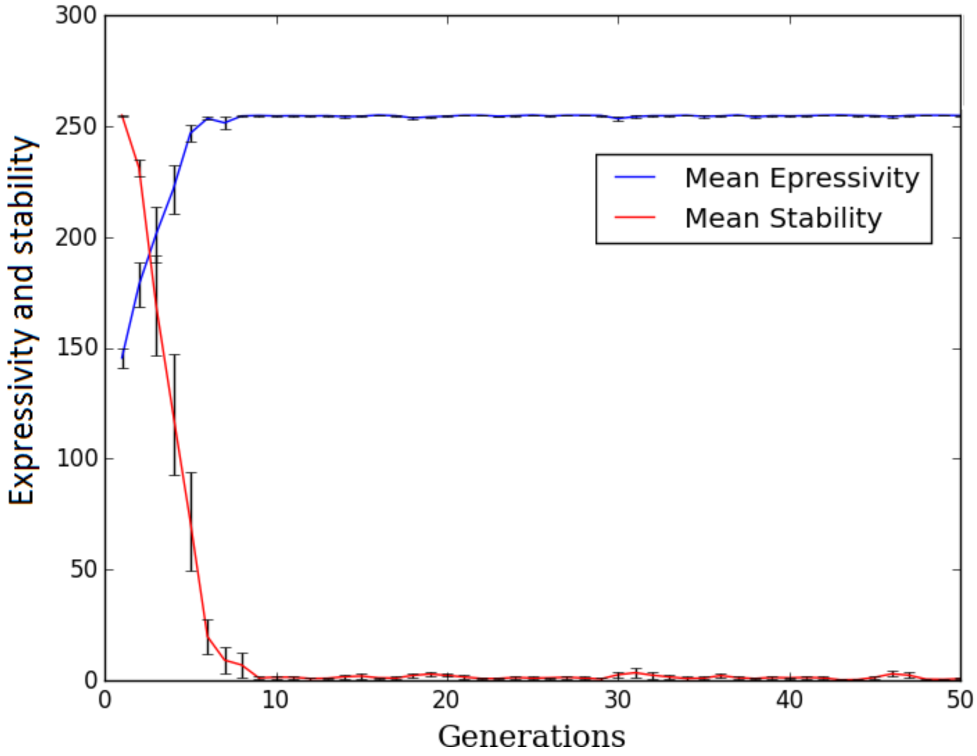


FIGURE 4.4: Averaged ILM replication behaviour. The blue line represents language expressivity and the red line represents the language stability. Where, N_M=1, N_I=1, B=50, E=100, M_T=1, I_T=0. The values are the average over 30 runs of the model.

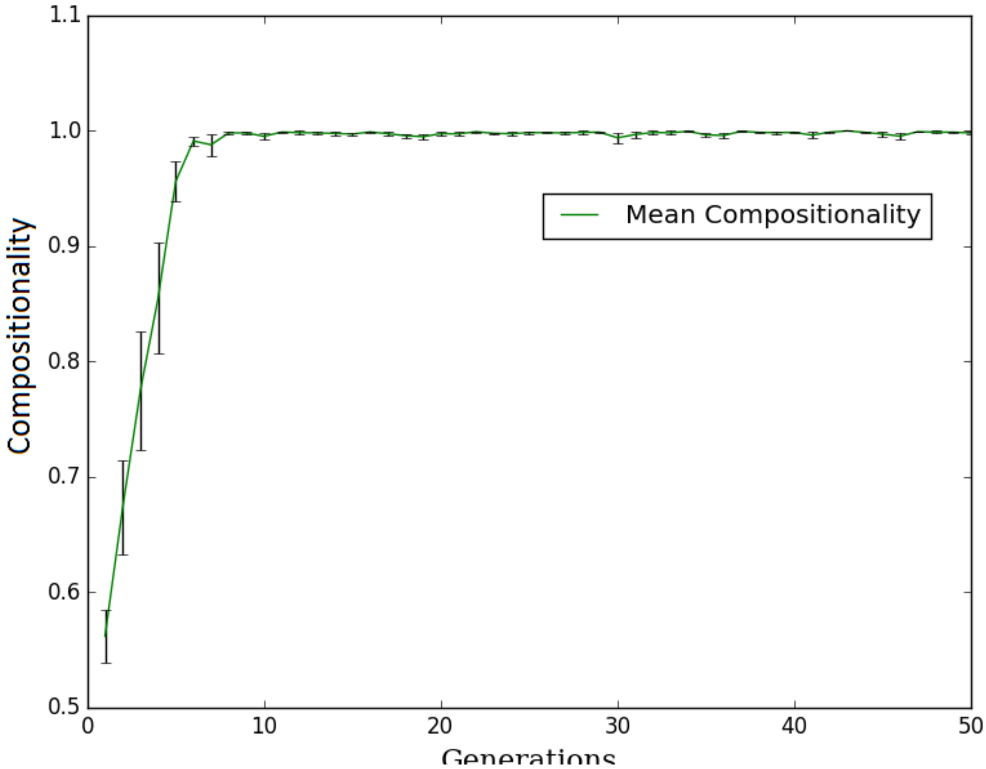


Figure 4.5: Langauge compositionality, over time for the ILM replication, where N_M=1, N_I=1, B=50, E=100, M_T=1, I_T=0. Here, the values are the average over 30 runs of the model.

4.5 Population-based Iterated Learning[6]

As stated in chapter 3, population structure has very rarely been taken into account in E/I models[7]. Given that iterated learning models aim to shed light on the relationship between the properties of individuals and the population-level behaviour that they exhibit, combined with the fact that much of the work done in this area thus far has been concerned primarily with vertical cultural transmission, it is of significant interest to explore the behaviour of this particular ILM within a population of agents.

The impact of population dynamics have been modelled previously within an iterated learning context. Griffiths (2007), explored iterated learning dynamics within a model where learning algorithms were based on the principles of Bayesian inference. By extending his framework to a population of such Bayesian agents, where each language user learns from a single member of the previous generation, he showed that iterated learning in this population produced language outcomes that could be understood as the result of solely the agent's individual learning biases. Therefore negating the role of other constraints, such as the transmission bottleneck.

---

[6] The following section draws from Brace et al. (2015).

[7] Kirby & Hurford (2002) themselves point out that complex population dynamics were traded off for computational power in the original model.

However, Smith (2009) argues that Griffiths's (2007) findings imply that it is possible to understand the prior biases of learners by looking at the typological distributions of languages. To support his argument, Smith (2009) presents a model of Bayesian agents in order to demonstrate that Griffiths's (2007) results are based upon the idealisation that a language user learns from a single teacher, and once multiple teachers are included, the mapping from the learner biases to typology breaks down. Based upon this result, Smith (2009) concludes that inferring learning bias from typology could yield unsafe results. Furthermore, Griffiths' (2007) model is limited by the fact that the agents use very specific statistical learning algorithms. They are therefore not applicable to cases where the subjects of study use more general-purpose learning algorithms, which are more akin to the general purpose cognitive architecture that is likely to underpin human language (Hurford 2014).

In a later work, Burkett & Griffiths (2010) explored the problems raised by Smith (2009) in developing a model where Bayesian agents were allowed to learn multiple languages. In doing so, they demonstrate that, so long as an agent's hypothesis space explicitly takes into account the possibility of receiving input from multiple speakers with potentially different languages, then Bayesian learning does tend to reflect the learner's inductive biases in the same manner as the single teacher model presented in Griffiths (2007). However, this model still makes the simplifying assumption that agents only receive input from vertical transmission. This is clearly not the case for real-life language users, who are likely to learn from their immature peers as well as from their mature role-models.

Thus, in the model presented here, each iteration consists of a population of N language users, comprising of N_M mature individuals and N_I immature individuals; where $N = N_M + N_I$. During each iteration of the model, every immature language user is assigned a number of trainers from whom they infer the structure of their language through a series of training episodes. This set of trainers may involve both a number of randomly chosen mature trainers, M_T, and also a possible number of randomly chosen immature trainers, I_T (fig. 4.6). The presence of immature trainers represents scenarios in which language learners are not kept isolated from one another, but may influence one another's language learning.

An immature individual's total number of training episodes, T, is evenly split between their trainers; with each trainer being involved in $\frac{B}{M_T+I_T}$ episodes per training epoch[8]. As in the original ILM, it remains the case that the total number of training episodes, T, is the product of the bottleneck size, B, and the number of training epochs, E. Hence, $T = B \times E$. The training episodes involving a specific trainer will involve the same set of randomly selected meanings in each training epoch. The set of B training episodes that comprise a single epoch are encountered in random order.

---

[8] Fractional numbers of training episodes are avoided by rounding up.

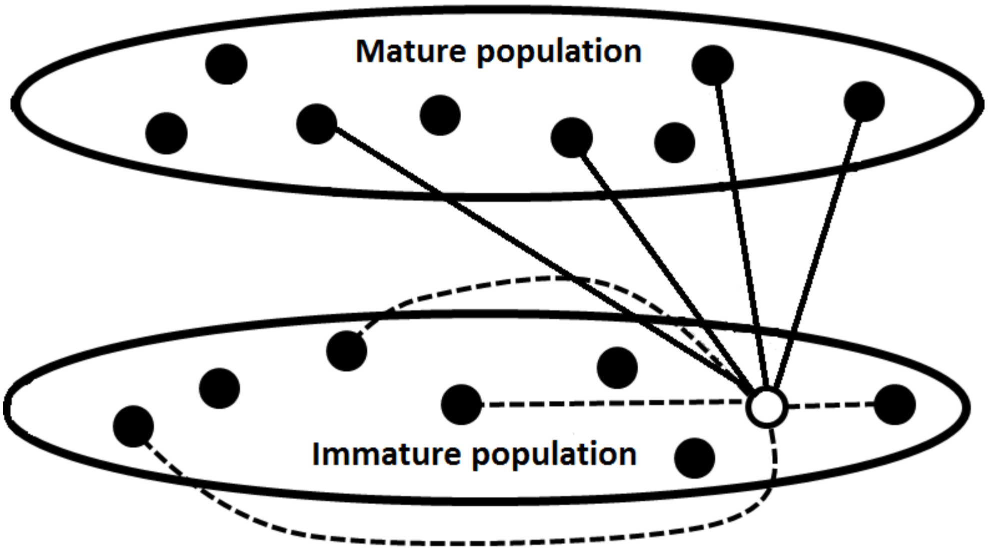


Figure 4.6: Diagrammatic representation of an ILM population divided into mature (upper set) and immature (lower set) agents, with $N = N_M + N_I$ agents per generation. Lines represent one immature agent's trainers: four mature trainers ($M_T = 4$, solid lines) and four immature trainers ($I_T = 4$, dashed lines).

4.6 Population-based Iterated Learning Model Results[9]

Figure 4.7 depicts a cross section of possible combinations of M_T and I_T, and how expressivity, X, and stability, S, develops in this population-based iterated learning model (PILM). In comparing figure 4.7A with figure 4.3, it is clear that training input from multiple mature agents has a significant impact upon the number of generations required for a fully expressive and stable communication system to arise. Unsurprisingly, given the nature of iterated learning, figure 4.7B shows how the system fails to improve above the scores obtained by random chance when M_T=0. Figures 4.7C and 4.7D depict how the system is able to produce a largely expressive and stable system when both M_T and I_T are set equal at 5 and 10, respectively.

To further explore the impact of multiple mature trainers on model behaviour, a series of tests were conducted with the aim of exploring the linguistic bottleneck. In figure 4.8, we see the result of different bottleneck sizes upon compositionality in a population. Here, I_T=0 and E=50; meaning that agents get half of the training sessions that they did in the original model, which should make learning far more difficult. The top graph of figure 4.8 shows the $M_T = 1$ case, while the bottom shows the $M_T = 10$ case. In both graphs, it can be seen that, unsurprisingly, the system does not learn when the bottleneck is set too low, the system does, unsurprisingly, not learn. When agents learn from only one mature trainer, a bottleneck of at least 80 meanings is required before a reasonably compositional language can survive. However, with ten mature trainers, a high level of compositionality can arise and survive with a much smaller bottleneck of

---

[9] Some of the following results are taken from Brace et al. (2015).

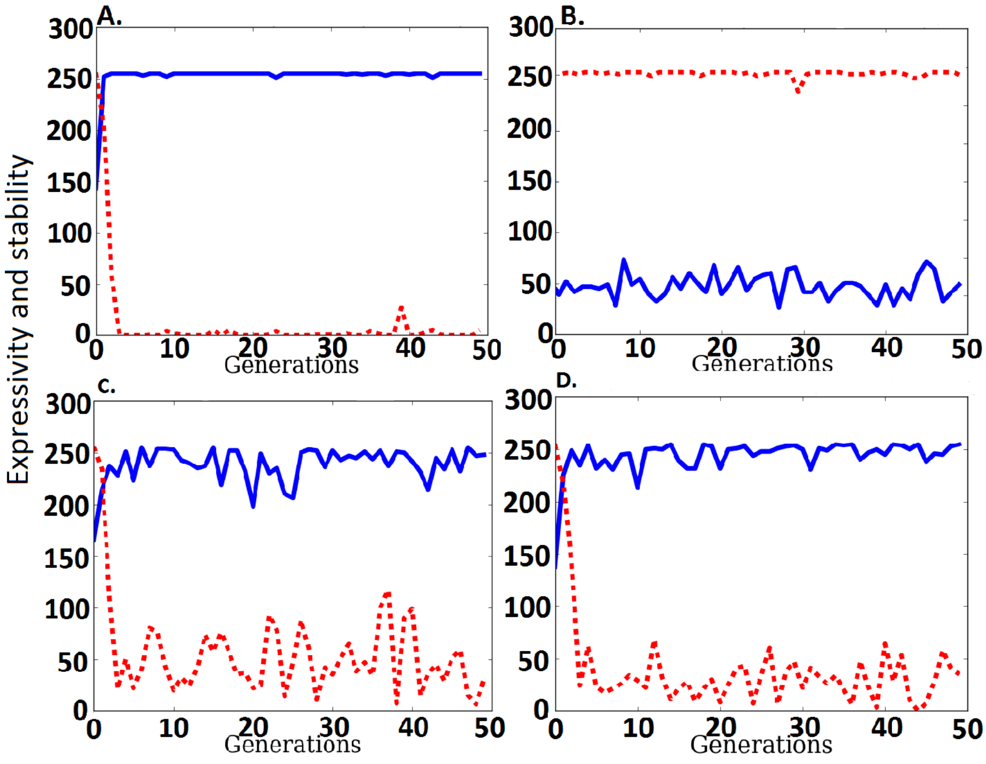


Figure 4.7: System behaviour for a single run of the PILM for various combinations of M_T and I_T. The blue lines depict expressivity and the red lines represents stability. Parameter settings for each: A. $M_T = 10$, $I_T = 0$; B. $M_T = 0$, $I_T = 10$; C. $M_T = 5$, $I_T = 5$; D. $M_T = 10$, $I_T = 10$; where N_M=15, N_I=15, B=50, and E=100 for all.

around 50. Moreover, when a compositional language arises, it does so at an increased rate when multiple trainers are present.

Figure 4.9 depicts analogous results for scenarios in which immature language users are allowed to influence each others' learning ($I_T = 5$). When immature trainers outnumber mature trainers (figure 4.9 top), language learning is compromised, with compositionality varying erratically over successive generations. This is due to the immature trainers, who are not yet fully linguistically competent themselves, essentially adding noise to the language-learner's input data. Despite this, it is notable that bottleneck size does influence the language; with larger bottlenecks allowing languages to achieve somewhat higher compositionality.

When immature trainers are outnumbered by mature trainers (figure 4.9 bottom), language learning is successful for scenarios with larger bottleneck sizes. Although compositionality does vary more from generation to generation in comparison with an equivalent scenario without immature trainers (compare figure 4.8 bottom). This is due to the fact that increasing the bottleneck size while allowing immature agents to interact with one another results in immature agents receiving extra signal diversity in their input training data.

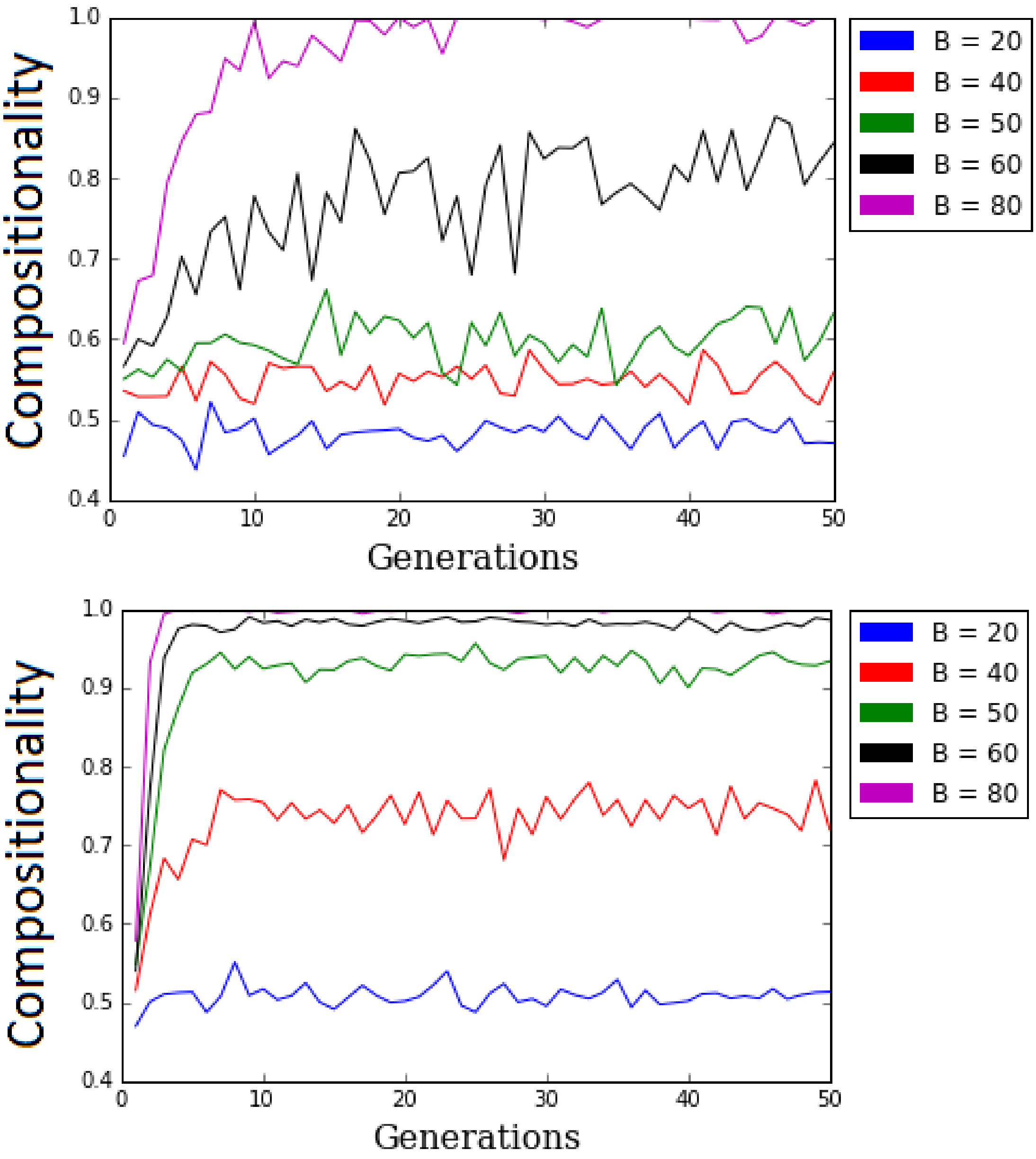


FIGURE 4.8: Graph depicting the impact of various value of B. *Top*: $M_T = 1$, with $B = 20$ (blue line: $\mu = 0.49$, $\sigma = 0.04$), $B = 40$ (red line: $\mu = 0.52$, $\sigma = 0.04$), $B = 50$ (green line: $\mu = 0.56$, $\sigma = 0.04$), $B = 60$ (black line: $\mu = 0.58$, $\sigma = 0.05$), and $B = 80$ (magenta line: $\mu = 0.65$, $\sigma = 0.06$). *Bottom*: $M_T = 10$, with $B = 20$ (blue line: $\mu = 0.52$, $\sigma = 0.04$), $B = 40$ (red line: $\mu = 0.63$, $\sigma = 0.05$), $B = 50$ (green line: $\mu = 0.80$, $\sigma = 0.06$), $B = 60$ (black line: $\mu = 0.90$, $\sigma = 0.08$), and $B = 80$ (magenta line: $\mu = 0.97$, $\sigma = 0.07$). With N_M=15, N_I=15, $I_T = 0$ and E=50 in both cases. Averaged over 30 runs.

Further evidence of the number of trainers having an impact on system behaviour can be seen in figure 4.10, which plots the average level of compositionality that the system exhibits over 50 generations for various combinations of M_T and I_T. In line with the above, it can be seen that compositional language tends to arise to the extent that the number of mature trainers is greater than the number of immature trainers, and that a

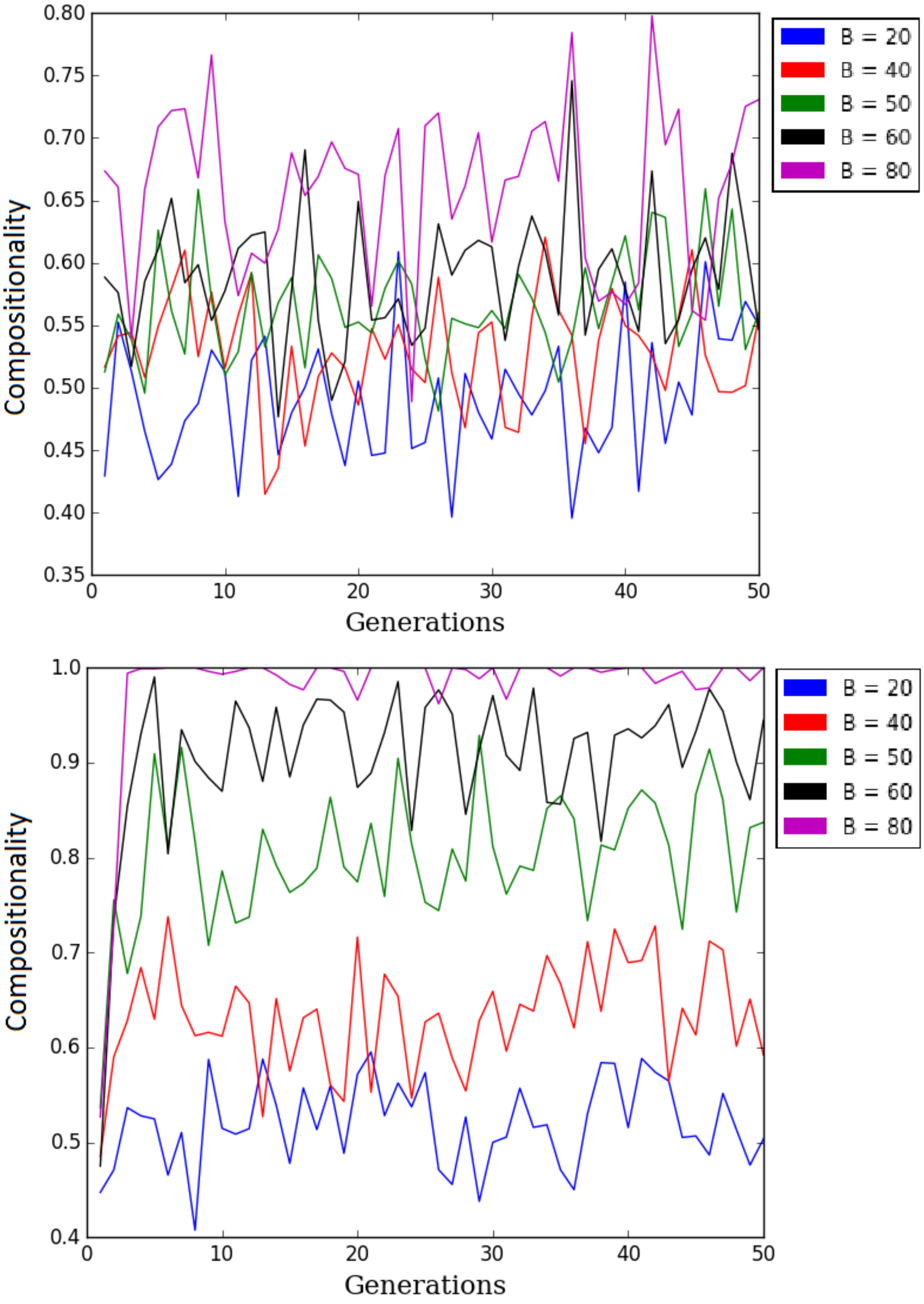


FIGURE 4.9: Graph depicting the impact of various values of B. *Top*: $M_T = 1$, with $B = 20$ (blue line: $\mu = 0.51$, $\sigma = 0.04$), $B = 40$ (red line: $\mu = 0.62$, $\sigma = 0.06$), $B = 50$ (green line: $\mu = 0.65$, $\sigma = 0.09$), $B = 60$ (black line: $\mu = 0.73$, $\sigma = 0.07$), and $B = 80$ (magenta line: $\mu = 0.96$, $\sigma = 0.08$). *Bottom*: $M_T = 10$, with $B = 20$ (blue line: $\mu = 0.58$, $\sigma = 0.05$), $B = 40$ (red line: $\mu = 0.76$, $\sigma = 0.05$), $B = 50$ (green line: $\mu = 0.92$, $\sigma = 0.05$), $B = 60$ (black line: $\mu = 0.98$, $\sigma = 0.04$), and $B = 80$ (magenta line: $\mu = 0.99$, $\sigma = 0.03$). With N_M=15, N_I=15, $I_T = 5$ and E=50 in both cases. Averaged over 30 runs.

greater number of mature trainers enable the system to develop and maintain a higher level of language compositionality.

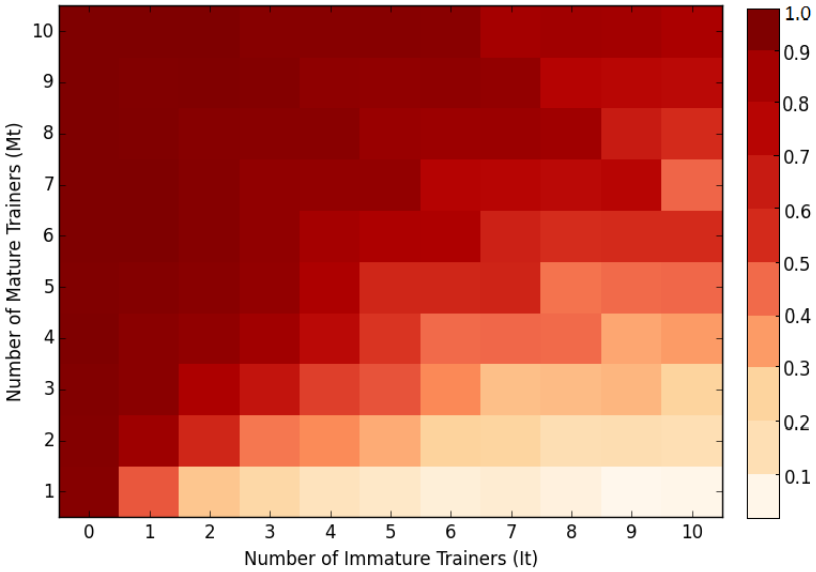


FIGURE 4.10: Heatmap of the average amount of compositionality over 50 generations, where N_M=15, N_I=15, B=50, and E=100, throughout.

Why might dividing the same number of learning episodes between a greater number of mature trainers lead to improved learning in language users? The traditional wisdom of those that have worked with E/I models is that they essentially simulate a standardisation in usage of meaning-signal pairings being established between agents. Thus, it follows that, in cases where $M_T = 1$, it takes a number of generations for an expressive and stable communicative system to arise (figure 4.3), due to the way in which there are multiple signals being mapped to single meanings during the early generations. Intuitively then, a situation with $M_T > 1$ would exacerbate this problem, due to the fact that such a scenario would see the various mature trainers behaving the same way, with the agents in the next generation being trained on multiple degenerate languages; due to the input data coming from multiple mature agents, each of which is producing its own version of a degenerate language. This factor should make it much more difficult for a conventionalised system to establish itself.

However, this appears to not be the case. Moreover, the results presented here demonstrate that, while the mature agents in the initial generation may indeed use multiple signals for the same meaning, when language-learners learn their 'language' from multiple tutors, such inconsistencies in the language are coalesced by the immature agents during their language training due to the obverter procedure. Thus, when the initial immature population is promoted to mature agent status, they receive varied, but similar input during their training episodes; meaning that they all posses similar weights. This, in turn, results in them producing linguistic output that is similar to one another; output that is then used to train the second batch of immature agents. This is why figure 4.11 shows that the immature agents in the second generation tend to have more unique signals in their training data when they learn their language from a higher number of different mature agents. Ergo, we see that a higher number of mature trainers results in a fully expressive, compositional, and stable language emerging in far fewer generational steps, than in the original model presented by Kirby & Hurford (2002).

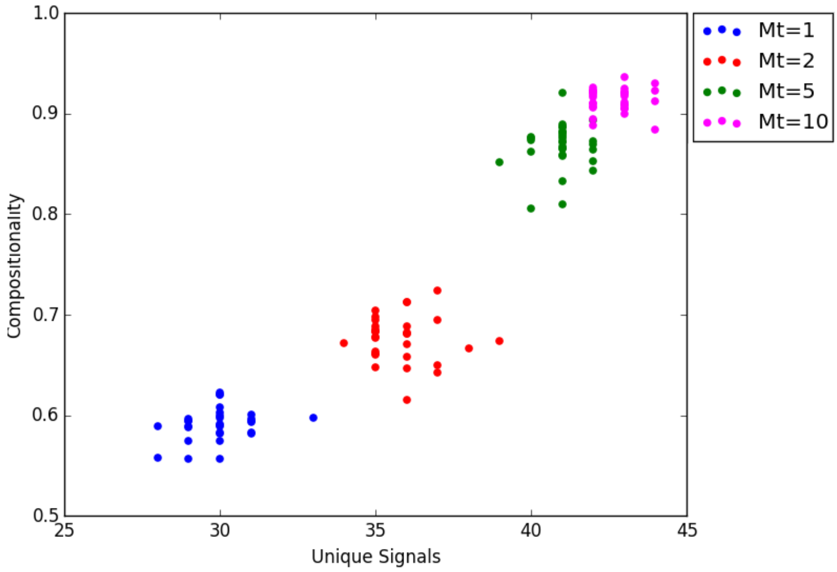


Figure 4.11: Graph depicting the average number of unique signals in the training sets for all immature agents at generation 2 and the average level of compositionality at generation 5 for $M_T \in \{1, 2, 5, 10\}$, where N_M=15, N_I=15, B=50, and E=100.

4.7 Conclusions

It has been shown here that our understanding of language can benefit from viewing human language as an emergent phenomenon of the three interacting complex systems of biology, culture and ontogeny. More specifically, it has been demonstrated that language learners can benefit from learning signals from multiple trainers due to the way in which they see that multiple signals can be used for a single meaning. However, due to the way in which the agent networks and obverter procedure operate, they are quickly able to pick the most commonly used signal for such a meaning.

These results provide an encouraging first step in our aim to use computational methods in order to explore the relationship that researchers using different methodologies belief to exist between population structure/social dynamics and linguistic form. In so much that it has shown that language learners can benefit from learning signals from multiple trainers.

In regards to E/I models more generally, such results can be seen to support the myriad of computational work that has demonstrated how cognitively rather sophisticated agents can successfully map a structured meaning space to a structured signal space in such a manner as to give rise to, and learn, a compositional language (Steels 1995, Hurford 2000, Kirby 2002*c*, Kirby & Hurford 2002, Barr 2004, Belpaeme & Bleys 2007, Steels 2007, Smith 2007).

More interestingly, traditional wisdom within the field of E/I research is that these models essentially simulate the establishment of a standardisation of usage of meaning-signal pairings between agents; meaning that having more mature speaker agents in the simulation would lead to different signals being used for the same meaning. Thus making it more difficult for a conventionalised system to establish itself. However, this has been shown to not be the case, and although the mature agents in the initial generation may indeed use multiple signals for the same meaning, when an immature agent learns its 'language' from multiple tutors, such inconsistencies in the language are coalesced by the immature agent during their language training. This, in turn, results in the immature agents producing linguistic output that is similar to one another; output that is then used to train the second batch of immature agents. This is why we see that a higher number of unique signals appearing in the training input for the second batch of mature agents in figure 4.11. Thus meaning that a higher number of mature trainers results in a fully expressive, compositional, and stable language emerging in the population-based iterated learning model much faster than in the original model presented by Kirby & Hurford (2002).

However, this work can be criticised from both a theoretical and technical stand point. Technical criticisms take the form of modelling choices in the implementation of the model. The agents within this ILM are feed-forward neural networks which use back-propagation. The first issue here is that standard multi-layer neural networks require the length of input data to be the same as that of the output data, but not all human utterances are the same length[10].

Secondly, as Turner (2002) observes, the neurological plausibility of the back-propagation algorithm is questionable for a number of reasons. It is not clear that synapses can transmit error backwards, in that we rarely have the opportunity to quantify our errors by comparing our actual behaviours with target behaviours; and, unlike real-world learning, back-propagation requires exposure to a vast number of examples.

Thirdly, there is a question as to whether the behaviour seen in the data is a genuine reflection of language dynamics, or a result of specific neural network behaviour; such as the way in which the agents coalesce multiple inputs into a single output.

Finally, the obverter procedure itself was designed to implement a bias towards an alignment between interpretation and production (Hurford 1989, Franke 2015); meaning that it produces behaviour that specifically aims to create efficient communication.

---

[10] This problem could potentially be overcome through the use of recurrent neural networks, which allow such sequential processing (Elman 1990, Turner 2002).

In conclusion, the results presented here provide a fascinating starting point for our exploration into the behaviour of E/I models and the impact of population dynamics upon linguistic form. However, there is a possibility that these results were obtained purely as a result of some idiosyncrasies of the model or its algorithm. This possibility will be explored in the next chapter.

Chapter 5

Convention Emergence Through Reinforcement Learning

5.1 Introduction

It was demonstrated in the last chapter that a population of agents learning their language from multiple mature trainers enables a compositional linguistic system to be established with greater ease. However, given the rather sophisticated algorithm employed by the PILM, it would arguably be beneficial to explore whether those results are generalisable. As such, this chapter will explore the same ideas, but with cognitively simplistic agents by looking at how the meanings of lexical items change over time, which is indicative of cultural factors playing a crucial role in the shaping of human language.

The establishment of the meanings of lexical items and the subsequent change in these meanings is in part what led Lewis (1969) to work on the conventionality of meaning. He introduced a signalling game in order to explore how meaningful language might evolve from the use of initially random signals. The motivation here was to demonstrate how interactions between rational individuals could give rise to conventions in meaning and usage with the aid of common understanding of other player's rationality and knowledge of the pay-offs.

Renewed interest in these ideas over recent years have focused on how such conventions can emerge from mindless evolutionary processes, and this has led to a body of work that has explored the evolution of term-based languages through coordination games (Skyrms 2004, 2009, 2010, Huttegger 2007, Barrett 2006, 2009, Argiento et al. 2009).

5.2 The Lewis Signalling Game[1]

In a Lewis signalling game there are two players, a sender and a receiver. A single bout of the game involves the sender knowing that the world is in some random state, t, but the receiver being ignorant of this information. The sender then selects a signal, s, with which to convey the world state to the receiver. The receiver observes s and has to pick an appropriate action, a. If the action chosen by the receiver matches the world state (i.e., $a = t$), the bout is considered to have been a success. In other words, the communicative episode is deemed to be a success if the hearer can accurately determine the world state, t, from the signal, s produced by the speaker. Here, t, s, and a are drawn from finite sets T, S, and A, respectively. All of these finite sets are of size n; in Lewis' (1969) original model $n = 2$.

Over successive bouts of the game, both players are expected to adapt their behaviour in order to increase the chance of achieving communicative success; typically through some kind of reinforcement learning.

The easiest way to conceptualise this is in terms of urns and balls. At the outset of the simulation run, an unbiased sender will have n urns, one for each state of the world; each of which will contain n balls, one associated with each of the n possible signals. Let's suppose that during the first bout of the game, $t =$ "*red*". The sender picks a random ball from their *red* urn. The symbol on this ball dictates the signal to be made, s; in this case, suppose $s =$"fah". Likewise, the receiver observes $s =$ "fah", and picks a random ball from their *fah* urn, which indicates the action to be taken, a. Both balls are then returned to their respective urns. If $a = t$, the interaction was a success, and in accordance with the principles of Roth-Erev reinforcement learning (Roth & Erev 1995), the sender adds extra balls of type s to urn t and the receiver adds extra balls of type a to urn s. The number of extra balls added to the urns, u, corresponds to the utility associated with the outcome of the signalling bout; in Lewis' (1969) original game $u = 1$ if a bout is successful and $u = 0$ otherwise.

More formally, at any point in time, $b(t, s)$ is the number of balls for signal s in the sender's urn for state t, and accordingly, $b(s, a)$ is the number of balls corresponding to act a in the receiver's urn s. Thus, the behavioural strategies for sender (σ) and receiver (ρ) are as follows:

$$\sigma(t, s) = \frac{b(t, s)}{\sum_{s' \in S} b(t, s')} \qquad \rho(s, a) = \frac{b(s, a)}{\sum_{a' \in A} b(s, a')} \tag{5.1}$$

Essentially, these equations describe the probability that an agent will produce a specific signal for a given meaning. At the start of its lifetime, each agent has an equally chance of producing any of the possible signals for any meaning presented to it. However,

---

[1] The following section is drawn from Brace & Bullock (2016).

throughout the course of its interactions with other agents, the probability of said agent producing the same signal every time it is presented with a specific meaning increases every time the agent uses the signal in question to successfully convey a specific meaning, world state, to another agent. Likewise, for the probability of an agent producing a specific action upon being presented with a signal.

There are a number of possible signalling equilibria that can arise in such a game. Perfect signalling strategies result in optimal pay-offs for the players by mapping each world state onto a unique signal and each signal onto the unique appropriate action (figure 5.1). This behaviour constitutes an evolutionarily stable strategy (ESS) because, when it is played by the whole population, there is no incentive for any individual to change their strategy.

However, players may spend significant time playing sub-optimal 'pooling' strategies, in which senders employ the same signal for multiple world states (pooling these world states together). Thereby making it impossible for receivers to determine the state of the world from the signals that they receive (figure 5.2). Pooling strategies are not ESSs since adjacent strategies often achieve equal fitness; i.e., the two pooling strategies in figure 5.2. The expected pay-off for such a pooling strategy is 0.5 when $n = 2$.

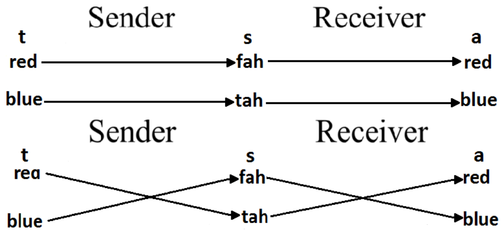


FIGURE 5.1: Optimal strategies for the $n = 2$ game.

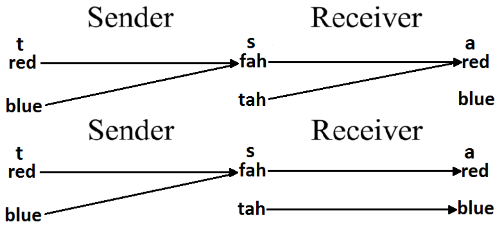


FIGURE 5.2: Two of the possible sub-optimal pooling strategies for the $n = 2$ game.

It has been shown by way of both computational simulations (Barrett 2006, 2009, Skyrms 2010) and mathematical modelling (Huttegger 2007, Argiento et al. 2009) that the $n = 2$ game will nearly always converge upon an optimal signalling system. Indeed, Skyrms (2009) went on to demonstrate that this behaviour also holds in cases where there are two senders and one receiver. These results are further supported by Table 5.1, where it can be seen that a replicated computational model of the $n = 2$ game being played for 10^6 bouts will almost always reach a perfect signalling equilibrium.

5.2.1 Higher-n games

However, a successful outcome is not always achieved when the game is played with $n > 2$; i.e., with a higher number of states, signals, and actions (Skyrms 2010, Huttegger 2007, Barrett 2006, 2009).

In adopting the methodology of Barrett (2006, 2009), the R-L model that formed the basis of the population-based R-L model was run multiple times for various values of n. Each run consisted of 10^6 bouts, B, of the game. A run of the simulation is considered to fail if the number of successful bouts is less than 90% of the total number of bouts.

Table 5.1 (left) shows the results of these runs, which agree with those of Barrett (2006, 2009). Table 5.1 (right) shows the results of a smaller sample of 100 runs, with all other parameters being held constant[2]. It can be seen clearly from Table 5.1 that, in a n=3 game, the players fail to achieve a high enough rate of signalling success roughly 10% of the time, and that this increases to ≈ 20% for n=4 games, ≈ 60% for n=8 games, and so on.

1000 runs	
	Success
2-n	0.999
3-n	0.881
4-n	0.784
8-n	0.391
10-n	0.281
20-n	0.264

100 runs	
	Success
2-n	0.99
3-n	0.87
4-n	0.84
8-n	0.33
10-n	0.22
20-n	0.23

TABLE 5.1: Table depicting the success rates of the replicated R-L model after 10^6 bouts for various values of n, with 1000 runs (left) and 100 runs (right).

5.3 Generational Population-based Reinforcement Learning Model[3]

Although interesting in their own right, the dyadic setting considered so far limits the conclusions that can be drawn from such models. After all, human language persists in a highly complex social milieu, and it has been shown that the structure and composition of a population can influence the dynamics of language change over time (Brace et al., 2015; chapter 4).

As such, the original reinforcement learning model (R-L) was extended in a number of ways. First, whereas the original model focused on a single sender and receiver playing

---

[2] The comparison in Table 5.1 is important to show, as the extended model presented later is run for 100 runs due to limits on computational power.

[3] The following section is drawn from Brace & Bullock (2016).

for B bouts, in the population model (R-L-P), there are a population of agents. This population is divided into a number of mature agents, N_M, and a number of immature agents, N_I. All agents start life as immature and then get promoted to mature status after the first epoch of their existence; in the same manner as the model from the previous chapter.

Mature agents play bouts of the language game with one another, updating their language behaviour according to game outcomes. By contrast, while agents are immature they merely observe the language bouts played by their mature parent, and update their language behaviour on the basis of the outcomes of these observed games. The lifespan of agents is two epochs, the first as an immature agent and the second as a mature agent; after which they are removed from the simulation.

It is important to emphasise here that, throughout the simulation, when new immature individuals are added to the population, as in the standard R-L model, they have no knowledge of the language currently being used. This is true for the initial population of mature agents, and also true for new immature agents born into all subsequent epochs. For each immature agent, each world state, t, is associated equally with each signal, s, when playing as Sender, and each signal, s, is associated equally with each action, a, when playing as receiver. In other words, each of an immature agent's n state urns contain a single ball for each possible signal, and each of their signal urns contain a single ball for each possible action. Thus, any change in communicative performance or language use over generations is the result of language evolution; there is no biological evolution on the part of the agents.

Furthermore, instead of agents merely interacting B times, the R-L model is extended to include a generational aspect. In other words, the model is set up to run for a number of epochs, E. During each epoch, every mature agent plays B bouts as the sender with other mature agents; with the amount of bouts it plays as the hearer being the result of how many other agents it is partnered with, divided by B. The number of different mature agents that a mature agent interacts with, P, is a key parameter of the model. Each mature agent's total number of interactions, B, being equally divided amongst its P unique partners, i.e., the number of interactions that a mature agent plays as a hearer is the result of B/P (rounded up).

The R-L-P model thus proceeds as follows. At the start of the simulation run, an initial population of $N_M = 15$ unbiased mature agents are created, with an equal chance of generating each signal for each world state. For each epoch, E, a fresh population of $N_I = 15$ unbiased immature agents is created, each having an equal chance of generating each signal for each world state. Each immature agent is assigned a randomly selected mature agent to act as their parent. Each mature agent is assigned P unique randomly selected mature partners with which to play the signalling game. Each of the mature agents then engages in B/P bouts with each assigned partner, with each participant updating their signalling or receiving strategy at the end of each bout through reinforcement learning. Each child will update their behaviour based on the outcome of the bouts that their parents are involved in; i.e., at the end of a successful bout, a sender's child will add a ball of type s to urn t, and a receiver's child will add a ball of type a to urn s. At the end of an epoch, all mature agents are removed, all immature agents are promoted to mature agent status, and a new set of unbiased immature agents are created for the next generation.

5.4 Genertaional Population-based Reinforcement Learning model Results[4]

The R-L-P model does not achieve a successful signalling system as often as the standard R-L model. Indeed, comparing table 5.2 to table 5.1 shows how, with $P = 1$, overall success rates are lower for all n-games than in the standard R-L model.

	Partners = 1	Partners =2	Partners = 5	Partners = 10
4-n	0.3	0.36	0.77	0.81
8-n	0.0	0.3	0.48	0.62
10-n	0.0	0.2	0.20	0.46
20-n	0.0	0.0	0.29	0.57

TABLE 5.2: Table depicting the success rates after 20 epochs for 100 runs of the R-L-P model for various values of N and P.

However, increasing the value of P does increase the rate of success (table 5.2 and figure 5.3). In figure 5.3, with $P = 1$ or 2, there is an initial level of success, which corresponds to the number of successful bouts that would be seen in the normal R-L model for a n=20 game, but this then drops during generation 2.

Here, low P values create a situation whereby mature agents form a communicative system based on conventions agreed upon between themselves and only a small number of other agents. Thus, in subsequent epochs, when the child of a mature agent has to interact with the child of another mature agent, who has not previously interacted with the mature agent in question, the agreed upon conventions that both parties formulated during the first epoch are likely to be of little use; due to different agents forming conventions based upon their idiosyncratic experiences. This gives rise to sub-optimal behaviour at the population-level.

However, any immature agents that are present learn from the successful bouts of their respective parents; hence the steady increase in success rates for these lower P values[5]. In contrast, with high P values, we see an obvious and immediate increase in communicative

---

[4] The following is drawn from Brace & Bullock (2016).

[5] Given enough epochs, it is likely that the agents would give rise to a successful communicative system.

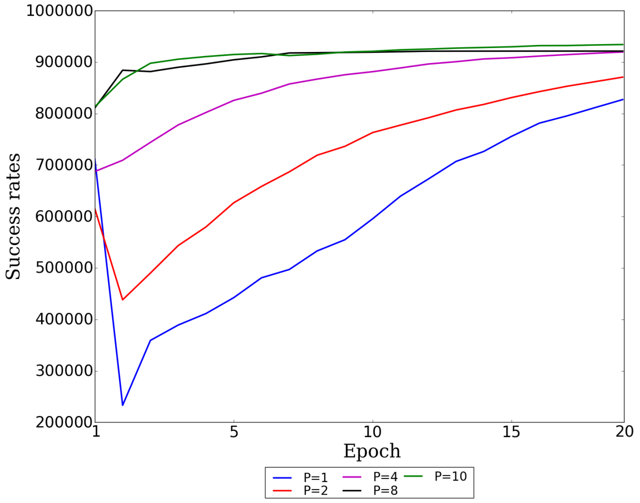


FIGURE 5.3: Graph depicting the average number of successful bouts across epochs for $P = 1$ (blue line: $\mu = 599028.9$, $\sigma = 12176.42$), $P = 2$ (red line: $\mu = 711869.4.9$, $\sigma = 14145.25$), $P = 4$ (pink line: $\mu = 857588.7$, $\sigma = 12754.06$), $P = 8$ (black line: $\mu = 906913.7$, $\sigma = 8977.71$), and $P = 10$ (green line: $\mu = 902448.8$, $\sigma = 14507.5$) for a $n = 20$ game with $N_M = 15$, $N_I = 15$, $B = 10^6$, and $u = 1$. Averaged over 30 runs.

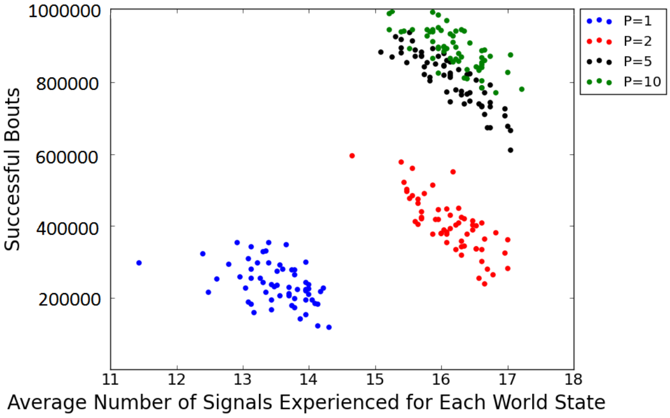


FIGURE 5.4: Graph depicting the average percentage of successful communicative bouts between all mature agents plotted against the number of unique signals presented to them during said bouts in the second epoch for a $n = 20$ game, for P=1, 2, 5, 10 and $N_M = 15$, $N_I = 15$, $B = 10^6$, and $u = 1$. Averaged over 60 runs.

success. This is due to the way in which an increase in P leads to the children of the mature agents having more diversity in their training input. This better enables these individuals to communicate with a larger number of other agents upon being promoted to mature agent status (figure 5.4). This greater chance of success then leads to this first batch of immature agents essentially coalescing the meaning-signal pairings; again, similar to the model from the previous chapter.

Imagine a hypothetical mature agent from epoch one, who is partnered with ten other randomly selected agents; who in turn, are partnered with ten other agents. In the simulation, bouts are scheduled in such a way that $agent_1$ will have one of the allocated

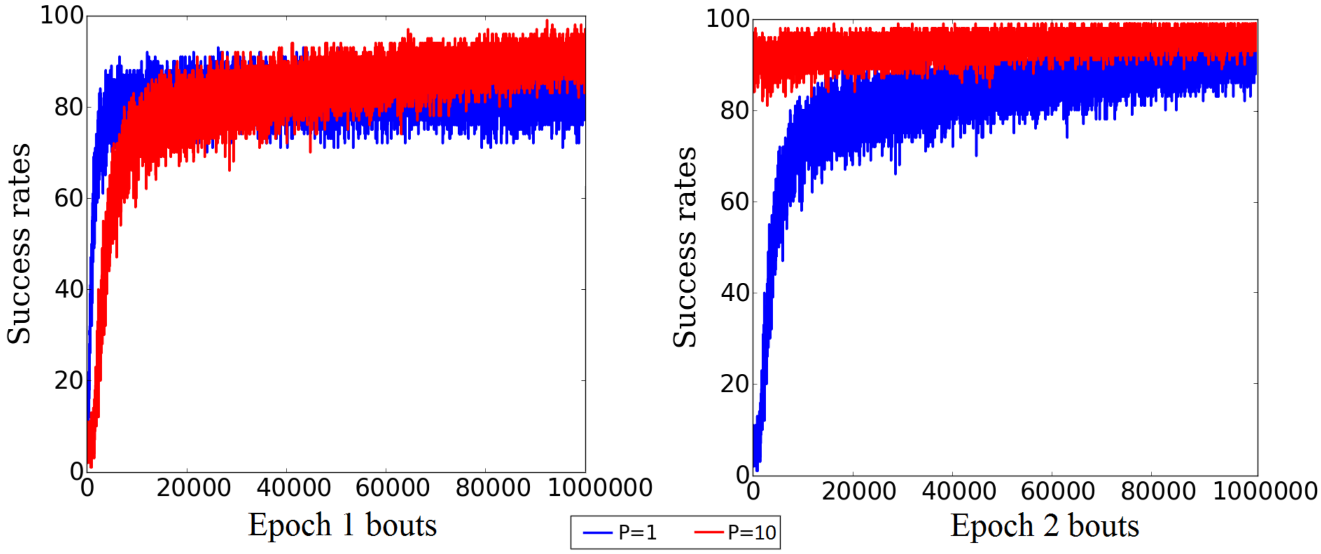


FIGURE 5.5: Graph depicting the number of successful bouts out of every 100 bouts, over all 10^6 bouts of a random agent during the first epoch (left) and the second epoch (right); with P=1 (blue line) and P=10 (red line). Where $n = 20$, $N_M = 15$, $N_I = 15$, $B = 10^6$, and $u = 1$.

bouts with one of its randomly selected partners, then $agent_2$ will do the same; and so on, until we reach $agent_{M_N}$. At which point we go back to $agent_1$ and allow it to have its second bout, again with a randomly selected partner; and so on until each partner of every agent has played B/P bouts with the agent.

In the P=1 case, unsurprisingly, we see higher levels of initial success during the first generation than in the P=10. This is due to the establishment of a convention involving fewer agents having to negotiate with one another (figure 5.5, left).

In contrast, with $P = 10$, it is slightly harder to establish a conventionalised usage because each agent has to negotiate with an increased number of different agents, which results in higher levels of signal diversity (figures 5.4 and 5.5, left). However, when the offspring of the first epoch's mature agents are forced to interact with a different subset of the population in the second epoch, populations with higher P values exhibit higher communicative success due to the increased signal diversity in the previous epoch. This is the result of the immature agents learning from the successful bouts of their parents. In turn, this results in these agents establishing a conventionalised usage that requires less renegotiating when speaking to previously unencountered agents than in the $P = 1$ case; where agents have a more idiosyncratic language that requires them to renegotiate the conventions established by their parents (figure 5.5, right).

This is why figure 5.4 shows an increase in communicative success with higher values of P, while also indicating a negative trend in each of the data clusters for each specific P value. Although it is harder to establish a language when negotiating meaning-signal pairs with more individuals, doing so makes it more stable across generations (figure 5.5). Indeed, as figure 5.5 (right) shows, the agreed upon convention of usage in cases of lower P values has to be renegotiated in subsequent epochs due to it offering little communicative success to agents when communicating with newly encountered individuals.

It is important to note that the increase in communicative success is the result of higher P values and not of another variable, such as B. Indeed, figure 5.6 demonstrates the average level of communicative success over twenty epochs is significantly lower for $P = 1$ or 2, as compared to $P = 4$; a trend that continues as P is increased. Furthermore, it can be seen from figure 5.6 that higher P values allow for an increased amount of communicative success, even when agents have significantly fewer training sessions (lower B values).

5.4.1 Impact of Immature Language Users

In the real world, children are not just passive receivers of linguistic input. They interact with others; including other children, who may not yet be fully linguistically competent. Thus, a number of model runs were conducted where immature agents had B bouts with P other immature agents while witnessing their parents bouts (figure 5.7). These immature agent bouts are scheduled in a similar manner to the mature agent bouts described above, in that we allow each agent to have one bout with a randomly selected partner; starting with $agent_1$ and cycling through to $agent_{N_I}$, before going back to $agent1$ again. In these runs, mature agents only interact with mature agents and immature agents only interact with other immature agents. Although, immature agents still learn from their parent's interactions.

Figure 5.7 demonstrates how performance in the $P = 10$ case is impeded by allowing interactions between immature agents. This is to be expected, as linguistically underdeveloped individuals interacting with one another will add a degree of noise into the communicative system. However, with $P = 1$, allowing immature agents to interact with one another dramatically increases communicative success. This difference in behaviour can again be attributed to signal diversity. While in the above results, immature agents only learned from the interactions of their parents, meaning they got a degenerative sample of the language because they only ever witnessed the same two individuals communicating during their first epoch, here they are also interacting and learning with another individual who is likely to have witnessed two different mature agent's interacting with one another. This would increase the amount of signal diversity in the immature agents training data.

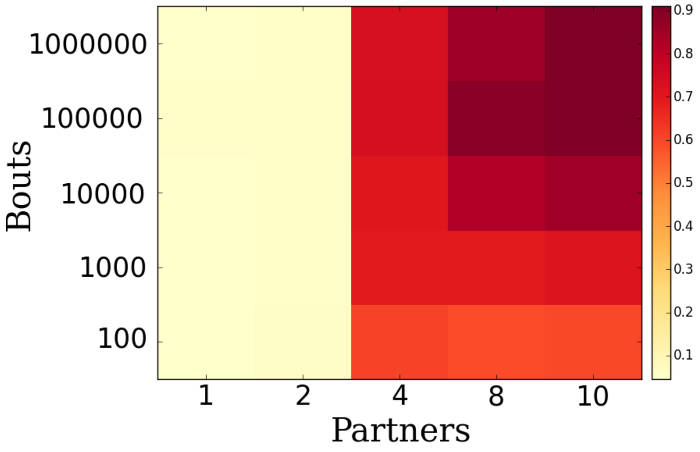


FIGURE 5.6: Graph depicting the average amount of communicative success over 20 epochs for various values of B and P. Where $n = 20$, $N_M = 15$, $N_I = 15$, $B = 10^6$, and $u = 1$. Averaged over 30 runs.

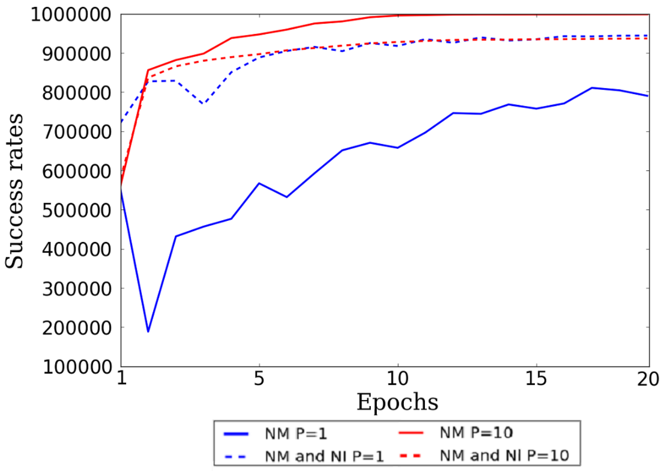


FIGURE 5.7: Graph depicting the average amount of communicative success over 20 generations for $P = 1$ where only N_M interact with one another (blue solid line: $\mu = 63342.1$, $\sigma = 15604.78$) and where both N_M and N_I interact with other mature and immature agents (blue dashed line: $\mu = 89468.85$, $\sigma = 6159.112$), and likewise for $P = 10$ where only N_M interact with one another (red solid line: $\mu = 94822.5$, $\sigma = 9862.42$ and where both N_M and N_I interact with other mature and immature agents (red dashed line:$\mu = 89468.85$, $\sigma = 6159.112$. Where $n = 20$, $N_M = 15$, $N_I = 15$, $B = 10^6$, and $u = 1$. Averaged over 30 runs.

5.5 Conclusions

The results presented here build upon a larger body of work, both in regards to signalling conventions (Skyrms 2004, 2010, Barrett 2006, 2009) and expression/induction model research in general (Hurford 2002). It has been shown that a signal can acquire a conventionalised meaning without the sender intending for it to do so, and that the meaning of such simple signals is dependent upon the stabilisation of usage conventions, which emerge from functional historical signal production. Thus, even the most automatic or reflexive signals can acquire meaning, so long as the production and response mechanisms are co-adapted to coordinate their behaviours in accordance with such an arbitrary signal (Harms 2004).

More interestingly, it has been shown that a population structure that allows for interactions between more of its members is beneficial in allowing it to evolve an efficient term-based language. Specifically, it has been shown that, as intuition dictates, while it is initially harder to establish a conventionalised system of usage with larger numbers of individuals, doing so enables the emerging language to persist in subsequent epochs. This is due to the input into language learners being initially more diverse, which prevents these learners from developing a more idiosyncratic communicative system that makes it harder to communicate with previously unencountered individuals.

In addition, the results reported here can be seen to be linked to the aforementioned concept of a linguistic bottleneck, which refers to how the input data for a language learner will only be a subset of the potentially large range of grammars of the language user from which it is learning. We have seen that ILMs have demonstrated how the linguistic bottleneck is crucially important in regards to whether or not language can be successfully passed from one generation to the next and, in situations where this transmission can be achieved successfully, show that it is also crucial to the linguistic structures that arise (Kirby 2002*b*,*a*, Kirby & Hurford 2002, Kirby et al. 2014, Smith 2002, Smith et al. 2003, Brace et al. 2015). Although a similar effect to the bottleneck is seen in uni-generational models, such as the *naming game* (Steels 1995)[6], the model presented here is novel in that it demonstrates the impact of bottleneck-like behaviour in a generational-based simulation that explores the use of term-based languages.

Here, this bottleneck-like behaviour takes the form of the way in which internal representations of individuals are induced from limited examples of the behaviour of other agents (Hurford 2002). This supports the results from chapter 4, which demonstrated a link between the linguistic bottleneck and the number of linguistic tutors (Brace et al. 2015).

Additionally, the behaviour seen in figure 5.7 indicates that the factors underpinning the cultural transmission of language change and linguistic variation are perhaps too complicated to be understood by analysing the nature of just inter- and intra-generational transmission; and that further research into linguistic change should focus on the makeup of the social group that underpins linguistic populations (Wichmann & Holman 2009, Lupyan & Dale 2010, Milroy 2013).

Indeed, it would be interesting to explore how an expanding and contracting population size, with varying numbers of mature language users and immature language learners,

---

[6] See chapter 3.

could impact the emergence and form of a language (Johansson 1997, Hurford 2002). An expression/induction model geared towards this interest could provide valuable insights for a growing body of research that is interested in the nature of the relationship between language and population change (Johansson 1997, Nettle 1999*a*, Wichmann & Holman 2009, Lupyan & Dale 2010, Milroy 2013, Trudgill 2013). These ideas will be the focus of the next chapter.

Chapter 6

Population Dynamics and Language Emergence

6.1 Introduction

Technological developments over the last couple of decades have resulted in a number of new methodologies becoming available to language evolution researchers; including the development of global linguistic databases that enable one to performs statistical analysis on linguistic structures of languages from across the world. These developments have led to a growing body of work that posits a relationship between changes in linguistic structure and population dynamics (Johansson 1997, Wichmann & Holman 2009, Lupyan & Dale 2010, Milroy 2013, Atkinson et al. 2015). Given the nature of expression and induction (E/I) models and their focus on the behaviour within populations of agents, computational simulations of this kind are well positioned to explore the nature of this theorised relationship. Especially given the results from the models in chapters 4 and 5 that indicate an intimate relationship between iterated learning dynamics and population structure.

However, the two models presented thus far have only looked at how linguistic structure has been affected by populations that have been static in nature. This is obviously not reflective of reality, where population sizes tend to be in a constant state of flux; often exhibiting periods of severe contraction of expansion. Furthermore, the previous two models have only explored how agents create a mapping between meanings and a pre-defined, and fixed, set of signals. As such, the model presented in this chapter will draw upon the reinforcement learning concept from chapter 5 and the iterated learning framework, specifically the iterated learning model as implemented by Kirby (2002*a*,*b*), in order to explore how agents create their own signals and syntactic rules and use these in order to convey meanings.

6.2 Linguistic Change And Social Groups

A number of scholars have attempted to explore links between aspects of societies and features of the languages spoken by the individuals therein; with much of this work focusing on aspects of culture (Trudgill 2013). In the vast majority of such work, the make-up of social groups are considered to be a determining factor in language change, contact, maintenance, and shift (Labov 2001, de Bot & Stoessel 2002, Ke et al. 2008).

Sociolinguists have produced empirical studies that focus on the social networks of small communities and the linguistic performance of the individuals within them (Milroy 1987). These have quantitatively demonstrated that the linguistic behaviours of individuals are highly correlated with their degrees of integration into the social group.

In other words, an individual is likely to have less variation in their language abilities, and will better conform to the speaking norms of the linguistic population, the more integrated they are within the network when linguistic variations are present within the social groups (Ke et al. 2008).

Some of this work has explored how differing social network structures and properties can explain the presence of language varieties that are specific to particular social groups. These network-orientated studies of linguistic change take the form of both sociolinguistic variationist work[1] and socio-historical studies of changes that occurred during the earlier stages of a language's life (Milroy 2013). However, for practical reasons, such empirical work tends to only look at small communities of individuals; and few have explored the effect of different network types upon linguistic change over longer historical periods (Ke et al. 2008). This is perhaps why only a few empirical studies have been able to quantitatively demonstrate that population make-up is an important factor in language change.

Another line of enquiry has also come about with the establishment of online linguistic databases, such as the World Atlas of Language Structures (WALS). Specifically, Bromham et al. (2015) demonstrates that there is the potential for demographic factors to influence language evolution, while Lupyan & Dale (2010) showed that languages that are spoken by larger groups of individuals, such as modern English, tend to have simpler inflectional morphology[2] than those spoken by smaller groups. Furthermore, Cysouw (2009) conducted a study looking at how the values of a language's features have low frequencies in the entire database; what he termed 'rarity'. In doing so, he found that, out of the fifteen languages in the WALS with the highest rarity, twelve had fewer than 6,000 speakers. However, instead of attributing this to population size, he argued in favour of geographical locations and genealogical inheritance as being the cause of this finding.

---

[1] The study of the way in which many languages have more than one way of saying the same thing.

[2] This refers to the processes that distinguish the forms of words in certain grammatical categories; such as affixation and suffixation.

More specific to this thesis, a significant amount of data has been gained through the use of computational simulations. However, these simulations have yielded mixed results on the relationship between populations and language change, with the results being heavily dependent upon the kind of social interaction used within the model (Wichmann & Holman 2009).

Nettle (1999*a*, 1999*b*, 1999*c*) was the first to raise questions concerning, and use computational simulations to explore the nature of population sizes and linguistic change. In doing so, he presented a model that demonstrated that higher rates of linguistic change occurred in smaller populations. He further supported these results using indirect empirical evidence, which suggested that languages with fewer speakers experience higher rates of linguistic change compared to those with many speakers. He demonstrated this by comparing the way in which languages in the Americas are divided into many unrelated families, all of which are characterised by having relatively few speakers. While languages in Eurasia and Africa are divided into a few, large, families and tend to have many speakers.

Initially, Nettle (1999*a*) argued that the cause of this stemmed from the way in which linguistic change occurs too rapidly in languages with a small number of speakers for them to grow very large before evidence of their relationships to other languages are lost. Whereas larger languages change slow enough for their relationships to be documented.

However, Nettle (1999*c*) presents a model that attributes the higher level of linguistic diversity in the Americas, not to population size, but to a rapid growth in diversity followed by a slower rate of decay; a model that allows higher diversity in geographical areas that were settled relatively recently than those settled in prehistory. Nettle's (1999*a*) logic is that it is easier to spread an innovation throughout a tribe of 500 people than it is through one of 5 million. This notion is explored through a simulation consisting of two linguistic features, which demonstrate that the rate at which the majority of the population switches between one of these two choices decreases to a small, but nonzero, limit when the population is increased from 120 to 500.

Wichmann et al. (2008, 2009) built upon the work of Nettle (1999*a*; 1999*b*; 1999*c*) by using a scale-free network, whereby the social influence of a particular individual increases during network growth in proportion with the amount of influence that this individual already exerts. Social distance in this model refers to the distances among individuals in the network.

In a variation of the model whereby a speaker randomly adopts variations from the entire population, the size of the population that has been given the linguistic variation indirectly impacts the probability that such a variation will diffuse further within the network. This version was tested with a number of different parameter settings, and it was found that with small diffusion rates, there was no impact on the rate of linguistic change from population sizes, but there was for larger diffusion rates.

In Wichmann et al.'s (2008) model, it was found that population size had no systematic effect on diffusibility, and that the degree to which languages undergo contact-induced change is likely to be dependent upon the particular history of interactions between the speakers. Although, their model did demonstrate that larger populations lead to slower rates of linguistic change. The authors found that, an increased population size will yield a lower rate of linguistic change in a situation where individuals adopt linguistic features from others anywhere in the speech community, where certain individuals are more connected than others, and where diffusion of innovations are high. Given these results, the authors go on to argue that languages such as English and Mandarin will change at a slower rate than languages spoken by populations in relative isolation to one another; as the situation might have been for some traditional societies.

In addition, Wichmann & Holman (2009) used a newer form of empirical test to that used by Nettle (1999*a*; 1999*b*; 1999*c*) to demonstrate that, during the breakup of dialects into separate languages, there is a tendency for a somewhat faster rate of change in the dialect of the majority. The tests here mainly show negligible effects of population. The exception being an apparently faster rate of change in the larger of two closely related variants. Wichmann & Holman (2009) argue that a possible reason for this could be the influence on emerging standard (or cross-regional) variants from speakers who shift from different dialects to the standard. The results of the work by Wichmann & Holman (2009) appear to demonstrate that the sizes of speaker populations do not in, and of, themselves determine rates of language change.

In sum, comparison of these empirical findings, accompanied with the results presented thus far in this thesis, suggest that the most plausible model for language change is one in which changes propagate on the level of individual interactions; in a network where individuals have different degrees of connectivity.

Furthermore, although the research output to date could be seen as a 'mixed bag' of data, there are some common themes. First, rates of language change tend to be lower in larger linguistic populations. Second, population size is not the sole factor that determines rates of linguistic change. Moreover, how integrated into their social group a language user is, or in our modelling terminology, how many linguistic tutors a language learner has, tends to be more important. Third, certain linguistic individuals being socially connected to more individuals than others also appear to play a crucial role in the rate of linguistic change.

These notions will be explored in this chapter through a model that incorporates the language learning mechanisms featured in the models presented in chapters 4 and 5; the behaviour of which have already been well documented in populations of static size. More specifically, this model will first explore the roles played by differing numbers of linguistic tutors and bottleneck sizes, before varying the amount of influence upon language learners. Finally, it will explore how the model behaves with larger population sizes that contract and expand in varying ways.

Although empirical studies using real-life participants have failed to provide convincing evidence that variability in speaker input influences language learning and evolution (Atkinson et al. 2015), the data presented in the aforementioned works, combined with the nature of E/I models and the results presented in chapter 4 and chapter 5 arguably make this a worthwhile endeavour.

6.3 The E/I Baseline Model

As mentioned above, the model presented here draws upon that presented by Kirby (2002*a*,*b*), in so much as it uses the ILM framework, a meaning-signal space that takes the form of predicate logic, and having both a production and learning phase. Although this last similarity is a common feature of all E/I models. Indeed, the algorithms used in these phases are quite different from those seen in Kirby's (2002*b*, 2002*a*) models, not least because they incorporate the principles of reinforcement learning.

As was the case with the previous two models, all the agents in every generation of this model are born as blank agents. They have no innate linguistic abilities and have to learn their language, in its entirety, in the manner described below.

At the start of a simulation run, an initial mature population, M, and an initial immature population, I, are created. Each mature agent is then assigned P_{MM} other randomly selected mature agents to play a language game with, and immature agents are each assigned P_{MI} randomly selected mature agents to act as their linguistic tutors. Due to mature tutors being selected at random, it is often the case that certain mature agents will train more immature agents than others, and will therefore exert more influence on the shape of the emerging language.

Each mature agent then plays a number of bouts, B, of the language game. Here, the number of bouts that each mature agent plays as the speaker with each of its assigned partners is the result of rounding up B/P_{MM}. After all the mature agents have played B bouts of the language game as the speaker, the mature population is removed and the immature population are promoted to mature agent status. Then, at the start of the next generation, a new immature population is created; and so on, until the maximum number of generations is reached.

Meanings in this model consist of one of five possible verbs and two of five possible names. When a meaning is created, these are chosen at random and used to construct the meaning. Meanings therefore take the form of `likes(Heather, John)`, for example.

In the discussion that follows, a distinction will be drawn between structure (or syntactic) rules and semantic rules[3], which will be created by the agents during their induction process. Semantic rules are denoted with an `A/` and consist of a meaning part and a signal part. Using the above meaning as an example, the semantic rule for 'Heather' would take the form of `A/ Heather, 'zkh'`; where `'zkh'` is a randomly generated string that is constructed by randomly selecting between one and three letters from the alphabet. In comparison, structure rules are denoted by an `S/` and dictate the way in which signal bits are ordered. These take the form of `S/ p(x, y), '/x/'('/p/', '/y/')`, for example.

The agents begin life producing complete `S/` rules, by generating a signal part for each meaning part; such as `S/ detests(John, Heather), 'jfi'('dol', 'ewp')`, for example. Then, throughout the course of its lifetime, the agent learns the given grammatical rules of its group. This allows it to produce specific signal parts for each of the different meaning parts, and order the whole signal in an manner that has, like the signal parts themselves, emerged through the agent's interactions to become a conventionalised system of usage.

As an example, imagine that we are looking at the grammar of a hypothetical agent after some language learning has taken place, this agent may have the following grammatical rules that pertain to the above meaning:

```
S/ p(x, y), '/x/'('/p/', '/y/')
A/ John, 'utr'
A/ Heather, 'zkh'
S/ p(x, y), '/p/'('/y/', '/x/')
A/ likes, 'dew'
S/ hates(x, y), 'tjh'('/y/', '/x/')
S/ hates(x, y), '/x/'('hfd', '/y/')
S/ admires(Pete, Gavin), 'ryt'('wde', 'zou')
.
.
.
.
A/ John, 'iws'
```

This would result in the agent being able to produce one of the following possible meaning-signal pairs when presented with the meaning `likes(Heather, John)`:

```
S/ likes(Heather, John), 'zkh'('dew', 'utr')
S/ likes(Heather, John), 'zkh'('dew', 'iws')
```

---

[3] Similar distinctions are seen in the models by Kirby (2002*a*,*b*). Although the manner in which they are created and used by the agents vary greatly.

```
S/ likes(Heather, John), ‘dew’(‘utr’, ‘zkh’)
S/ likes(Heather, John), ‘dew’(‘iws’, ‘zkh’)
```

The agents develop these grammars by playing a language game that involves producing and inducing meaning-signal pairs. This language game has a number of steps to it. First, a meaning is randomly generated using the possible meaning parts. The mature agent currently under consideration takes the role of the speaker and one of its randomly selected partners takes the role of the hearer. Both produce a signal for the meaning.

There are a number of steps involved in producing a signal. First, if the agent's grammar contains no entries, a signal part between one and three characters in length are created for each of the meaning bits using randomly chosen letters from the alphabet, these are then collated to create the signal for the meaning. Otherwise, the agent cycles through each of the rules within its grammar, and if a rule begins with `S/` (a structure rule), then the agent compares each semantic part of this structure rule to the corresponding semantic part of the meaning. If the two match, the rule is given a point. If they do not match, and the semantic part of the grammar rule is a variable, the agent will cycle through each of its grammatical rules and find a rule that is prefixed with `A/` (a semantic rule). If the semantic part of this rule matches the meaning part that is currently under consideration, then the rule is added to a temporary urn; an urn here is used in the same sense as in the reinforcement-learning model from the previous chapter. A semantic rule is then randomly selected from this temporary urn, and its semantic part temporarily replaces the variable in the main `S/` rule that is currently under consideration, and the signal part of this semantic rule is placed in the part of the `S/` rule's signal part that corresponds to the variable. A point is awarded. The highest scoring structure rule is then used as the agent's meaning-signal pairing.

Once both speaker and hearer have produced a meaning-signal pair, one of two things happens. If the two pairs match, the bout is considered to have been successful. In which case, a copy of each of the rules used by the speaker in constructing its meaning-signal pairing, both structure and semantic in nature, are added to the speaker and the hearer's grammar, as well as to the grammar of any immature agents that are assigned to the speaker or hearer; all of which then also apply the induction algorithm to the meaning-signal pair produced by the speaker. If, however, the hearer's pair does not match the speaker's pair, then only the hearer, applies the induction algorithm to the meaning-signal pair produced by the speaker.

The motivation behind this design choice is context. Up until now, the models presented here have involved one agent producing a signal for a meaning and the other agent trying to infer said meaning. However, in real-life, so very much of the meaning conveyed in conversation is drawn from context, with speakers using the specific words they use so as to ensure that the listener will understand. Indeed, both parties producing signals for the same meaning is one of the ways in which infants learn languages. We also see this behaviour when two individuals, without any common language attempt to communicate with one another; a phenomenon that this model will be used to explore in the next chapter. Ergo, both agents producing a signal for a given meaning and then comparing them can be seen as a short hand for providing agents with context. In reality, this is likely to make little difference when compared to the procedures of the models presented in the two preceding chapters, since all three mechanisms essentially involve the hearer knowing which meaning is trying to be conveyed by the speaker. However, adopting this specific procedure here enables this model to be used in subsequent chapters to explore the nature of contact language emergence.

The induction algorithm also has a number of steps to it. First, a copy of the meaning-signal pair produced by the speaker is created. Then, each signal part in this copy is compared to all of the semantic rules in the learning agent's grammar. If the signal part in the semantic rule that is currently under consideration matches the signal part of the speaker's produced signal that is currently being considered, and if the semantic part of this semantic rule matches the corresponding semantic part in the speaker's meaning-signal pair, then the signal part and its corresponding semantic part in the copy of the speaker's meaning-signal part is replaced by a variable and an extra copy of this semantic rule is added to the learning agent's grammar. Secondly, for any signal and semantic parts that do not match in this manner, a semantic part and signal part are randomly selected from the copy of the speaker's pairing. Both of which are not variables, and are used to construct a new semantic rule. This is then added to the learning agent's grammar. The semantic and signal bits used to construct this new semantic rule are then replaced in the copy of the speaker's meaning-signal pair by variables. Finally, a copy of the speaker's meaning-signal pair, which by now will consist solely of variables, is added to the learning agent's grammar.

6.4 Baseline Model Results

Following on from the findings of the previous two models, the analysis of this baseline model predominantly focuses on the impact of population size and different numbers, and types, of linguistic partners, as well as the size of the bottleneck (number of bouts) upon the amount of communicative success achieved. There will also be an analysis of contracting and expanding population sizes and the impact that this has upon language learning.

6.4.1 Baseline Social Interactions

It was shown in the previous models that the amount of diversity within an immature's linguistic training data played a crucial role in how quick, if at all, an expressive and stable linguistic system emerged within a population of language users. As such, the first step in the exploration of this baseline model's behaviour was to look at how different values of P_{MM} affected the amount of communicative success.

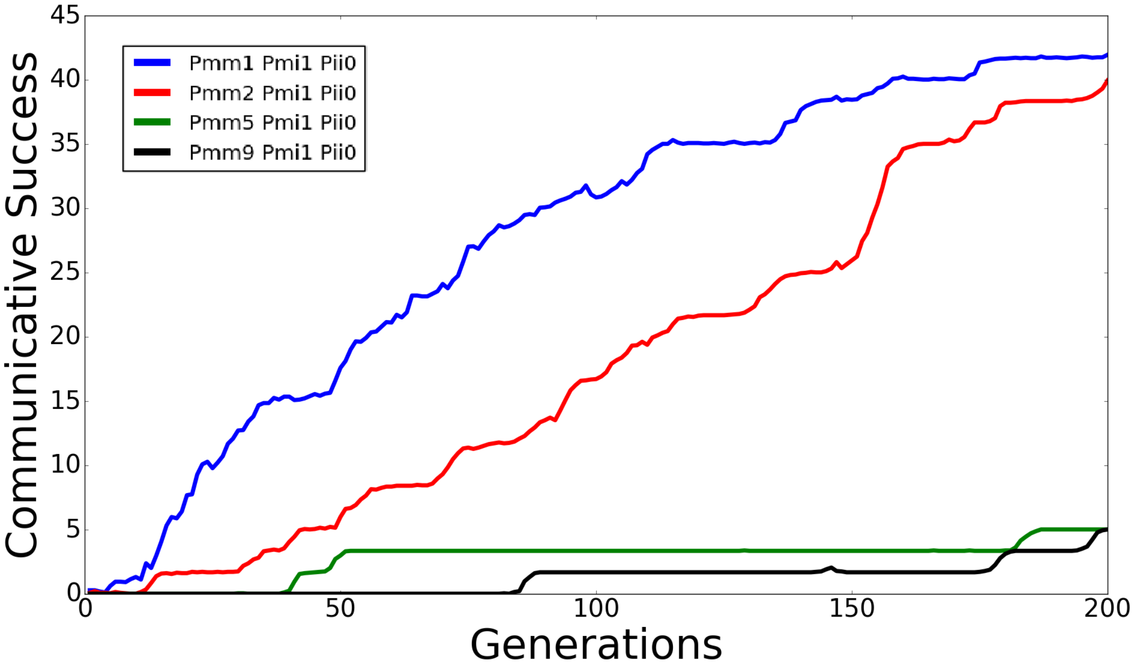


FIGURE 6.1: Graph depicting the amount of communicative success achieved with different values of P_{MM}, with P_{MI}=1, M and I=10, and B=50. Averaged over 30 runs. P_{MM}=1, P_{MI}=1, P_{II}=0 (blue line):$\mu = 27.76$, $\sigma = 1.47$. P_{MM}=2, P_{MI}=1, P_{II}=0 (red line):$\mu = 17.82$, $\sigma = 1.40$. P_{MM}=5, P_{MI}=1, P_{II}=0 (green line):$\mu = 2.73$, $\sigma = 0.79$. P_{MM}=9, P_{MI}=1, P_{II}=0 (black line):$\mu = 1.17$, $\sigma = 0.52$.

It can be seen from figure 6.1 that higher values of P_{MM} are actually detrimental to the population being able to achieve a high level of communicative success when P_{MI}=1. In comparison, figure 6.2 shows how higher values of P_{MI} have a significantly positive impact upon the communicative success of the population.

These results make logical sense. In figure 6.1, the higher values of P_{MM} and low level of P_{MI} result in a situation whereby mature agents communicating with one another leads to more signal diversity within the population, but the immature agents are only learning a very limited sub-sample of the linguistic system. In ILM terms, the bottleneck is too small to enable a stable communicative system to emerge. In contrast, the exact opposite is true in figure 6.2, where there is less signal diversity due to mature agents only communicating with one other mature agent, and immature agents experiencing a much wider spectrum of potential signals for each specific meaning. Thus improving the chances of these immature agents having more successful communicative episodes once they become mature. This is further supported by figure 6.3.

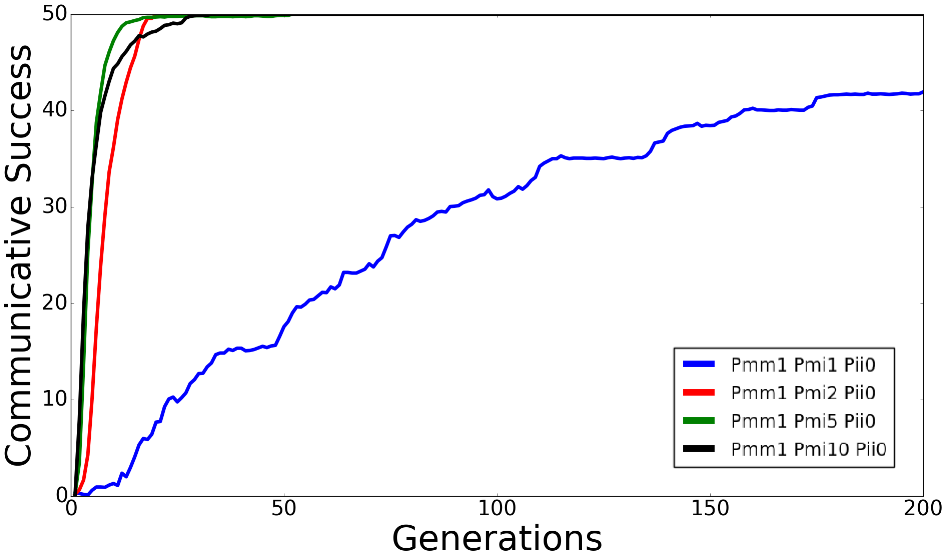


FIGURE 6.2: Graph depicting the amount of communicative success achieved with different values of P_{MI}, with P_{MM}=1, M and I=10, and B=50. Averaged over 30 runs. P_{MM}=1, P_{MI}=1, P_{II}=0 (blue line):$\mu = 27.76$, $\sigma = 1.47$. P_{MM}=1, P_{MI}=2, P_{II}=0 (red line):$\mu = 48.07$, $\sigma = 0.26$. P_{MM}=1, P_{MI}=5, P_{II}=0 (green line):$\mu = 48.89$, $\sigma = 0.11$. P_{MM}=1, P_{MI}=10, P_{II}=0 (black line):$\mu = 48.77$, $\sigma = 0.11$.

By looking at figure 6.4, it can be seen how the language during a successful run of the simulation changes over generational time. During early generations, an agent has numerous meaning-signal pairings for both syntactic (`S/`) and semantic (`A/`) rules. The number of different rules for each meaning-signal mapping then decreases until the population has a linguistic system that consists of a single meaning-signal mapping for each of the potential meaning parts and a single syntactic rule; a linguistic system that achieves 100% success rate. It should be noted here that no meanings or signals are being explicitly removed from the simulation by the researcher. Moreover, as the simulation progresses over generations, the agents give rise to a language that has fewer and fewer signals associated with any specific meaning. This continues until there is a situation akin to what is seen in generation 38 of figure 6.4, where the agents only have one choice of signal for each specific meaning. This is the result of the agents giving rise to a convention of usage through the course of their interactions.

The next step in the exploration of this model's behaviour was to investigate as to what kind of impact allowing the immature agents to communicate with one another would have upon the communicative success of the population. Figure 6.5 shows how, although the population does continue to develop the linguistic system in the same manner as when P_{II}=0, allowing the immature agents to learn from one another does impact the amount of generational time it takes for a fully stable language to arise.

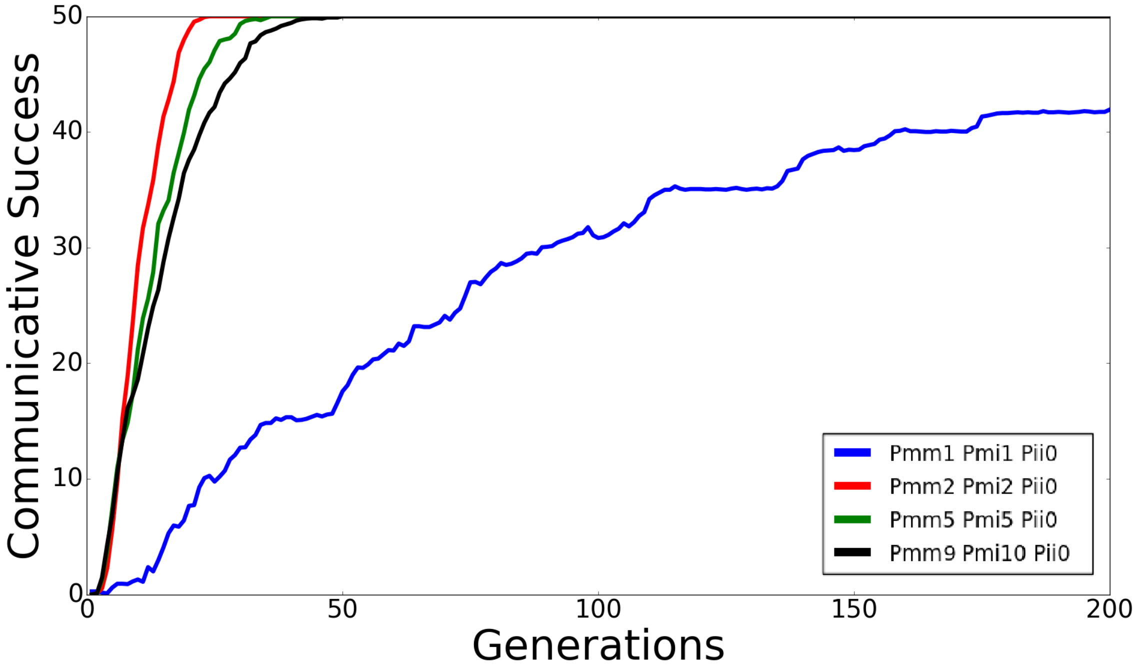


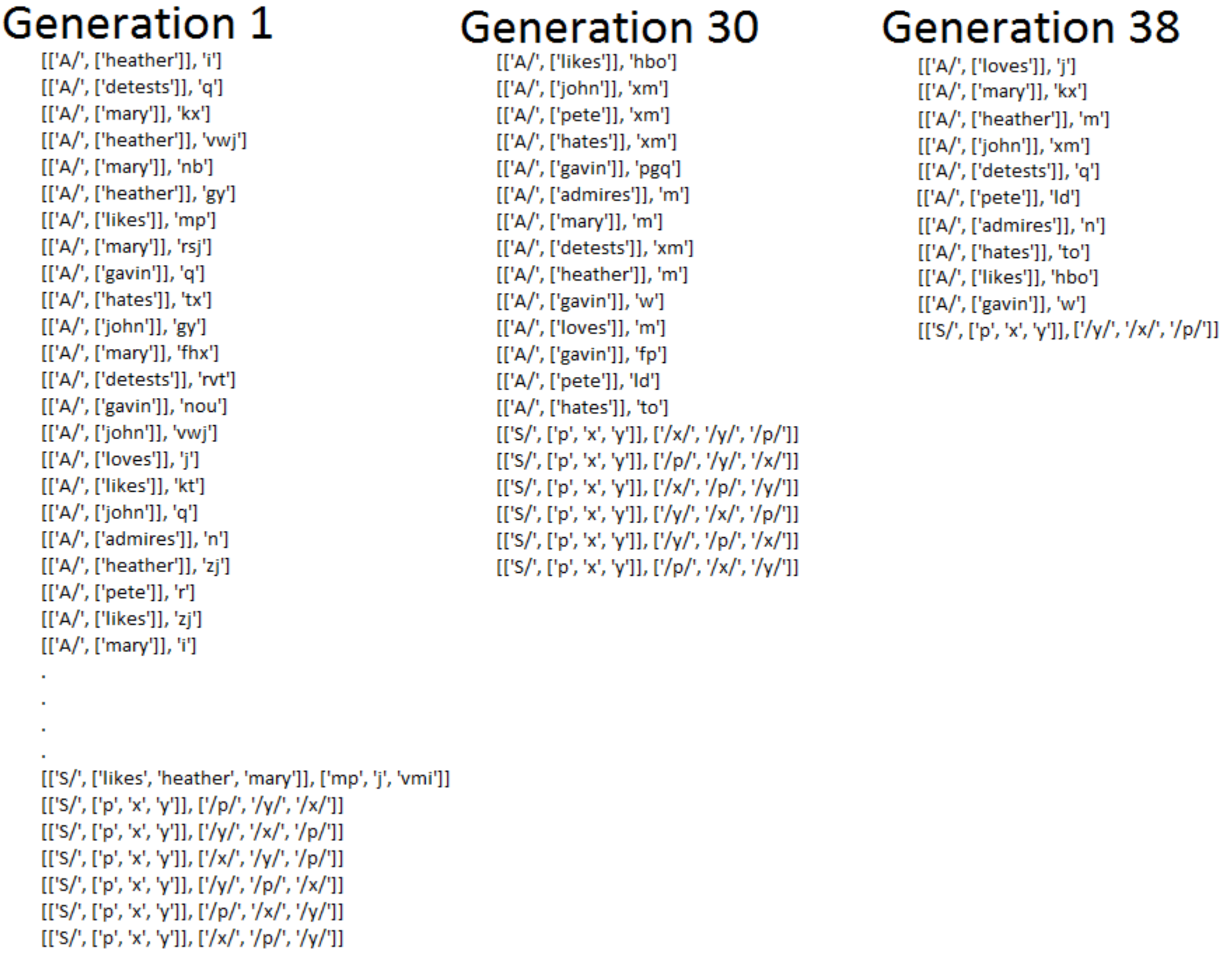
FIGURE 6.3: Graph depicting the amount of communicative success achieved with different values of P_{MM} and P_{MI}, with M and I=10, and B=50. Averaged over 30 runs. P_{MM}=1, P_{MI}=1, P_{II}=0 (blue line):$\mu = 27.76$, $\sigma = 1.47$. P_{MM}=2, P_{MI}=2, P_{II}=0 (red line):$\mu = 47.57$, $\sigma = 0.32$. P_{MM}=5, P_{MI}=5, P_{II}=0 (green line):$\mu = 46.95$, $\sigma = 0.40$. P_{MM}=9, P_{MI}=10, P_{II}=0 (black line):$\mu = 46.46$, $\sigma = 0.45$.

6.4.2 Baseline Bottleneck Exploration

In exploring the bottleneck, the two extremes in terms partner combinations (P_{MM} and P_{MI}= 1 and P_{MM} and P_{MI}= 10), were taken and B was increased. Figure 6.6 demonstrates how with higher P_{MM} and P_{MI}, there is no significant change in behaviour; with the success rate increasing less smoothly. However, in the P_{MM} and P_{MI}=1 case, there is a significant impact upon communicative success. This is understandable, given that a larger number of bouts, which essentially acts as a bottleneck, offers the agents a larger sub-sample from which to learn.

6.4.3 Larger Population Sizes

As was seen above, many researchers have used a number of methods in order to explore the nature of the relationship between population size and linguistic change. This body of work appears to indicate that it is not the size of the population *per se* that has influence on linguistic change. Moreover, given how different social network structures yield different results, it is arguable that language change is more heavily influenced by the connectivity of a linguistic social group. An argument that is supported by the data



Generation 1

```
[['A/', ['heather']], 'i']
[['A/', ['detests']], 'q']
[['A/', ['mary']], 'kx']
[['A/', ['heather']], 'vwj']
[['A/', ['mary']], 'nb']
[['A/', ['heather']], 'gy']
[['A/', ['likes']], 'mp']
[['A/', ['mary']], 'rsj']
[['A/', ['gavin']], 'q']
[['A/', ['hates']], 'tx']
[['A/', ['john']], 'gy']
[['A/', ['mary']], 'fhx']
[['A/', ['detests']], 'rvt']
[['A/', ['gavin']], 'nou']
[['A/', ['john']], 'vwj']
[['A/', ['loves']], 'j']
[['A/', ['likes']], 'kt']
[['A/', ['john']], 'q']
[['A/', ['admires']], 'n']
[['A/', ['heather']], 'zj']
[['A/', ['pete']], 'r']
[['A/', ['likes']], 'zj']
[['A/', ['mary']], 'i']
.
.
.
.
[['S/', ['likes', 'heather', 'mary']], ['mp', 'j', 'vmi']]
[['S/', ['p', 'x', 'y']], ['/p/', '/y/', '/x/']]
[['S/', ['p', 'x', 'y']], ['/y/', '/x/', '/p/']]
[['S/', ['p', 'x', 'y']], ['/x/', '/y/', '/p/']]
[['S/', ['p', 'x', 'y']], ['/y/', '/p/', '/x/']]
[['S/', ['p', 'x', 'y']], ['/p/', '/x/', '/y/']]
[['S/', ['p', 'x', 'y']], ['/x/', '/p/', '/y/']]
```

Generation 30

```
[['A/', ['likes']], 'hbo']
[['A/', ['john']], 'xm']
[['A/', ['pete']], 'xm']
[['A/', ['hates']], 'xm']
[['A/', ['gavin']], 'pgq']
[['A/', ['admires']], 'm']
[['A/', ['mary']], 'm']
[['A/', ['detests']], 'xm']
[['A/', ['heather']], 'm']
[['A/', ['gavin']], 'w']
[['A/', ['loves']], 'm']
[['A/', ['gavin']], 'fp']
[['A/', ['pete']], 'ld']
[['A/', ['hates']], 'to']
[['S/', ['p', 'x', 'y']], ['/x/', '/y/', '/p/']]
[['S/', ['p', 'x', 'y']], ['/p/', '/y/', '/x/']]
[['S/', ['p', 'x', 'y']], ['/x/', '/p/', '/y/']]
[['S/', ['p', 'x', 'y']], ['/y/', '/x/', '/p/']]
[['S/', ['p', 'x', 'y']], ['/y/', '/p/', '/x/']]
[['S/', ['p', 'x', 'y']], ['/p/', '/x/', '/y/']]
```

Generation 38

```
[['A/', ['loves']], 'j']
[['A/', ['mary']], 'kx']
[['A/', ['heather']], 'm']
[['A/', ['john']], 'xm']
[['A/', ['detests']], 'q']
[['A/', ['pete']], 'ld']
[['A/', ['admires']], 'n']
[['A/', ['hates']], 'to']
[['A/', ['likes']], 'hbo']
[['A/', ['gavin']], 'w']
[['S/', ['p', 'x', 'y']], ['/y/', '/x/', '/p/']]
```

FIGURE 6.4: Example of how a language evolves over generational time during a successful run, taken from a random mature agent in each generation.

thus far presented in this thesis. Thus, it was deemed worth while to explore whether or not increasing M and I from 10 to 30 had any significant impact upon language emergence within the current model.

Figure 6.7 demonstrates how, with this larger population size of 30 mature and 30 immature agents, it is much harder for the model to give rise to a stable communicative system. Indeed, the P_{MM}=1, P_{MI}=1, P_{II}=0 and P_{MM}=1, P_{MI}=1, P_{II}=1 cases fail completely. However, the P_{MM}=9, P_{MI}=10 P_{II}=9 does achieve some level of stability, and the P_{MM}=9, P_{MI}=10, and P_{II}=0 case could potentially reach complete stability after a prolonged period of time. Despite this, it is clear that, in this model, larger populations struggle to give rise to an expressive and stable linguistic system.

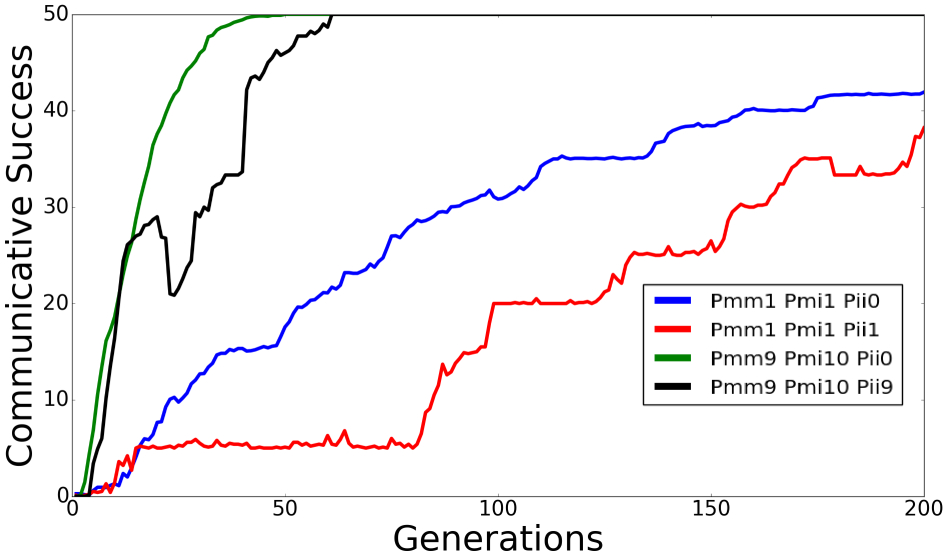


FIGURE 6.5: Comparison of the amount of communicative success achieved when immature agents are permitted to play the language game with one another, for P_{MM} and P_{MI}=1 and P_{II}=0 (blue line) and P_{II}=1 (red line), likewise for P_{MM}=9 and P_{MI}=10 (green and black lines, respectively). Averaged over 30 runs. P_{MM}=1, P_{MI}=1, P_{II}=0 (blue line):$\mu = 27.76$, $\sigma = 1.47$. P_{MM}=1, P_{MI}=1, P_{II}=1 (red line):$\mu = 18.28$, $\sigma = 1.48$. P_{MM}=9, P_{MI}=10, P_{II}=0 (green line):$\mu = 46.46$, $\sigma = 0.45$. P_{MM}=9, P_{MI}=10, P_{II}=9 (black line):$\mu = 32.43$, $\sigma = 1.13$

6.4.4 Even Amounts Of Influence

Given the above discussion, and the way in which the rate of linguistic change and communicative success is tied to the amount of diversity featured in an immature language learner's input training data, it is worthwhile exploring what the impact would be if all of the mature language users trained the same number of immature agents. In other words, if all of the mature agents exerted equal influence on the next generation of language users.had the same influence.

The reader may remember that, for whatever value P_{MI} is, each immature agent selects that number of mature agents at random to act as a linguistic tutor; while not selecting the same mature agent more than once. This results in a situation where some mature agents have more linguistic influence than others; a factor that the above literature indicates could play a significant role in rates of linguistic change. Thus, the behaviour of the model was tested in situations whereby the P_{MI} assigning procedure cycled through each of the mature agents in turn and assigned P_{MI} immature agents, selected at random. Again, ensuring not to select the same immature agent more than once[4].

---

[4] This method of assigning language tutors is only applied to data presented in this sub-section. The results that are featured in subsequent sections use the original assigning method.

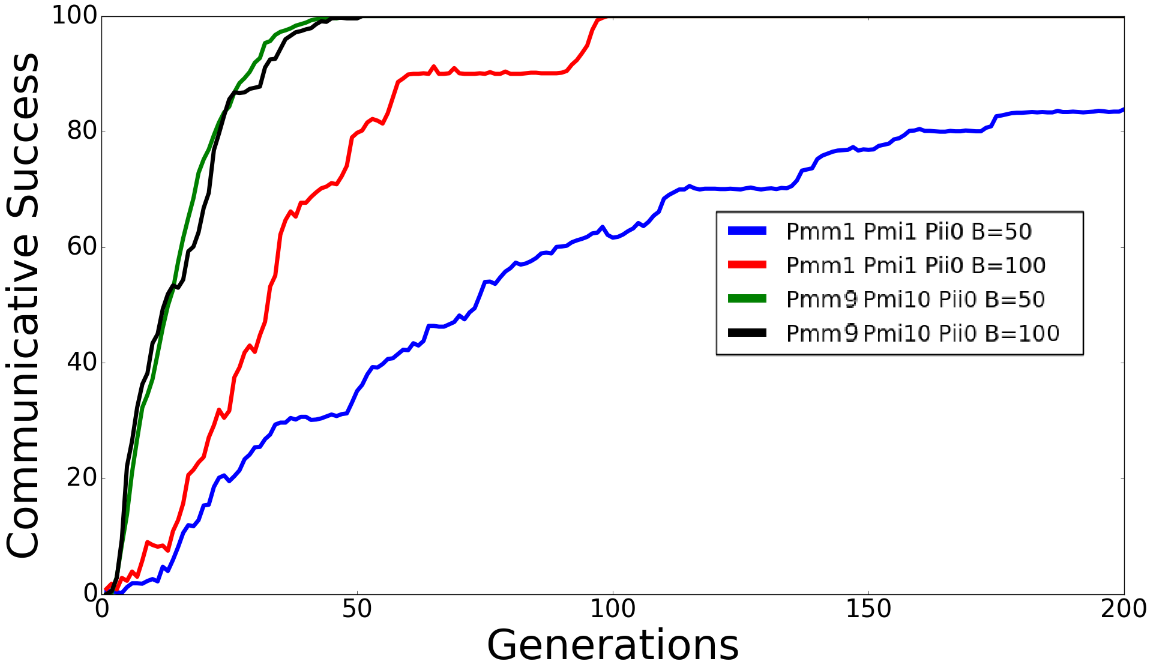


FIGURE 6.6: Graph depicting the amount of communicative success achieved with P_{MM} and P_{MI}=1 with B=50 (blue line: μ = 27.76, σ = 1.47) and B=100 (red line: μ = 38.82, σ = 1.79), and P_{MM}=9 and P_{MI}=10 with B=50 (green line: μ = 46.46, σ = 0.45) and B=100 (black line: μ = 43.52, σ = 0.54). With M and I=10. Averaged over 30 runs.

Figure 6.8 depicts the amount of communicative success achieved for two different combinations of linguistic tutors when they influence the same number of immature language learners, compared to that achieved when the linguistic tutors are assigned randomly; as seen in figure 6.5. Here, it can be seen that the success rate for P_{MM}=1, P_{MI}=1, and P_{II}=1 is severely lowered when all mature agents have the same amount of influence. This is in line with arguments from the above literature, which suggests that a number of individuals being more connected within the social network than others, results in them having more influence; and that this is a key factor in the rate of linguistic change. The P_{MM}=9, P_{MI}=10, and P_{II}=9 remains relatively unchanged; likely due to the limited population and meaning space sizes of the model.

6.4.5 Population Contraction And Expansion

All three of the models presented thus far have had one feature in common. Although they all had population turnover, the size of the population was always constant, with there being no fluctuation in size. This is clearly not the case in the real world, where social groups go through population growth and decline; a phenomenon that has an impact upon the demography of the linguistic community. Thus, two extensions were

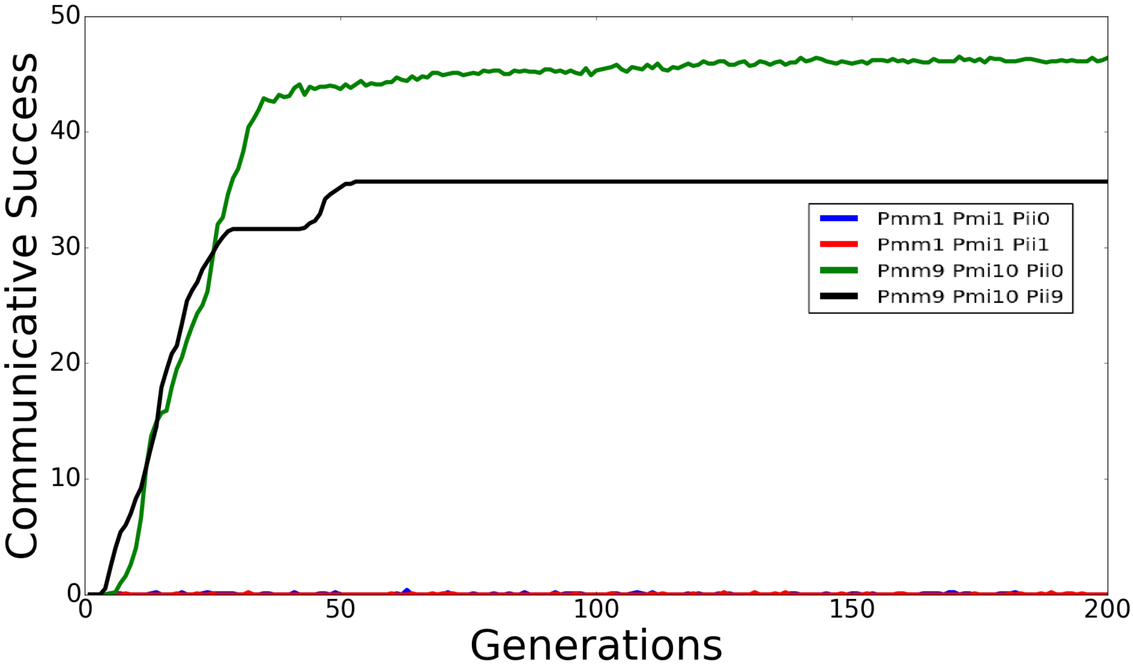


FIGURE 6.7: Graph depicting the amount of communicative success achieved with P_{MM}=1, P_{MI}=1, and P_{II}=0 (blue line:μ = 0.03, σ = 0.01), P_{MM}=1, P_{MI}=1, and P_{II}=1 (red line:μ = 0.01, σ = 0.01), P_{MM}=9, P_{MI}=10, and P_{II}=0 (green line:μ = 40.84, σ = 0.45), and P_{MM}=9, P_{MI}=10, and P_{II}=9 (black line:μ = 25.48, σ = 0.35), with M and I=30 in all cases. Averaged over 30 runs.

implemented with this baseline model. The first aimed to simulate the impact of population contraction, and the second explored the impact of population expansion. Both of these extensions involved the model going through exactly the same steps as detailed above, except for the differences discussed below.

After the first generation of the contraction model, if the length of the total population (M + I) is equal to what the total size of the population was set to during the first generation (i.e. $M + I$ in generation one), then the contraction procedure begins before any other actions are taken in this generation. Otherwise, a new batch of agents are created as before.

The contraction procedure removes a proportion, $0 \leq C < 1$, of the population, but does not reduce the population size below 4, since at least two agents of each type are required for the process of iterated learning to proceed in this model. The proportion of agents to be removed, at random, from the total population (M + I) is set at the start of the simulation. If R is the number of agents to be removed, then:

$$R = \min(M + I - 4, C(M + I))$$

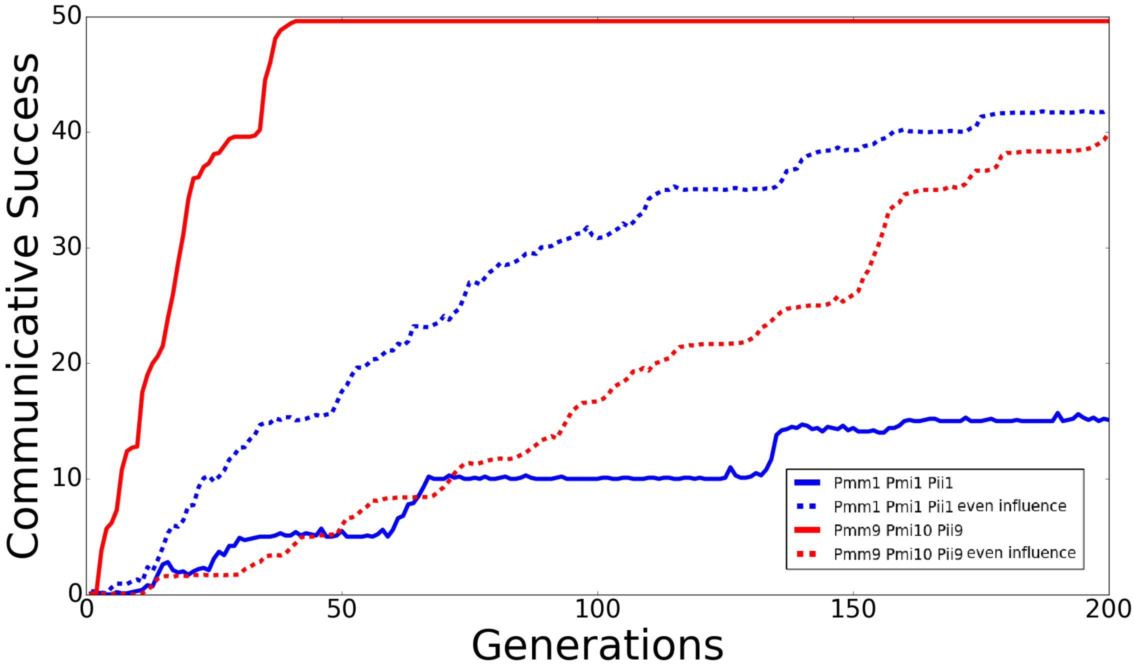


FIGURE 6.8: Graph depicting the amount of communicative success achieved with P_{MM}=1, P_{MI}=1, and P_{II}=1 with even influence (blue solid line:μ = 9.57, σ = 1.32) and normal influence (blue dashed line:μ = 18.28, σ = 1.48), compared to P_{MM}=9, P_{MI}=10, and P_{II}=9 with even influence (red solid line: μ = 31.62, σ = 1.12) and normal influence (red dashed line: μ = 32.43, σ = 1.13). When M and I=10 and B=50. Averaged over 30 runs.

The expansion version of the model works in a similar way. In that after the first generation, the size of the total population (i.e. the current values of M and I) are taken. A parameter set at the start of the simulation, E, is then taken and converted into a proportion the current total population size, and this number of new agents are created and added to I. More mathematically, where A is the number of new agents to add to I:

$$0 \leq E$$
$$A = E(M + I)$$

In both versions of the model, the values of P_{MM}, P_{MI}, and P_{II} that were set at the start of the simulation run are then taken and converted into a percentage of their originals values for the original half of the population that they applied to. This percentage of partners is then assigned in accordance with the size of the respective half of the population. If, for example, P_{MM}=5 and M and I=10 during the first generation, then the contraction procedure would result in M and I=5, P_{MM} would be adjusted to equal 3; due to the value of P_{MM} being rounded up. In mathematical terms, taking P_{MM} as an example and where M' is the new value of M:

$$P'_{MM} = (P_{MM}/M) \times M'$$

Due to how the population size fluctuates during a run of either of these model versions, the rest of the graphs presented in this chapter, unless stated otherwise, are the result of a single typical run of the model. This enables an analysis of the way in which population size and interconnectivity can impact the emergence of a conventionalised language.

The 'partner percentage', PP, measure used below is the result of taking the total number of trainers that each immature agent has (P_{MI} + P_{II}) for the current generation, and calculating this as a percentage of the total population size for this generation (M + I). More concisely, where M' and I' are the new values of M and I, respectively:

$$P = (P_{MI} + P_{II})/(M + I)$$
$$PP = P \times 100$$

In addition, the communicative success and partner percentage shown in the graphs below are normalised, so as to allow comparisons to be made between different population sizes. To do this, data was collected for four aspects of the models; P_{MM}=1, P_{MI}=1, with P_{II}=1 and P_{MM}=9, P_{MI}=10, with P_{II}=9. The amount of communicative success achieved and partner percentage for each generation of these four runs were then taken and normalised between the lowest and the highest population percentage achieved during every generation of all four runs of the models; likewise for the communicative success.

Figure 6.9 shows how contracting the population of the model is far more successful in achieving a stable linguistic system with P_{MM}, P_{MI}, and P_{II}=1 than in the standard baseline version of the model (figure 6.5). This increased success rate is due to the way in which the overall population size tends to be much smaller each generation than in the baseline model, while the number of training partners each immature agent has remains the same; one mature and one fellow immature trainer. In other words, the model performs better under these parameters due to the way in which the population has a higher percentage of overall interconnectivity than in figure 6.5.

This argument is supported by figure 6.10, which shows the same two parameters being tested on the expansion version of the model. Here, it can be seen that when P_{MM}, P_{MI}, with P_{II}=1 the model achieves no level of communicative success; a performance that is worse than in the baseline model (figure 6.5). As in the contracting model, this difference in performance is due to the way in which the total number of different communicative partners from which each immature agent learns is significantly lower in this much larger population size; which averages 38 throughout the course of the simulation run. The spike in the success rate for the P_{MM}=9, P_{MI}=10, with P_{II}=9 case shows how, when agents achieve successful communication 50% of the time, the agents can add a certain degree of noise to the linguistic system. This is due to the way in which, at this point, there are a number of conventions in place that are being used regularly by all agents, but there are also some agent-specific linguistic conventions in the system. The latter of which can still spread throughout the system due to the way in which the hearer will add an idiosyncratic convention used by a sinlge speaker to its own grammar; even if the hearer's best guess as to the meaning that the speaker was referring to is incorrect. However, the more widely agreed upon conventions will continue to be used more regularly than these more idiosyncratic linguistic conventions; which is why the system ultimately achieves 100% communicative success.

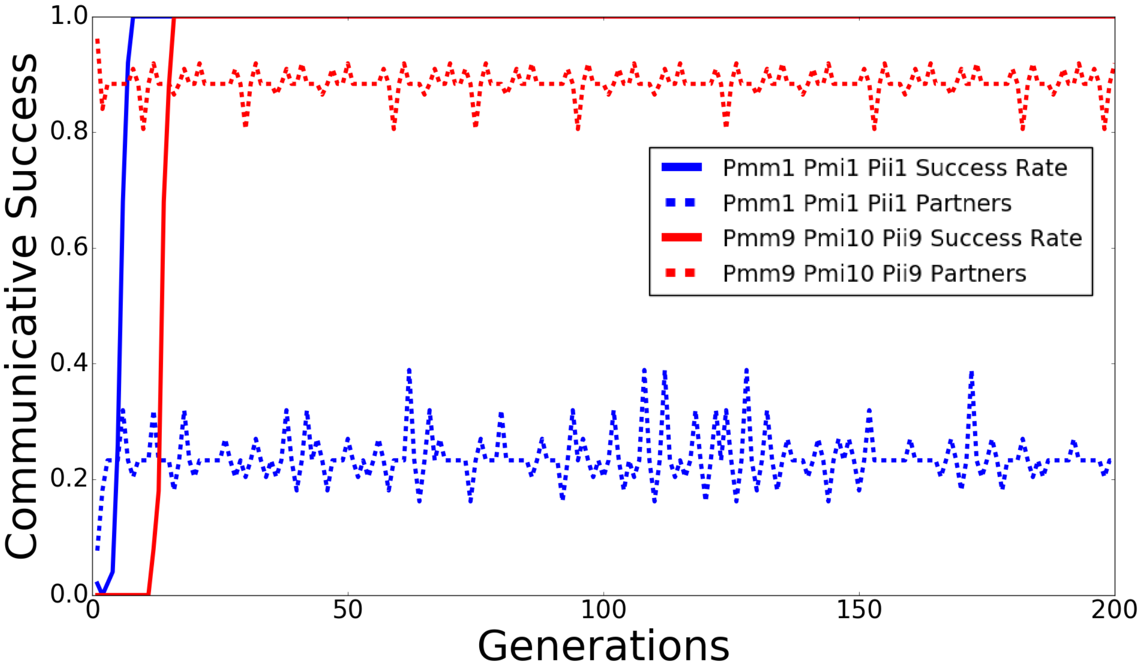


FIGURE 6.9: Graph depicting communicative success plotted against partner percentage, with all scores normalised. The blue lines representing the case whereby P_{MM}=1, P_{MI}=1, and P_{II}=1 for the first generation, with success represented by the solid lines ($\mu = 0.97$, $\sigma = 0.14$) and partner percentage represented by dashed lines ($\mu = 0.23$, $\sigma = 0.04$). Likewise, red lines represent the case whereby P_{MM}=9, P_{MI}=10, and P_{II}=9 for the first generation, with success rate represented by the solid lines ($\mu = 0.93$, $\sigma = 0.24$) and partner percentage represented by dashed lines ($\mu = 0.88$, $\sigma = 0.02$). Here, C=2, and both M and I are initially set to 10, with the contracting procedure occuring every generation after the first.

The difference in the amount of communicative success achieved by the contracting and expanding models can be seen more clearly seen in figure 6.11, which demonstrates how a higher overall percentage of population connectivity results in a completely stable linguistic system emerging within a couple of generations.

However, populations do not always have sudden shifts in size like this; sometimes it is a gradual process. Therefore, similar tests were conducted on versions of the model whereby the total population started low and slowly increased before slowly decreasing again (figure 6.12), and where the total population size started high and slowly decreased before increasing again (figure 6.13). Combined, figures 6.12 and 6.13 demonstrate how,

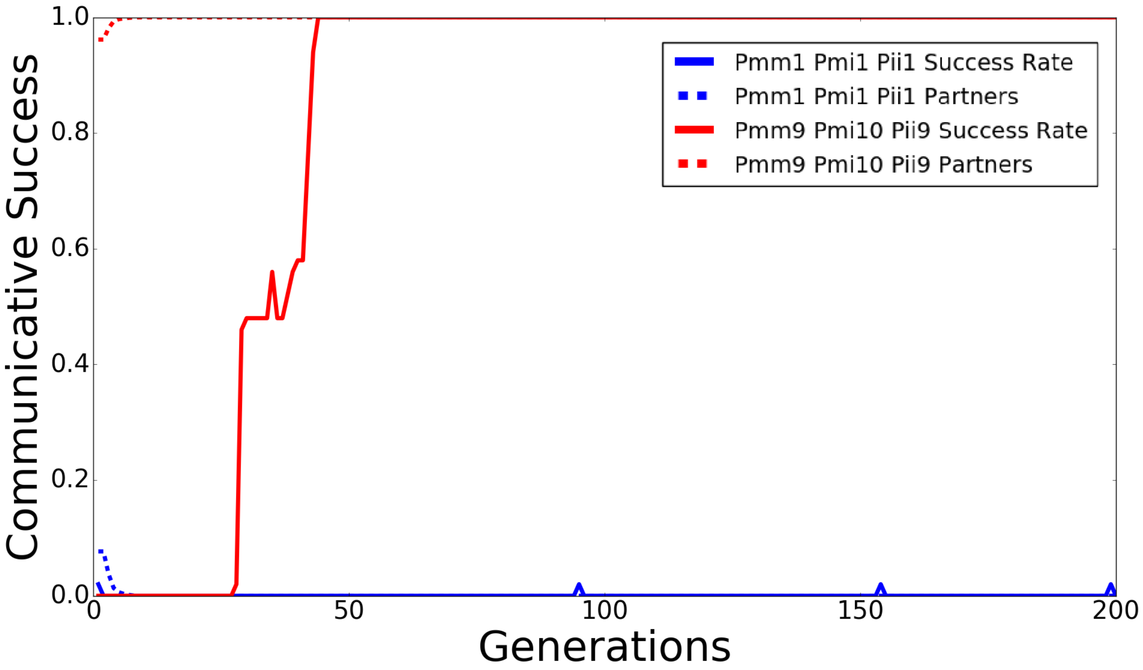


FIGURE 6.10: Graph depicting communicative success plotted against partner percentage seen in a single run of the expansion model, with all scores normalised. The blue lines representing the case whereby P_{MM}=1, P_{MI}=1, and P_{II}=1 for the first generation, with success represented by the solid lines (μ = 0.00, σ = 0.00) and partner percentage represented by dashed lines (μ = 0.00, σ = 0.00). Likewise, red lines represent the case whereby P_{MM}=9, P_{MI}=10, and P_{II}=9 for the first generation, with success rate represented by the solid lines (μ = 0.82, σ = 0.35) and partner percentage represented by dashed lines (μ = 0.99, σ = 0.00). Here, E=30, and both M and I initially set to 10, with the expansion procedure occuring every generation after the first.

regardless of the number of partners agents are assigned, a successful linguistic systems fails to evolve in any system whereby the population size changes gradually.

6.5 Conclusions

In line with the results presented elsewhere in this thesis, the model presented in this chapter has demonstrated that a key parameter in the emergence of a stable linguistic system within a population of language users is the number of trainers from which immature language users learn their linguistic system.

Specifically, it has been shown that the agents in this simulation are far more successful in establishing a conventionalised language when mature agents only interact with a small number of other mature agents, while immature agents get their linguistic training data from a large number of different mature agents, due to this improving their chances of experiencing successful communicative episodes upon their promotion to mature agent

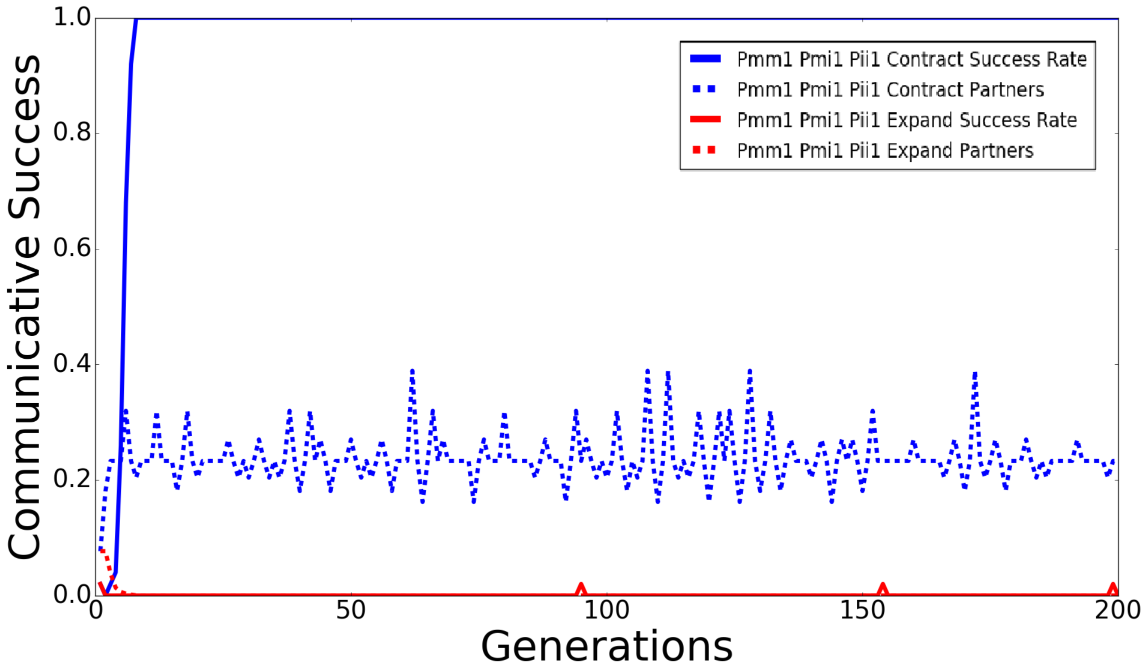


FIGURE 6.11: Graph depicting communicative success plotted against partner percentagel, with all scores normalised. The blue lines showing the communicative success (solid line: $\mu = 0.97$, $\sigma = 0.14$) and partner percentage ($\mu = 0.23$, $\sigma = 0.04$) for a run of the contraction version of the model when P_{MM}, P_{MI}, and P_{II}=1, with C=2 in the first generation. Likewise, the red lines showing the communicative success (solid line: $\mu = 0.00$, $\sigma = 0.00$) and partner percentage ($\mu = 0.00$, $\sigma = 0.00$) for a run of the expansion version of the model when P_{MM}, P_{MI}, and P_{II}=1, with E=30 in the first generation. Here, both M and I initially set to 10.

status. In addition, the agents in this model struggled to achieve communicative success with greater population sizes.

Taken together, these results lend support to the notion that population size, in and of itself, does not have a direct impact upon linguistic evolution and change. Moreover, it is the make-up and connectivity of said population that is the determining factor, with individuals displaying less variability in their linguistic conventions the more integrated they are in the social group (Labov 2001, de Bot & Stoessel 2002, Ke et al. 2008). This argument is further supported by the results yielded by the contraction and expansion models, which demonstrated how smaller population sizes with a higher percentage of connectivity perform much better when establishing communicative norms than larger populations with proportionally smaller rates of connectivity.

A major criticism that could be levelled against the model presented in this chapter is that, when an agent has no grammar whatsoever during the first generation of the simulation run, the production algorithm constructs a randomly generated signal part for each of the meaning parts presented in the meaning; and these early signal parts are then organised in the same manner as the meaning. As a result, the agents should inadvertently be more inclined to a specific syntactic rule; namely $S/p(x, y), `/p/'(`/x/', `/y/')$.

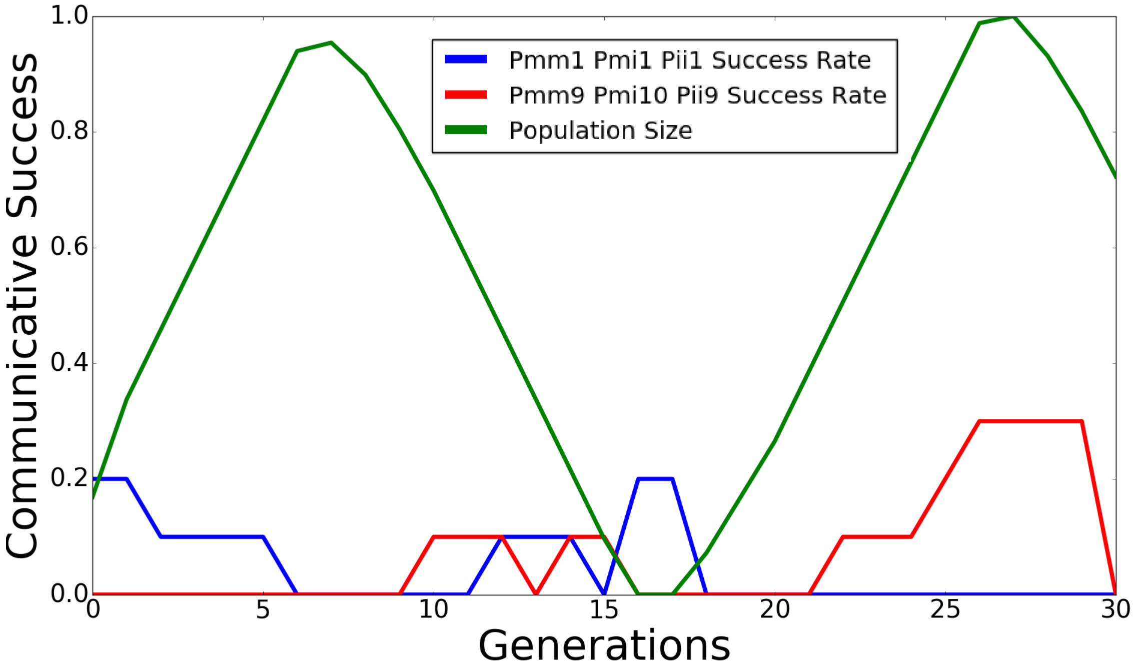


FIGURE 6.12: Graph depicting the amount of communicative success achieved for P_{MM}=1, P_{MI}=1, and P_{II}=1 (blue line: $\mu = 0.04$, $\sigma = 0.07$) and P_{MM}=9, P_{MI}=10, and P_{II}=9 (red line: $\mu = 0.07$, $\sigma = 0.10$) and how the total population size $(M + I)$ changes over generational time (green line: $\mu = 0.55$, $\sigma = 0.31$), with both M and I initially set to 4 and slowly increasing first. Averaged over 30 runs and with both population size and communicative success normalised between 0 and 1.

However, the results presented above show that the structure rule that ends up being dominant, and therefore dictating the syntactic tendencies of the population, is often very different to this. Thus, it is arguable that the syntactic structure of the linguistic systems in this model are the result of social evolution, which in turn stems from individual learner biases.

It should also be noted that a number of the graphs depicting system behaviour within this chapter demonstrate step-like changes in communicative success. After close examination of the data, it became apparent that this is caused by severe fluctuation in the number of generational time-steps it takes for runs included in the average performance calculation to reach various levels of communicative success. Thus, while the graphs in question accurately depict the average success of the linguistic systems within these various runs, future work should include re-running these experiments and averaging over more than 30 runs. This would eliminate this step-like behaviour.

In sum, the model and results presented here support previous work in demonstrating that the most successful model for language emergence and linguistic change is one that sees linguistic variation occurring and propagating at the local level of individual interactions; where individuals are well integrated into the social network (Ke et al. 2008, Wichmann & Holman 2009).

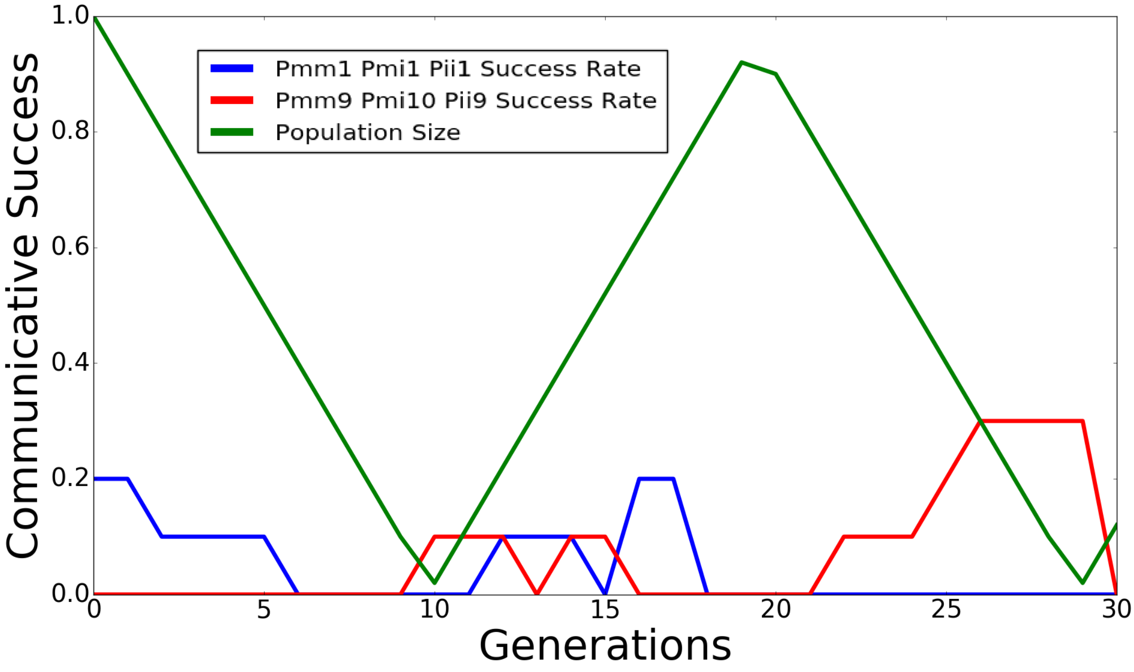


FIGURE 6.13: Graph depicting the amount of communicative success achieved for P_{MM}=1, P_{MI}=1, and P_{II}=1 (blue line: μ = 0.04, σ = 0.07) and P_{MM}=9, P_{MI}=10, and P_{II}=9 (red line: $\mu = 0.07$, $\sigma = 0.10$) and how the total population size ($M + I$) changes over generational time (green line: μ = 0.46, σ = 0.29), with both M and I initially set to 25 and slowly decreasing first. Averaged over 30 runs and with both population size and communicative success normalised between 0 and 1.

Chapter 7

Linguistic Populations And Contact Languages

7.1 Introduction

A contact language is a linguistic system that tends to take its vocabulary from one language and its grammar from another. Indeed, it is regularly the case that the great majority of a contact language's lexicon is drawn from just one language; with a much smaller, roughly 20%, contribution from other languages (Sebba 1997). In such a case, the language that contributes the majority of the lexicon is referred to as the *lexifer* or *superstrate*, while the *substrate* refers to the language from which the smaller proportion of grammar is drawn. The dominant theory here is that the users of the substrate language normally possess less social power and are therefore likely to be more accommodating and use words from the superstrate language due to its user having more social power. Although, the meaning, form, and use of these words may be influenced by the substrate language (Holm 2000).

In the case of Fanagalo, for example, grammatical aspects resemble those seen in English much more than those seen in Zulu, while the bulk of the vocabulary stems from Zulu. Thus, Zulu is regarded as the lexifer and English and/or Afrikaans the substrate (Mesthrie 1989). As with so many other cases, it is difficult to state here as to what extent the similarities between the pidgin grammar and the substrate grammar derive from the substrate grammar itself, and to what extent it is present in the pidgin as a result of simplification of the lexifer in accordance with 'universal' principles of simplification that are applicable to any language (Sebba 1997). Indeed, how influential the substrate can be is an active research question in pidgin and creole research.

There is a body of work which claims that developing our understanding of the processes that underpin contact language emergence has the significant potential to improve our undestanding of linguistic change as a whole (Botha 2003, 2006, Roberge 2009, Tallerman & Gibson 2013*a*). Furthermore, as will be seen below, these contact languages go through a refinement process; and it stands to reason that this refinement could be dictated by learner bias. As such, iterated learning-based models are well suited to exploring the process of contact language emergence.

7.2 Pidgins

A pidgin is the linguistic result of two communities, without any prior common language, coming into contact with one another over a prolonged period of time, and having the need to communicate with one another for a specific purpose; such as trading.

They are restricted linguistic systems in that they have limited lexical range and structural means when compared to full languages (Holm 2000, Roberge 2013)[1]. In other words, the resultant make-shift language will reduce the number of words and drop unnecessary complications, such as inflections; i.e. 'two knives' becomes 'two knife'. However, these reductions will be compensated for by expanding the usage of the words that do feature in the pidgin (Holm 2000).

It should be noted that not all diminished languages qualify as a pidgin. For example, an Englishman using broken Spanish while on holiday in order to communicate with the locals does not constitute a pidgin; but is referred to as *jargon*. The difference here stems from the way in which a pidgin is a stabilised convention of usage.

7.2.1 Monogenesis And Relexification Orgins

The theory of a monogenetic origin of pidgins essentially argues that all modern-day pidgins and creoles are linked, either directly or indirectly, to a Mediterranean pidgin; most likely Savir (Sebba 1997). According to those who subscribe to this theory, the proto-pidgin in question mostly drew its lexicon from Portuguese, and was subsequently spread across the world as a result of European colonisation.

A major component of this theory is the concept of *relexification*. The idea here is that, although Portuguese was the original lexifer, as this proto-pidgin came to be used among different groups of English, French, Dutch, and Spanish colonisers, it was subjected to an almost complete substitution of vocabulary. The operative word here being 'almost', due to there being a few words of Portuguese origin that are incredibly wide spread, and continue to pervade within the world's languages to this day (Sebba 1997).

---

[1] Although, it has been argued that there is a tendency to overlook the way in which pidgins can develop a considerable degree of complexity (Botha 2009).

However, the notion of relexification can only be relevant to pidgin genesis when explored in relation to pidgins evolving in a particular socio-temporal context. Indeed, although relexification is crucial to the monogenesis theory, it does have a number of weaknesses. Including being unable to explain why the proto-pidgin had the features that it did, or take into account that relexification could occur locally under the right conditions (Muhlhausler 1986). Furthermore, despite the monogenesis argument accounting for the similarities seen in pidgins and creole grammars, it has been argued that these similarities could be the result of a common linguistic 'blueprint', or universal redundancy to simplify a language in a specific manner (Sebba 1997).

The concept of *nautical jargon* is not too dissimilar to that of monogenesis, in so much as it posits that pidgins and creoles around the world are related to either English or French seafarer's jargon. It has even been shown that seafaring English could account for much of the lexical similarity seen in pidgins and creoles of the Pacific region; some of which are shared as far away as the Caribbean and the Atlantic (Clark 1979, Sebba 1997). However, a key criticism of this notion is that there are many pidgins and creoles that share grammatical similarities in geographical regions outside of the Atlantic and Pacific. This latter point could, in turn, be seen to support the above 'blueprint' argument.

Despite these criticisms, such ideas do still offer a slight insight into the historical relationships between various contact languages. Indeed, there is no reason to doubt that nautical language played some role in the global emergence of pidgins and creoles, given that some of them were certainly spread by seafarers; even if they were not responsible for their ultimate origin (Hall 1966).

7.2.2 Polygenesis And Parallel Origins

The logical counter-argument to the idea of monogenesis is that of *polygenesis*. In the pidgin and creole literature, this is largely an umbrella term for a number of different theories. One such notion is that of the *universalist theories*, which draw upon the 'linguistic blueprint' argument in positing that humans possess universal strategies for simplifying languages, and that these are part of our innate language learning abilities. Hall (1966), for example, describes a process whereby the similarities between contact languages around the world can be attributed to similar processes acting upon the different languages. As Hinnenkamp (1984) states:

> *"languages that are structurally quite different may be simplified in very similar ways, yielding nearly identical surface structures [...] Every language has its way of arriving at this kind of surface structure."*
> *- Hinnenkamp (1984:160)*

Thus, it has been argued that there is the possibility of innate universal principles being at work in the pidginisation process. Sebba (1997) highlights how such universal principles might take different forms, which would affect the pidginisation process in various ways:

1. They could manifest themselves as restraints on what adult language users are capable of learning, or at least provide a scale of learnability whereby some types of structure or vocabulary would be too difficult to learn under certain circumstances. This would then account for the absence of complex sentences (i.e. relative clauses) or words of more than two syllables in rudimentary pidgins.

2. They could manifest themselves as constraints on permissible relationships between form and meaning.

3. They could take the form of strategies for simplifying language during interactions with other speakers, so as to achieve maximum communicative efficiency; resulting in foreigner talk on the one hand and simplified 'learners' interlanguage' on the other.

If one were to accept the notion of universals in human language and Chomsky's idea of a parameter setting model, then it is arguable that pidgins would choose the simplest, minimal, specification of each of these parameters. This would result in pidgins being the 'default case' of human grammar (Sebba 1997).

Another noteworthy polygenesis theory is the *common core* concept, which attempts to account for similarities in the combinations of grammars that are involved in the emergence of pidgins. The idea here is that the pidgin which emergences from the interactions between social group A and social group B will only have in its grammar the part of the grammar of the language of social group A that is also found in language of social group B. In other words, the overlapping parts of the two grammars.

A final, and popular, hypothesis that is of interest is referred to as *foreigner talk*. Inspired by the concept of *baby talk*, the logic here is that, when we interact with individuals who are non-natives, and who are attempting to communicate with us in our own language, we use a simplified version of our language in an attempt to reduce errors in communication. In other words, foreigner talk is a special communicative style by which native speakers communicate with foreigners, and not vice-versa; much like the locals interacting with the aforementioned Englishman on holiday in Spain.

Foreigner talk-based systems have a number of similarities to pidgins, such as being grammatically simplified, possessing a limited vocabulary, and only being employed in a limited range of interactions; mostly where communication is essential and no alternative method of communication is available (Sebba 1997).

This is interesting, given that contact situations between European colonisers and indigenous people must have been a rather common occurrences during the age of discovery.

In an effort to communicate with one another, it is entirely possible that both groups spoke their own version of foreigner talk to the non-native speakers. Thus, while each side would have walked away from the interaction with a practical understanding of important vocabulary, and perhaps a rough idea of word order, which would no doubt prove useful in future contact situations, they probably also had a false understanding of the actual nature of the language used by these 'others'.

In studying the role of English-based foreigner talk in the emergence of Tok Pisin, Muhlhausler (1986) concluded that foreigner talk played a very crucial role in the very early stages of the emergence of the pidgin, even if it was of little relevance in its later development. However, Hinnenkamp (1982, 1984) argued that his own work demonstrated that foreigner talk was too inconsistent and unpredictable to provide an adequate model for second language acquisition. Although, he did observe that, through an interactive process, native speakers in a Turkish town using a conventionalised foreigner talk (known as *Tarzanca* or 'Tarzan talk') and German tourists using broken Turkish were giving rise to a form of 'tourist pidgin'. It is arguable that a process of conventionalisation was crucial in this case. If there was a tradition in place of using foreigner talk to communicate on a regular basis with foreigners, as appears to be the case with *Tarzanca*, then it is logical that a convention of usage may emerge and be consistent enough to act as a language model for non-native speakers.

These polygenesis-based theories, and others like them, have gained significant support in recent years. This is largely due to the discovery of a number of African (i.e. *Lingala* of Zaire and *Sango* of the Central African Repbulic), Asian (*Nagamese* of India), and Oceanic (*Pidgin Fijian*) pidgins, which are completely void of any European connections, and yet, still exhibit basic similarities to their European-inspired counterparts. Prior to these discoveries, it was possible for monogenesis advocates to find an historical, European-based, link between all pidgins and creoles (Sebba 1997). However, it can now safely be argued that any theory of pidgin origins must be polygenetic in nature.

In summary, it is safe to argue that foreigner talk is likely to have played a crucial role in the emergence of pidgins across the world, at least to some degree. However, it is not sufficient, in and of, itself to enable a pidgin to develop (Hinnenkamp 1984), and there is currently not enough data to safely conclude that it is a necessary pre-condition either. Indeed, despite the amount of academic interest in the process of pidginisation, the way in which they develop remains uncertain for a simple reason; there are no first-hand observer accounts of the phenomenon from trained linguists (Sebba 1997).

7.3 Pidgin To Creole Transitions

Both Tok Pisin and Russenorsk originally emerged in the South Pacific and Artic waters, respectively, as forms of maritime jargon used by sailors and traders. The former evolved to become the national language of Papua New Guinea, while the latter died out after a few centuries without ever developing into anything more complex than maritime jargon. This indicates that pidgins go through a form of *developmental continuum* (Sebba 1997).

Muhlhausler (1986) presented a schematic representation of this pidgin-creole continuum (figure 7.1). The first stage, aptly named *the jargon stage*, is a pidgin in its most rudimentary form; with both limited structure and range of functions. It is therefore only used in the most restricted of circumstances. These jargons tend to be characterised by great instability, both in terms of vocabulary and grammar:

> *"[they are] unstable linguistically and socially. Moreover, they are not transmitted in any consistent way from generation to generation, but invented in an ad-hoc fashion".*
> *- Muhlhausler (1986:147)*

These jargons tend to be the result of contact between two groups, both of which possess their own language. To advance beyond this stage, it seems necessary for more groups and languages to become involved (Sebba 1997).

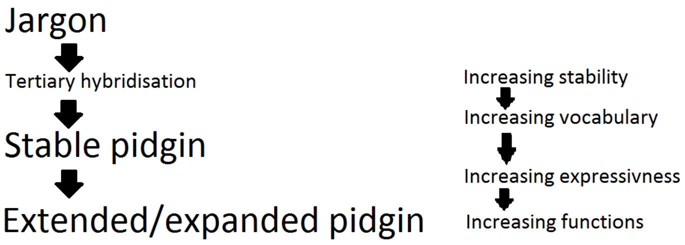


FIGURE 7.1: Steps involved in the transition from an informal jargon through to a stabilisaed pidgin. Replicated from Sebba (1997).

It has been argued that this leads to a process known as *tertiary hybridisation*, which draws influence from the realm of biology. In a biological context, *primary hybridisation* refers to the development of multiple species from a single ancestral species. This translates linguistically to the idea of a 'family tree model' of human languages. Similarly, the term *secondary hybridisation* refers to the interbreeding of two distinct species to form a new one.Whinnom (1968) argues that pidgins could be seen as the result of secondary hybridisation, in a linguistic sense, as they appear to be the result of 'interbreeding' between European and indigenous languages. However, Sebba (1997) observes how this idea may be too simple due to the way in which, while mixed communication systems do arise from contact between two languages, Whinnom's (1968) theory suggests that one side is invariably trying to learn the other's language. The resulting mixture of this language would lack stability because, while there is still a target language available as a model to learn from, the speakers of the other language will continue to learn it in its normal form. This would result in a continuum of learners' interlanguages, rather than a stable language with its own norms. Thus, we would have a range of different 'versions' of the target language, which are dependent upon the proficiency of the individual speakers.

Whinnom (1968) then goes on to argue that the emergence of a 'proper' pidgin results from *tertiary hybridisation*, whereby the pidgin comes to be used for communication between speakers who are not speakers of the original target language. Thus, there is now a language that is not native to any of its speakers being used as *lingua franca* by people who have no other language in common. By this stage, the pidgin would have to be stable, in that the norms for vocabulary and grammar would have to be fixed, due to the groups no longer having access to the target language.

The basic argument here is that, if speakers of a superstrate become the least important part of such a pidgin triangle, and if close contact is established and maintained between speakers of two different substrate languages over a prolonged period of time, then an *expanded pidgin* results (Holm 2000). This is achieved through a process whereby the original pidgin is continually developed in order to meet more demanding communicative needs (Muhlhausler 1986).

Holm (2000) observes two more important distinctions between pidgins and other forms of contact languages that should be noted when looking at this continuum. First, it is important that an amount of social distance is kept between speakers of the superstrate and speakers of the substrate languages, otherwise the latter would likely acquire enough information to learn the superstrate language in its entirety. Secondly, the two languages must not be closely related.

In additionally, it should be noted that contact languages can emerge between trading partners of equal social power. In such cases, vocabulary is drawn equally from both languages; such as *Russenorsk*.

In short, most of the contemporary arguments concerning pidgin formation view these languages as developing through a series of steps:

Stage 1: Two groups, A and B, communicate by means of one another's languages; A attempts to learn B's language and vice-versa. However, both A and B are simplifying their languages for the benefit of the other group. Thus, A is modelling a foreigner talk-based version of B's language, and vice-versa.

Stage 2: The norm for communication becomes a version of one of the languages, say language A. This is likely to be the language whose speakers dominate the interactions by virtue of greater social power, while group B continues to learn the languages of group A but does not have full access to it, and may only ever be exposed to the foreigner talk version. Thus, people in B never learn A's complete language, only a reduced version of it, which suffices for their interactions with A. This reduced version of language A becomes the norm for inter-group communication between A and B.

Stage 3: Group B begins to use their reduced version of language A to communicate with members of another group, C. B and C now communicate with one another by means of a pidgin form of language A. Members of group A are not involved and may no longer be present at all. So there will now be no model for the native speaker version of language A, and no pressure to learn it completely and correctly. This process would result in a stable version of pidgin A.

This idea that 'proper' pidgins will only develop with stable norms, grammar, pronunciation, and vocabulary once they are taken up by a third party appears to be largely supported by the fact that most of the successful pidgins, those with the longest histories and spoken most widely, in the world today are also *lingua francas* used by speakers of large numbers of mutually unintelligible languages.

The second stage sees an unstable jargon develop into a stable pidgin, with its own grammatical norms, and lexicon. Linguists are torn as to the process behind this transformation. Some argue that it may be through a process akin to tertiary hybridisation, while others believe that the emergence of established norms are what counts. Muhlhausler (1986), for example, argues that the following are the crucial features of the stabilisation stage:

1. The reduction of variability found in preceding jargon stages.

2. The establishment of relatively firm lexical and grammatical conventions.

3. The development of grammatical structures independent from possible source languages.

However, whatever mechanisms one subscribes to, the crucial point is that for a pidgin to stabilise, it must develop its own norms of grammar, lexicon, and pronunciation. At which point, the target language becomes the pidgin and not the lexifer language (Sebba 1997).

After that stage, we have the extension or expansion stage. Stable pidgins emerge during the course of a generation or so, but once stable norms have begun to emerge, the process of development does not necessarily stop there. What happens at this point largely depends upon the functions that the pidgin is used to perform and the areas of life in which it is used (Sebba 1997).

If it continues to only be used in a limited number of scenarios, such as trading, it is unlikely to develop due to it being unable to move forward unless its users extend its usage and functions by using it in new social domains. If this does not happen, it may remain stable for some time, but its usefulness may eventually start to decline and it will die out; as was the case with Russenorsk. Alternatively, the pidgin may become extended or expanded, whereby it is used in a multilingual environment. Due to its usage here, it is extended and used beyond the original limited function that caused it to come into being (Todd 1990). These fully extended pidgins may then begin to act like first languages, and may develop into a full creole depending upon historical circumstance (Muhlhausler 1986).

Finally, the pidgin becomes a creole when children are born into a pidgin-speaking social milieu and end up requiring the pidgin as their primary language. It appears as though this could happen at any stage in the pidgin developmental process (Sebba 1997). Here, as a pidgin develops, it may become the primary language of a social group, none of whom speak it natively. The pidgin becomes the only common medium of communication, and is used in everyday interactions in many social domains, even in interactions between members of the same house. Then, when children come along, the pidgin will be learned by them as their first language; these children are then first generation creole speakers.

7.4 Creoles

The term *creole* refers to a class of vernaculars that emerged during the 17th-19th centuries, under similar social conditions. Namely, colonies settled by Europeans, who typically spoke non-standard varieties of metropolitan languages, and who utilised non-indigenous labour (Mufwene 2008, Roberge 2013).

Despite the similarities among pidgins and creoles, there is one key difference. As we have seen, pidgins are acquired as second languages, while creoles become such when they are learned as first languages. In other words, creoles have a jargon or pidgin in their ancestry, and are spoken natively by an entire speech community whose ancestors are likely to have been geographically displaced.

It has been argued that, in order to be learnable, the structure of the language has to adhere to a strict set of constraints being placed upon the possible structures that could be formed with it. These act as a kind of linguistic 'blueprint', and a language that violates these constraints would not exist because it would be un-learnable (Sebba 1997). However, as is demonstrated by the ample body of ILM work, languages themselves can evolve in accordance with the selection pressures placed upon them by language user bias, so as become more learnable. Given the above discussion, there is a chance that this same process occurs during pidgin to creole transformation.

The idea that children on plantations quickly create creole languages, complete with all of the features of human language that pidgins lacked, is still accepted in some language research (Pinker 2003). However, more recent thinking has begun to distance itself from the argument that creole formation requires the special intervention of children and that creole grammar reflects some form of Universal Grammar (Roberge 2013).

The argument put forward by this more contemporary view is that structural expansion is not dependent upon some form of nativisation process, and that the emergence of a creole language is not necessarily abrupt. Moreover, "complex linguistic systems can arise out of long-term encounters between adults in contexts of increasing use of a restricted but developing [medium of inter-ethnic communication]" (Roberge, 2013:540); with child language learners levelling out variability and producing more regular grammars, while not actually being the innovators (Mufwene 2008, Roberge 2013). This process is remarkably similar to what we have seen in the process of iterated learning. Thus making it hard to ignore the possibility that such phenomena are the result of learner bias.

7.4.1 Theories Of Creole Origins

As with most areas of human language research, there is agreement that pidgins and creoles possess some properties that call for explanation, but there is disagreement upon the nature of such properties. Nevertheless, there are four properties of creole languages that are believed to play a role in the genesis of such a communicative system:

1. Creole languages are assumed to be more akin to other creole languages, in terms of their structural features, than to other 'normal' languages; even when compared to the lexifer and substrate languages that form the basis for a creole. It has been largely argued that this cannot be coincidental.

2. Creoles are often seen as being simpler in terms of morphology, phonology, and syntax when compared to full languages.

3. Creoles often exhibit more mixed grammars than full languages. This observation has led to numerous parallels being drawn between biological processes and creole genesis. Most notably, it has been argued that, just as individuals are the result of a mix of African, European, and Asian ancestry, so too are their languages likely to be a mix of European vocabulary, with some African and Asian syntax and semantics.

4. Creoles are often seen to exhibit more internal variability than full languages. Indeed, they are assumed to be highly dynamical linguistic systems, which often co-exist with their lexifer languages in the same speech community.

Calvin & Bickerton (2000) argue that the transition from pidgin to creole is akin to the transition from protolanguage to full language. In that creoles have not normally existed for more than a single generation as a secondary language before becoming the primary linguistic input data for children.

In the case of plantation labourers, the children could have acquired the language of their parents, but there was little incentive for them to do so in such a diverse social milieu. Instead, they formulated a language based upon the mixed and inconsistent linguistic input that they received (Roberge 2013). If one were to look at this phenomenon through an ILM lense, then it would make sense that immature language learners who receive inconsistent learning data would give rise to a more coherent linguistic system. It would also support the observation that creole grammars tend to have more in common with other creole languages than they do the grammars of their lexifers or substrates (Bickerton 1995, Sebba 1997).

7.5 Model Motivation

There are three main reasons as to why this thesis will explore the nature of contact languages by way of an iterated learning-based E/I model. First, there is a growing body of work which argues that researchers stand to gain valuable insights into the nature of linguistic evolution as a whole by studying the nature of contact language emergence and change (Botha 2003, 2006, Roberge 2009, Tallerman 2013).

Second, doing so would constitute a novel application of the iterated learning framework. One that could offer valuable insights into the nature of the self-organisational properties of contact languages by exploring whether or not learner biases play as much of a role there as they do in 'normal' linguistic evolution.

Third, as Holm (2000) observes, the validity of theories put forward to account for the genesis and development of pidgins and creoles crucially depend upon their ability to account for all of the various socio-linguistic situations within which they come into being and pervade. However, developing such a theory is notoriously difficult given that creole researchers have not yet been able to observe the process of abrupt creolisation first-hand, and in cases of gradual creolisation of expanded pidgins, competent adult non-native speakers always co-exist with the children growing up as native speakers. Thus making it incredibly difficult to observe which group is contributing what to the emerging structure.

Additionally, there always seems to be outliers to any theories or rules put forward. Tertiary hybridisation, for example, is often thought to be a crucial factor in a contact language's stabilisation, but the existence of Chinese pidgin English suggests that this may not be the case (Holm 2000). As another example, we have seen that social power is often cited as being a critical factor in determining which language becomes the superstrate and which one becomes the substrate. However, the existence of contact languages, such a Russenorsk, raises doubts about this claim.

Furthermore, as discussed above, there is a strong argument made in the literature in favour of a linguistic ‘blueprint’ which ensures that languages adhere to a strict set of principles, so as to make them learnable. However, the work presented in this thesis thus far has built upon a larger body of research that shows how learner bias acts as a selection pressure upon languages; resulting in linguistic structures themselves evolving so as to be easier to learn. Combined with the observations that creoles do not typically exist for more than a single generation as a secondary language before becoming the primary linguistic input for immature language users, and that creole grammars tend to have more features in common with one another than they do with their lexifers and substrates, there is a compelling case for the origin and evolution of pidgins and creoles to be investigated by way of the ILM framework.

As such, this chapter will use an iterated learning-influenced model in order to explore a number of the aforementioned subjects seen in the contact language literature. Specifically, varying the make up of linguistic tutors in order to explore what impact his has on the behaviours of the model described below, before exploring what impact one linguistic social group possessing more social power than another during contact situations has upon the emergence of contact languages, as well as the nature of tertiary hybridisation. Most significantly, however, this model will feature both a pidgin and creole phase, so as to emulate an aforementioned argument in the literature. Namely, that a successful pidgin remains as such until a time when it is used by immature language users as their primary language. At which point, the pidgin goes through a transformation into a linguistic system that is more akin to a full language. The idea here is that this transformation could be the result of mechanisms seen elsewhere in linguistic evolution. Specifically, the pidgins evolving in accordance with learner bias so as to be easier to learn.

7.6 The Pidgin And Creole Model

Here, the baseline model discussed in the previous chapter is extended to explore the nature of contact languages. This model has number of parameters that are similar to those used in the model from the previous chapter. In order to make this model amenable to study, it consists of three distinct phases (figure 7.2). The first phase proceeds as detailed in the last chapter, without any population contraction or expansion. There are, however, three distinct populations of agents; each evolving their own unique linguistic system in isolation from the other two.

That is to say that, at the start of the simulation, three populations are created; each consisting of M mature agents and I immature language learners. Each of the mature agents in the simulation are then assigned P_{MM} other mature agents, selected at random from the same population, to play the language game with. Each immature agent is then assigned P_{MI} mature agents, selected at random from the same population, to act as its linguistic tutors. As before, each mature agent then plays B number of bouts of the language game. As before, the number of bouts that each mature agent plays as the speaker with each of its assigned partners is the result of rounding up B/P_{MM}. After all the mature agents have played B bouts of the language game as the speaker, the mature population is removed and the immature population are promoted to mature agent status. Then, at the start of the next generation, a new immature population is created; and so on, until the maximum number of generations is reached[2].

Once all three isolated populations achieve a perfect score for their respective linguistic systems, we allow them to continue for a further five generations before beginning phase two, where we introduce mature agents from population one to mature agents from population two (figure 7.2). During this second phase, which can also be seen as the pidgin phase, each member of the mature half of population one is randomly assigned a number of mature agents from population two to act as linguistic partners, P_{CMM}, in the pidgin game. Likewise, mature agents in population two are randomly assigned mature agents from population one. Meanwhile, all three populations continue to develop their social group-specific linguistic systems each generation, as discussed in the previous chapter. Crucially, while populations one and two are interacting with one another, population three remains isolated from both groups. Keeping population three in isolation, and not allowing its agents to interact with agents from either of the other two populations is crucial to the future steps in the simulation as it allows the model to simulate the concept of tertiary hybridisation.

For each bout of the pidgin game, the number of which is set by the parameter B_p, we allow each distinct agent in population one and two, selected in a random order, to play the pidgin game as the speaker; with the hearer being randomly selected from the speaker's assigned pidgin partners. Each of the immature agents in both populations two and three are assigned P_{CMI} mature agents, selected at random from the same population as the immature agent in question, to act as their linguistic tutors. It is important to note here that a mature agent's pidgin partners are a different group from the partners that were drawn from its own population and assigned to the agent to act as linguistic partners for the baseline language game; which is still on-going during this phase of the simulation.

However, unlike the baseline game, not all speakers playing this pidgin game play the same number of bouts as the speaker. In order to simulate the proposal in the literature that one of the populations involved in pidgin formation has a higher level of social power than the other, a dominance parameter, D, was implemented. If D is set to 0.8 and B_p is set to 1000, for example, mature agents in population one will play 800 bouts as the

---

[2] In order to avoid repetition, the reader is advised to consult Chapter 6 for a detailed discussion as to the nature of the language game used during this first stage. The important point to remember is that this first stage essentially involves three of these models from Chapter 6 learning at the same time, and that there is no interaction between these three different populations.

speaker while those in population two will only play 200. More specifically, where B_{P1} is the number of bouts that each mature agent in population one plays as the speaker and B_{P2} is the same for population two:

$$B_{P1} = D \times B_P$$
$$B_{P2} = B_P - B_{P1}$$

The pidgin language game then proceeds by creating a randomly generated meaning, as in the baseline phase. The speaker then produces a signal for this meaning. During this phase of the simulation, each agent has two distinct language stores. The first is the linguistic system that it has learned/developed while interacting with other agents from its own population; henchforth referred to as L1. The second stores the linguistic system that it has developed while interacting with agents from the other population, this constitutes the pidgin language and will be referred to as the agent's L2 language.

As such, in order to produce a pidgin-based signal for this meaning, the speaker first cycles through each of the meaning parts in turn, and gathers together all of the semantic rules within its L2 that match the meaning part currently under consideration, and places them in a temporary 'urn'. As before, a random rule is them selected from this urn and is used to convey the meaning part currently under consideration. However, if there is not a semantic rule in the speaker's L2 that matches the current meaning part, then the speaker conducts the same semantic rule search procedure on its L1. If no appropriate rule can be found there, then a random signal bit is produced, as in the baseline phase. This speaker production algorithm was designed so as to simulate the logical process of "this individual is from a different social group with a different language in place, and we do not share a word for this meaning, so I will attempt to teach him the word that I know for it".

One of the major differences between the pidgin game and the baseline language game concerns how the hearer in the pidgin game has a different production algorithm to that of the speaker. In so much as it only conducts the search for signal parts for the various meaning parts on its L2, but does not check its L1. If a suitable semantic rule cannot be found in its L2, then no signal part is produced for the meaning part currently under consideration.

It is important to note that meaning-signal pairings in the pidgin game consist solely of semantic rules, there are no structure rules. This is done to simulate the concept of *foreigner talk*. It could be argued that this foreigner talk-like behaviour should not be explicitly encoded into the agent's behaviour. However, this behavioural characteristic is so prolific in human interaction that it is unlikely to be a result of the kind of social-cultural processes that are being studied here. Thus making this a justified design choice of the model.

If one of the semantic rules presented by the hearer is present within the collection of rules produced by the speaker, the interaction is deemed to have been a success. In which case, a copy of each of the rules used by the speaker is added to the L2 of the speaker, hearer, and any immature agents for whom the speaker or hearer are acting as linguistic tutors. If, however, the bout is unsuccessful, then a single rule used by the speaker is selected at random and added to the L2 of the hearer only.

This process of all three populations playing their own population-specific language games, and then populations one and two playing the pidgin game continues for a single generation. In the results presented below, B_P=1000; and under such a parameter setting, the agents achieve >90% success rate in the pidgin game by the end of this single generation.

The simulation then moves onto the third, creole, phase of the simulation. The intention behind having both a pidgin and creole phase to the simulation was to emulate the common agreement within the literature that a useful pidgin remains as such until a time when it is used by immature language users as their primary language; which is typically a single generation. At this point, the pidgin goes through a transformation into a linguistic system that is more akin to a full language.

In this third phase of the simulation, known as the creole phase, population one is removed from the simulation. Mature agents in populations two and three are each assigned P_{CMM} number of linguistic partners from the other population; in the same way as mature agents from populations one and two were during the pidgin phase. Then, for the first generation of the creole phase, the model proceeds exactly the same way as it did in the pidgin phase; with one notable exception. Namely, any immature agents assigned to the mature agents update their L1 with the successful meaning-signal pairs used by the speaker during the pidgin game instead of adding them to their L2.

After this first generation of the creole phase, the simulation procedure changes. From here on out, neither of the remaining populations play the baseline language game. In other words, they no longer develop their population-specific linguistic systems. Instead, each mature agent is assigned P_{CMM} linguistic partners selected at random from the mature half of the other population. The mature agents then play the creole game. During this game, as with the pidgin game, the agents are subjected to the D parameter whereby the mature agents of population two act as the speaker more than the mature agents from population three if $D \geq 50$.

The creole game proceeds in exactly the same fashion as the baseline language game, just with a signal production algorithm that is different in one respect from that used in the baseline language game. Namely, if there is no structure, $S/$, rule in the speaker's L1 at the start of the signal production procedure, then a semantic rule is randomly selected as described above for each part of the meaning, and this signal is added to the agent's grammar and used as the linguistic output.

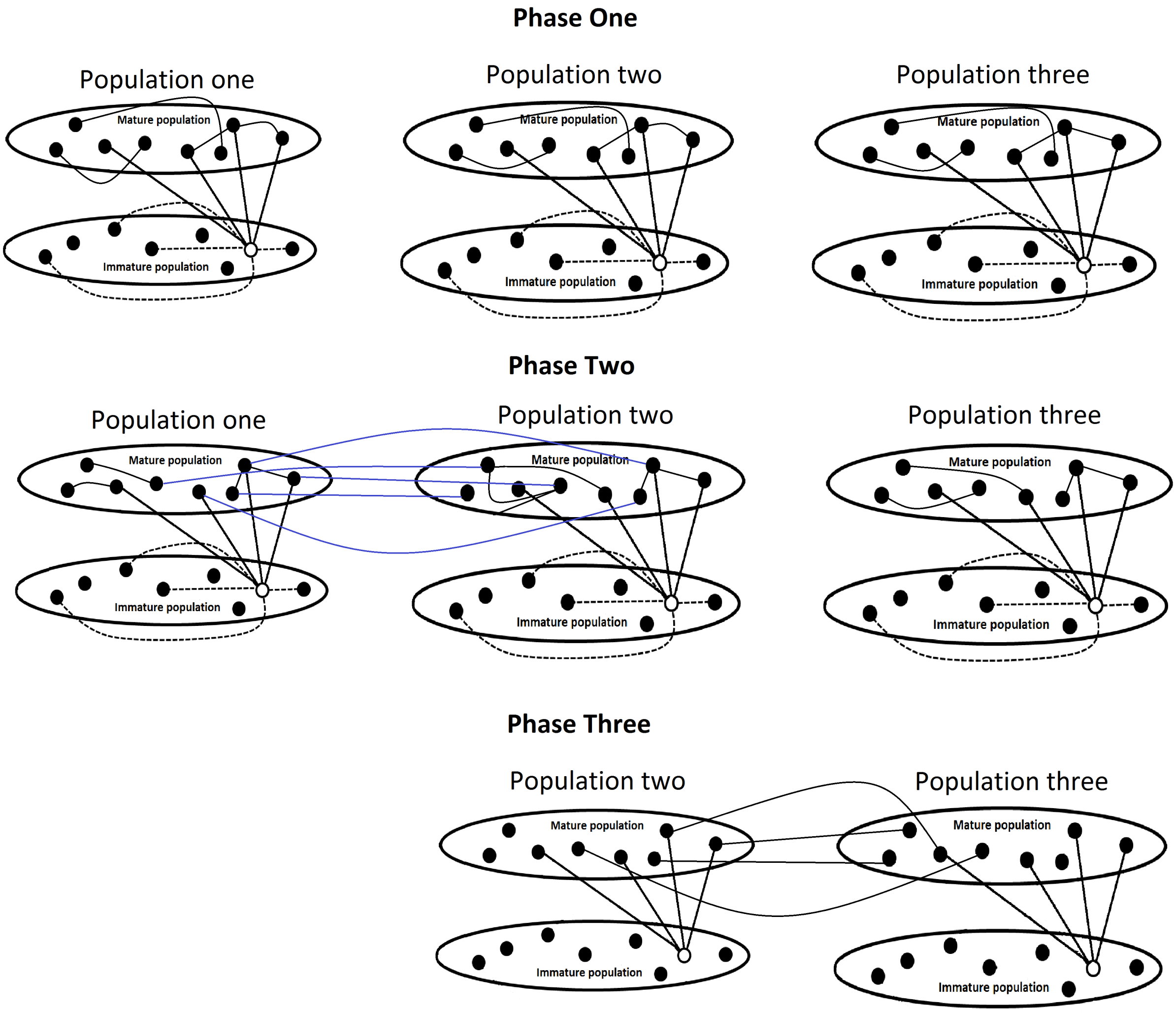


Figure 7.2: A diagrammatic representation of the various phases featured in the pidgin and creole model. The black solid lines representing interactions where mature agents update their L1 grammar store, the black dashed lines representing interactions where immature agents update their L1 grammar store, and the blue solid lines representing interactions where mature agents update their L2 grammar store.

7.7 Pidgin And Creole Model Results

A number of experiments were conducted with the pidgin and creole model, using various parameter settings, so as to explore the arguments in the above literature.

7.7.1 Impact Of Social Interactions

Following on from the analysis of the baseline model, figure 7.3 demonstrates that, throughout the course of an entire simulation run, from the baseline through to the creole phase, that higher values of P_{CMI} lead to a stable linguistic system arising. It can also be seen that higher values of P_{CMM} offer an additional advantage in the P_{CMM}=9 with P_{CMI}=1 case. Again, this is due to the immature agents receiving diminished training data from a linguistic system with a large amount of signal diversity. Thus leading to them being able to communicate successfully a larger percentage of the time.

Two other interesting behaviours can be seen in figure 7.3. First, the P_{CMM}=9, P_{CMI}=1, with P_{CII}=9 case plateaus at 20 bouts of success out of every 50. This is due to the fact that there are a large numbers of mature agents interacting with each other, enough for the immature agents to learn enough of the languages to be successful at communicating ∼50% of the time; with the average over 30 runs being a constant 20 across generations.

Second, in the P_{CMM}=1, P_{CMI}=1, with P_{CII}=0 case, the communicative success rate always drops to 0 once the simulation reaches the creole phase. However, the number of generations it takes for the simulation to get to the creole phase fluctuates between runs, meaning that different runs drop to 0 at different generational time steps. Thus, the slow, step-like, decline that that can be seen in the P_{CMM}=1, P_{CMI}=1, with P_{CII}=0 case is caused by some of the runs in the average score still having a communicative success of 100% where they have not yet entered the pidgin and creole phases of the simulation. In contrast, other runs are outputting a 0% communicative success rate during these same generational time steps due to them already being in these latter phases of the simulation, whereby they fail to estblaish a successful contact language and therefore continue to have a 0% success rate.

Indeed, it should be noted here that figure 7.3 is the result of averaging communicative success over multiple simulation runs. However, as stated above, the agents in this simulation do not begin to play the pidgin and creole games until all three populations have had a communicative system with a 100% success rate in place for five generations. The amount of time that it takes for all three populations to achieve this varies drastically between runs, and this is why figure 7.3 does not show any of the trend lines reaching a perfect 50 (which is what B was set to during these runs) before the agents go on to play the creole game. It is also why the P_{CMM} with $P_{CMI} = 1$ case shows a slight increase in success in the last few generations. Thus, unless explicitly stated otherwise, the rest of the graphs contained within this chapter will depict the results for a typical run of the model. This will allow for an analysis of creole emergence. Additionally, results are only shown for population two, as the results presented here will be the same for both populations two and three, due to the analysis being conducted on the same communicative system.

It should also be noted that, during all of these runs, P_{MM} and P_{MI} were all set to the highest value to enable the evolution of the base languages to occur relatively quickly. Once the simulation begins the pidgin phase, however, the P_{CMM} and P_{CMI} variables are used in their stead, with the former denoting the number of mature agents from the other population each mature agent is assigned as a partner to, and the latter operating

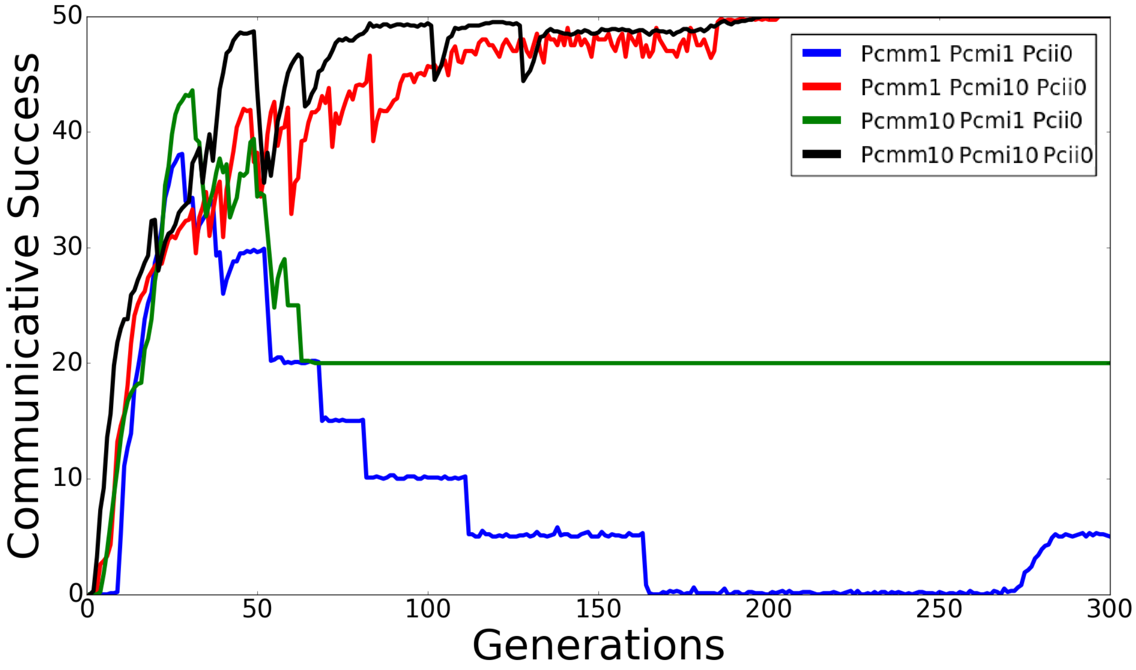


FIGURE 7.3: Graph depicting the amount of communicative success achieved throughout the course of the simulation run by the mature agents of population 2, for P_{CMM}=1, P_{CMI}=1, and P_{CII}=0 (blue line: $\mu = 9.79$, $\sigma = 0.90$); P_{CMM}=1, P_{CMI}=10, and P_{CII}=0 (red line: $\mu = 43.93$, $\sigma = 0.56$); P_{CMM}=9, P_{CMI}=1, and P_{CII}=0 (green line: $\mu = 21.52$, $\sigma = 1.35$); P_{CMM}=9, P_{CMI}=10, and P_{CII}=0 (black line: $\mu = 45.87$, $\sigma = 0.47$). Here, M and I=10 and B=50, and D=50. Averaged over 30 runs. Note that none of the trend lines reach a perfect 50 (which is what B was set to during these runs) before the agents go on to play the creole game due to this graph displaying the average communicative success rate per generation, over a number of runs; with there being fluctuation in behaviour between these runs.

in the exact same way as P_{MI}. Due to newly created immature agents always being blank, and language evolution within this model being a purely social phenomenon, the fact that the partner parameters during the baseline phase are always set to the highest value does not impact the behaviour we see once the agents engage in the pidgin and creole phases.

By looking at figure 7.4, we see that with P_{CMM} and P_{CMI}=1, there is an instant drop in communicative success when the agents enter the pidgin/creole phases. Despite a couple of successful communication attempts, the two populations are never able to develop a linguistic system that enables them to communicate effectively. The same is also true when P_{CMM}=9 with P_{CMI}=1. However, as with the baseline model results, we see that P_{CMM}=1 with P_{CMI}=10 quickly give rise to a successful communicative system. While the P_{CMM}=9 with P_{CMI}=10 case never actually drops all the way to zero during the first creole generation; meaning that the agents are somewhat successful at communicating even during the pidgin game generation. Thus, once again, it appears

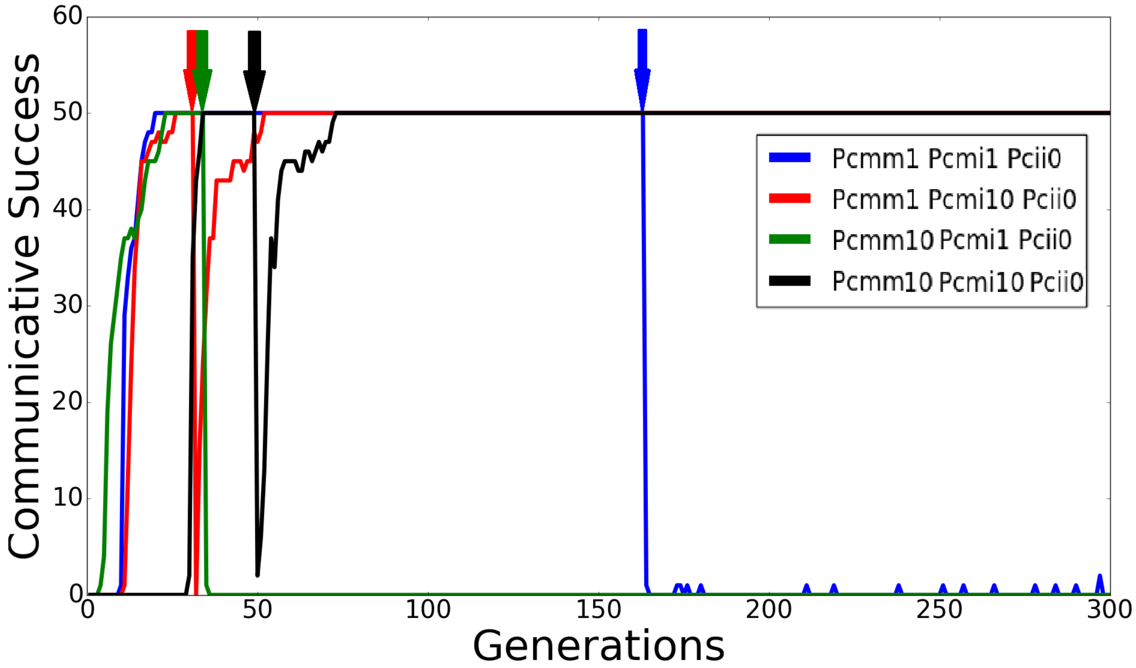


FIGURE 7.4: Graph depicting the amount of communicative success achieved by population 2 durng a single run of the model, for the various possible combinations of P_{CMM}=1 and 9 and P_{CMI}=1 and 10; with M and I = 10, B = 50, and D = 50. For each data set, the first sharp decrease in success rate is the pidgin phase generation, every subsequent generation then sees the agents playing the creole phase. The coloured arrows indicate when the version of the simulation represented by the corresponding coloured line enters phase two and plays the pidgin game, before entering phase three and playing the creole game in all subsequent generations.

that the crucial parameter in the model is the number of mature trainers from which the immature agents learn their language.

The reason that the system exhibits these various behaviours is, again, due to signal diversity. By taking the P_{CMM} parameter as an example, picking an agent at random during generation 205 of the simulation run and looking at each of the grammatical rules within its grammar store, we see that the agent in the P_{CMM} and P_{CMI}=1 case has multiple semantic rules for each of the possible meaning bits that it can experience and multiple structure rules for ordering these semantic rules, while the P_{CMM}=1 with P_{CMI}=10 only has one (figure 7.5). In short, a higher value of P_{CMI} enables a stable creole to emerge by making it easier for a population to establish a convention of usage in regards to symbols (the semantic rules) and syntax (the structure rules).

However, this raises the question as to why this is the case. By comparing figures 7.6 and 7.7, which look at the same behaviour as the last two graphs, but during the first generation of the creole phase of the simulation, we see something interesting. In contrast to the baseline model, increasing P_{CMM} from 1 to 9 has an obvious beneficial impact upon creole semantic emergence; regardless of the value of P_{CMI}.

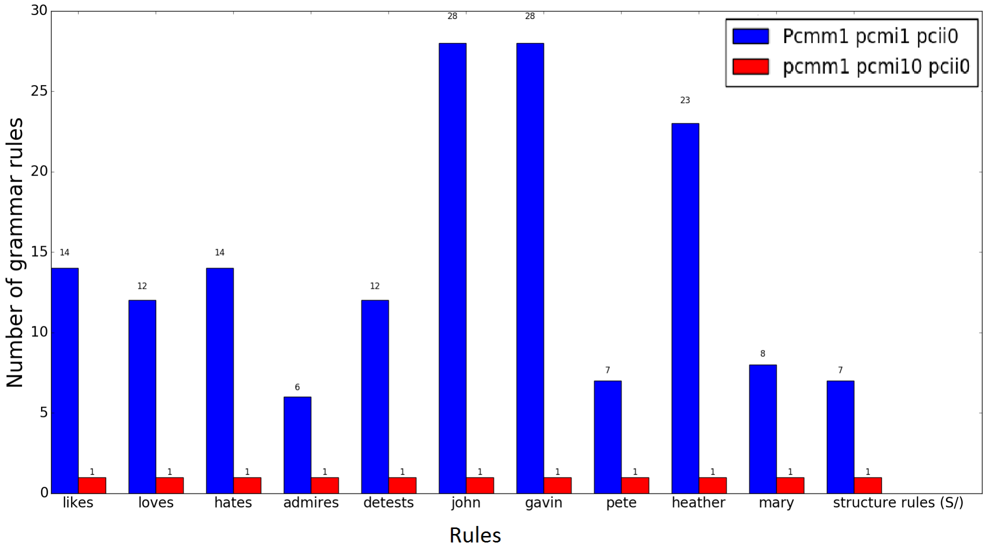


FIGURE 7.5: Graph depicting the number of different rules that a randomly selected agent from generation 205 has within its grammar store for each of the different meaning bits and possible structure rules used by the simulation for P_{CMM}=1 and P_{CMI}=1 (blue bars) and P_{CMM}=1 and P_{CMI}=10 (red bars), with M and I=10, B=50, and D=50.

This is due to the fact that each of the mature agents in this generation were immature agents during the previous generation, which was also the generation where the agents played the pidgin game. This results in this particular batch of agents receiving linguistic training data in the form of pidgin-based signals that originate from a larger number of mature agents playing the pidgin game. This, in turn, results in greater signal diversity, making the agents less idiosyncratic in their communicative system. These less idiosyncratic communicators are then better at communicating with one another as mature agents.

However, in the P_{CMM}=9 case, the mature agents have a better chance of being more successful in using the same signals as their partners than in the P_{CMM}=1 case due to the fact that there is a better chance that the hearer has already encountered the signals used by the speaker. Such success then obviously increases the chance of the agent using the meaning-signal pairs it encountered in future communicative episodes. This, combined with the way in which their children learn from the successful interactions, leads to these immature agents essentially acting as a filter, due to them being more acquainted with the more successful meaning-signal pairs. Thus resulting in the behaviour seen in figure 7.5. This is the exact same behaviour that we saw in the reinforcement learning model.

Although we see that, under certain circumstances, these agents are able to give rise to a successful communicative system, the question remains as to whether or not this new linguistic system is a creole in nature. Figure 7.8 depicts the grammar of a randomly selected mature agent from populations two and three during the final generation of

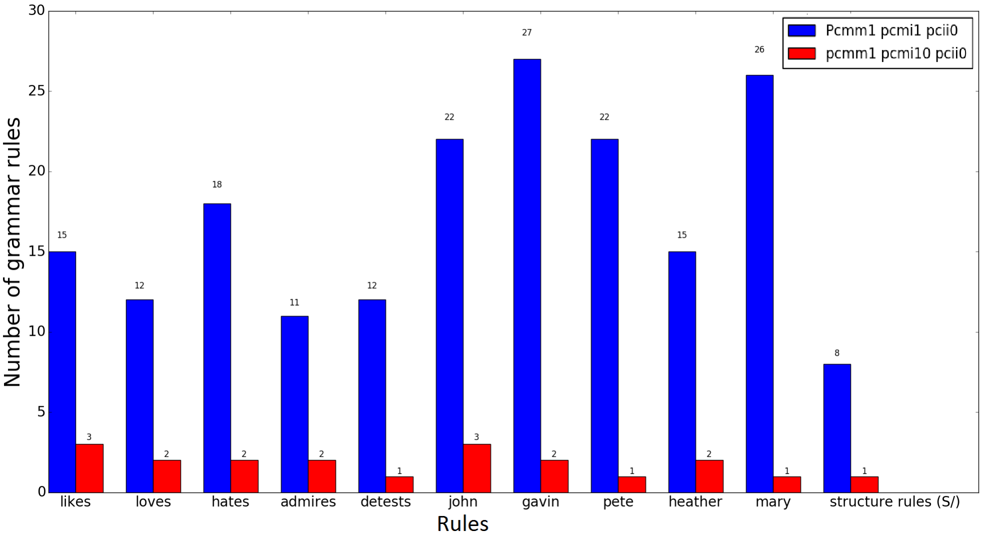


FIGURE 7.6: Graph depicting the number of different signals that a randomly selected agent from the first generation of the creole phase of the simulation has within its grammar store for each of the different meaning bits and possible structure rules used by the simulation for P_{CMM}=1 and P_{CMI}=1 (blue bars) and P_{CMM}=1 and P_{CMI}=10 (red bars), with M and I = 10 and B = 50 and D = 50.

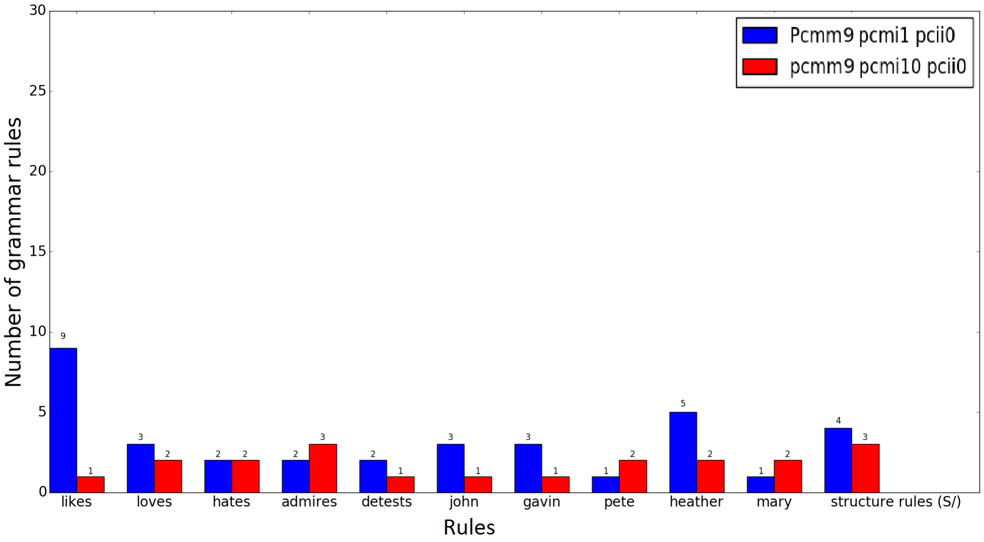
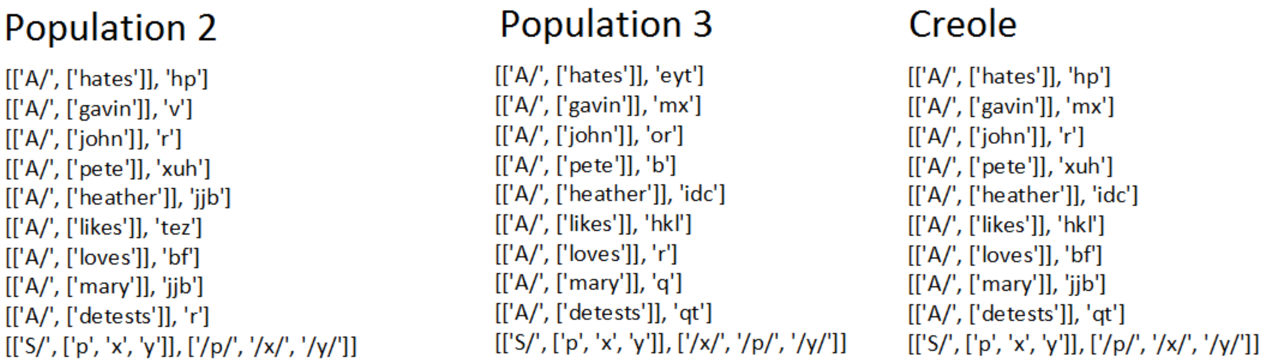


FIGURE 7.7: Graph depicting the number of different signals that a randomly selected agent from the first generation of the creole phase of the simulation has within its grammar store for each of the different meaning bits and possible structure rules used by the simulation for P_{CMM}=9 and P_{CMI}=1 (blue bars) and P_{CMM}=9 and P_{CMI}=10 (red bars), with M and I=10 and B=50 and D=50.

the baseline language game, before the simulation moves onto the pidgin phase. By comparing these two grammar lists to the creole list, which was taken from a randomly selected mature agent once the creole had reached 100% success rate at generation 210, we see that the resulting creole is a mix of the linguistic systems of populations two and three.



Population 2	Population 3	Creole
[['A/', ['hates']], 'hp']	[['A/', ['hates']], 'eyt']	[['A/', ['hates']], 'hp']
[['A/', ['gavin']], 'v']	[['A/', ['gavin']], 'mx']	[['A/', ['gavin']], 'mx']
[['A/', ['john']], 'r']	[['A/', ['john']], 'or']	[['A/', ['john']], 'r']
[['A/', ['pete']], 'xuh']	[['A/', ['pete']], 'b']	[['A/', ['pete']], 'xuh']
[['A/', ['heather']], 'jjb']	[['A/', ['heather']], 'idc']	[['A/', ['heather']], 'idc']
[['A/', ['likes']], 'tez']	[['A/', ['likes']], 'hkl']	[['A/', ['likes']], 'hkl']
[['A/', ['loves']], 'bf']	[['A/', ['loves']], 'r']	[['A/', ['loves']], 'bf']
[['A/', ['mary']], 'jjb']	[['A/', ['mary']], 'q']	[['A/', ['mary']], 'jjb']
[['A/', ['detests']], 'r']	[['A/', ['detests']], 'qt']	[['A/', ['detests']], 'qt']
[['S/', ['p', 'x', 'y']], ['/p/', '/x/', '/y/']]	[['S/', ['p', 'x', 'y']], ['/x/', '/p/', '/y/']]	[['S/', ['p', 'x', 'y']], ['/p/', '/x/', '/y/']]

FIGURE 7.8: List of each of the different grammar rules within a random mature agent from population 2 (left) and population 3 (middle) during the final generation before the agents enter the pidgin phase. The list on the right shows a randomly selected agent's language at generation 210, and with tertiary hybridisation playing a role in the emergence of the creole. It can be seen here that the resulting creole language has adopted features of the languages of both populations two and three. For P_{CMM}=1 with P_{CMI}=10, with M and I=10 and B=50 and D=50.

7.7.2 Exploring The Bottleneck

Given how the bottleneck size affected the rate of compositionality emergence in the iterated learning model from chapter 4, it is worth exploring whether or not we see a similar behaviour here. Thus, figure 7.9 shows the average amount of communicative success achieved over 30 runs with P_{CMM} and P_{CMI}=1 (blue line) and P_{CMM}=1 with P_{CMI}=10 (red line) with a larger bottleneck (B=100).

It can be seen here that the larger bottleneck does not impact the baseline model behaviour very much, but there is a distinct difference in success during the creole phase of the game. Namely, the P_{CMM} and P_{CMI}=1 case achieves a fully stable communicative system after a number of generations. However, this larger bottleneck size does result in the P_{CMM}=1 with P_{CMI}=10 case taking more generations to give to rise to a stable creole language. It should be noted that the step-like decline and increase in communicative success seen in the P_{CMM} and P_{CMI}=1 case is caused by the same behaviour that caused the step-like decline in figure 7.3. Namely, the average score for each generation includes a number of simulation runs that are still achieving 100% success rate due to them not yet entering the pidgin and creole phases of the simulation, while other runs have already entered these phases and are achieving a very low, often 0%, communicative success rate during these same generations. However, unlike in figure 7.3, the larger bottleneck size in the runs from figure 7.9 enables a successful linguistic system to emerge.

This difference in behaviour can be seen more clearly by looking at a single run. Figure 7.10 shows how, in the P_{CMM} and P_{CMI}=1 case, the system is unable to give rise to a successful language in conditions where $B \approx< 80$, but with $B \geq 80$, the system can

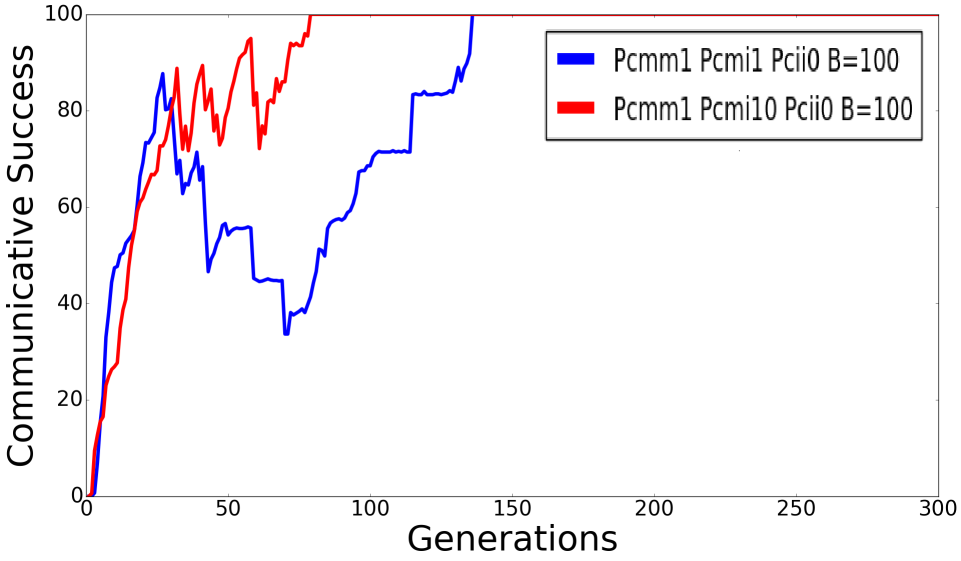


FIGURE 7.9: Graph depicting the difference in the percentage of communicative success for P_{CMM}=1 and P_{CMI}=1 (blue line: $\mu = 82.01$, $\sigma = 1.61$) and P_{CMM}=1 and P_{CMI}=10 (red line: $\mu = 91.97$, $\sigma = 0.85$), with M and I=10 and B=100 and D=50. Averaged over 30 runs. Note that none of the trend lines reach a perfect 100 (which is what B was set to during these runs) before the agents go on to play the creole game due to this graph displaying the average communicative success rate per generation, over a number of runs; with there being fluctuation in behaviour between these runs.

give rise to a linguistic system after a certain number of generations. However, this language is not a creole, but a completely novel linguistic system that has been developed by these agents; a phenomenon that will be discussed in more detail below. In contrast, in the P_{CMM}=1 with P_{CMI}=10 case, the system achieves 100% success rate the vast majority of the time (figure 7.11).

Although these results are interesting insomuch as they support the results and conclusions drawn from the previous two models, the aim of this model was to explore some of the open questions within the creole literature. Specifically, there were two questions that this model aimed to explore; whether or not one group has to be more socially powerful than another for a pidgin and creole to emerge, and whether or not tertiary hybridisation was a necessary prerequisite for creole emergence.

7.7.3 Social Power

As stated above, one of the key ideas encountered in the pidgin and creole literature concerns the idea of social power. This is the belief that the reason creoles tend to gain most of their semantics from the lexifer language is due to the users of that language having a greater amount of social influence over those using the substrate languages; as often seen in colonising scenarios.

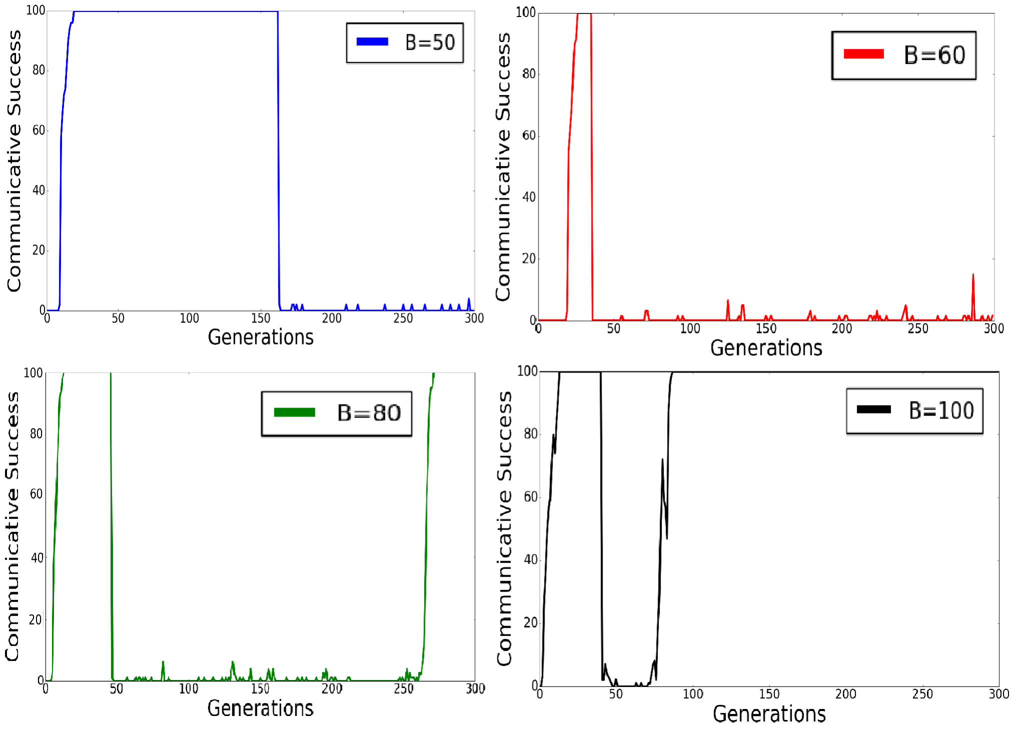


FIGURE 7.10: Graph depicting communicative success of a single run for P_{CMM} and P_{CMI}=1 with P_{CII}=0, when B=50 (top left), 60 (top right), 80 (bottom left), and 100 (bottom right); with M and I=10 and D=50. For each data set, the first sharp decrease in success rate is the pidgin phase generation, every subsequent generation then sees the agents playing the creole phase. The coloured arrows indicate when the version of the simulation represented by the corresponding coloured line enters phase two and plays the pidgin game, before entering phase three and playing the creole game in all subsequent generations.

In order to explore this notion, the simulation was built with a D parameter, as described above. Although multiple tests were conducted, altering this parameter did not result in any change in the amount of communicative success (figure 7.12). Indeed, as above, the crucial parameter is still the number of mature agents from which each immature agent learns their language; and as this increases, so does the amount of communicative success.

7.7.4 Tertiary Hybridisation

The role of tertiary hybridisation was explored by running the simulation without population three. In other words, the simulation proceeds exactly as described above, going through the baseline, pidgin and creole phases. The only difference being that, during the creole phase, populations one and two go straight from playing the pidgin game to playing the creole game with one another.

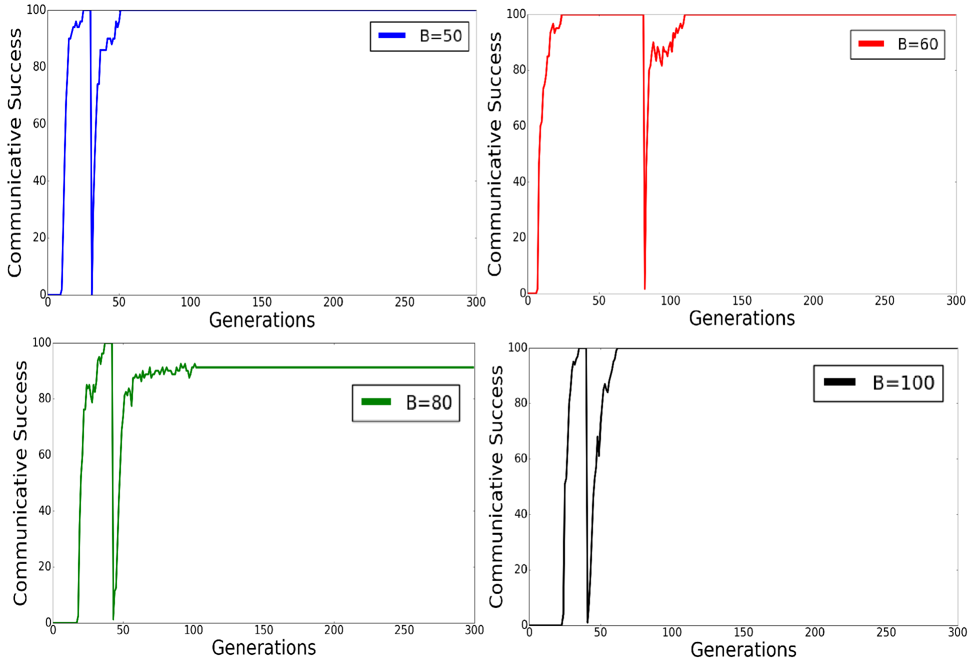


Figure 7.11: Graph depicting communicative success of a single run for P_{CMM}=1 with P_{CMI}=10 with P_{CII}=0, when B=50 (top left), 60 (top right), 80 (bottom left), and 100 (bottom right); with M and I=10 and B=100 and D=50. For each data set, the first sharp decrease in success rate is the pidgin phase generation, every subsequent generation then sees the agents playing the creole phase. The coloured arrows indicate when the version of the simulation represented by the corresponding coloured line enters phase two and plays the pidgin game, before entering phase three and playing the creole game in all subsequent generations.

It can be seen from figure 7.13 that a lack of tertiary hybridisation has a notable effect upon communicative success of the system in cases with a higher value of P_{CMI}; with tertiary hybridisation enabling the populations to give rise to a stable creole language.

This arguably lends some support to the tertiary hybridisation argument seen in the literature. The literature thus far has been unable to present any empirical evidence as to why tertiary hybridation would play the crucial role that many believe to, mostly due to the problems that pidgin and creole researchers experience in documenting creole emergence. However, by again looking at the internal grammars of randomly selected agents, this simulation offers some insight into how this phenomenon works within this particular model. In contrast to figure 7.8, figure 7.14 shows the same information but for a simulation run where there was no tertiary hybridisation. Here, we see that the resulting creole language is essentially an entirely new language, with only a couple of grammar rules originating from the lexifer or substrate languages. This is due to the way in which, despite playing the pidgin game with one another, both populations were still using their already well-established baseline languages, the signals of which continued to be more dominant than the linguistic conventions established during the pidgin-game

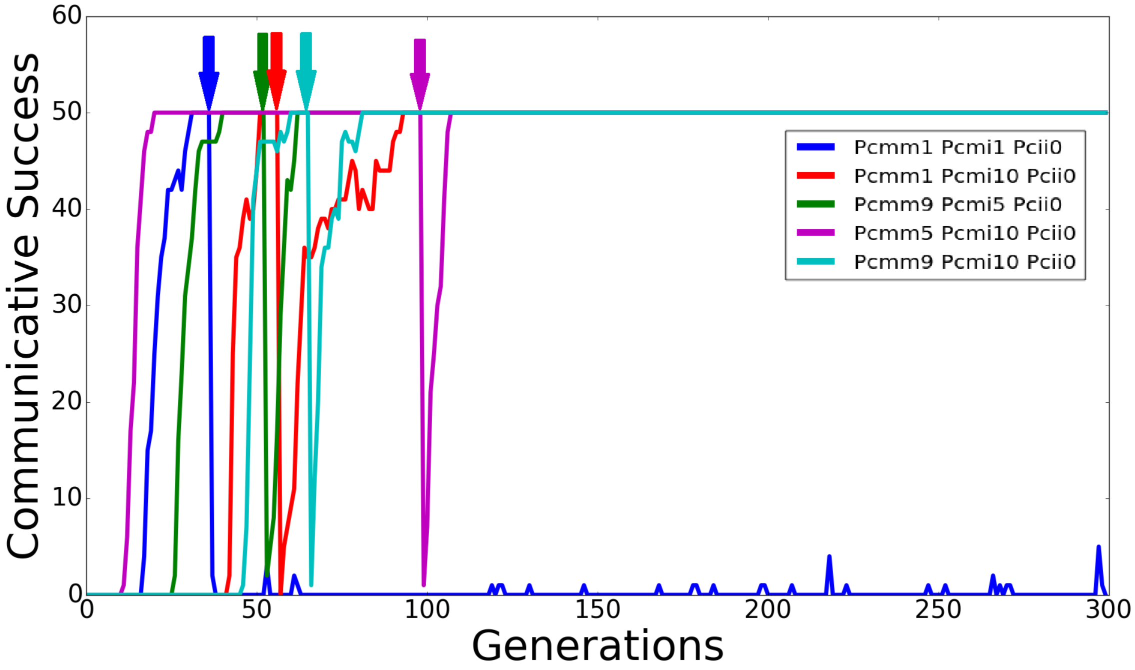


FIGURE 7.12: Graph depicting the amount of communicative success achieved by a single of the model for P_{CMM} and P_{CMI}=1 (blue line), P_{CMM}=1 with P_{CMI}=10 (red line), P_{CMM}=9 with P_{CMI}=5, P_{CMM}=5 with P_{CMI}=10, and P_{CMM}=9 with P_{CMI}=10, with M and I=10 and B=50 and D=100. For each data set, the first sharp decrease in success rate is the pidgin phase generation, every subsequent generation then sees the agents playing the creole phase. The coloured arrows indicate when the version of the simulation represented by the corresponding coloured line enters phase two and plays the pidgin game, before entering phase three and playing the creole game in all subsequent generations.

generation of the simulation; largely due to the amount of exposure that agents had to them.

Thus, when the simulation went on to the creole phase, both populations continued to use the meaning-signal pairs that they knew to be most successful. This resulted in the two populations having to re-negotiate and create an entirely new language in order to conduct inter-population communication.

This is supported by figures 7.15 and 7.16. Here, it can be seen that both with and without tertiary hybridisation, the agents in both populations have a pidgin that allows them to communicate with one another more than 90% of the time. However, without hybridisation, we get the resulting language we see in figure 7.14.

Figure 7.17 shows the no tertiary hybridisation version of the model being run under the same parameter settings, but with D=100. In comparing this graph to figure 7.12, we see that there is no difference in communicative success for the P_{CMI}=10 case. After a prolonged period of time, the P_{CMI}=1 case appears to achieve a 100% success rate. However, as in the above case, this is not a creole, but the result of the agents giving rise

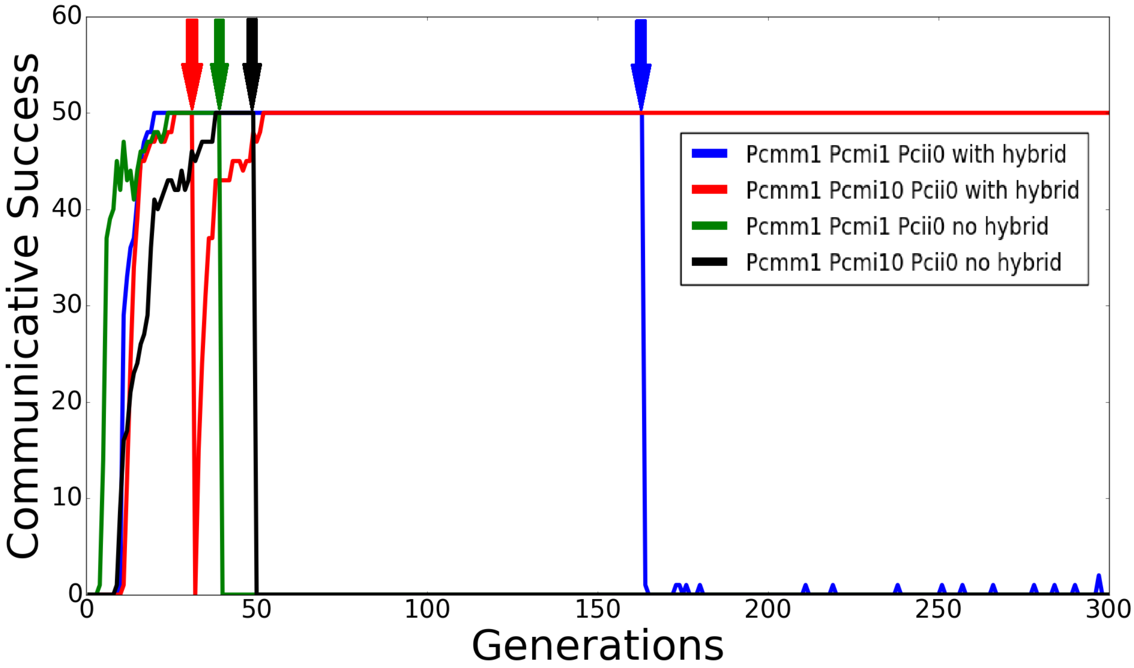


FIGURE 7.13: Graph depicting the amount of communicative success achieved by a single run of the model for $P_{CMM}$and P_{CMI}=1 with tertiary hybridisation playing and role (blue line) and without tertiary hybridisation (green line), and likewise for P_{CMM}=1 with P_{CMI}=10 (red line and black line, respectively), with M and I=10 and B=50 and D=50. For each data set, the first sharp decrease in success rate is the pidgin phase generation, every subsequent generation then sees the agents playing the creole phase. The coloured arrows indicate when the version of the simulation represented by the corresponding coloured line enters phase two and plays the pidgin game, before entering phase three and playing the creole game in all subsequent generations.

to a completely new linguistic system that is different from either population's previous linguistic systems.

7.8 Conclusions

The model presented in this chapter yielded a number of interesting results. The first concerns the role of social power, whereby the speakers of the lexifer adopt a passive role in the linguistic process. It was shown that was not the case in the model presented here, and that pidginisation had to be an interactive process involving multiple communicative efforts on the part of all parties involved. This is a finding that is supported by the fact that contact languages in the real world can emerge between trading partners of equal social power, with vocabulary being drawn equally from both languages. Indeed, the results gained from the pidgin/creole model in this chapter demonstrated that altering the dominance, D, parameter had no significant impact upon the behaviour of the model.

Population 1

```
[['A/', ['loves']], 'm']
[['A/', ['john']], 'ay']
[['A/', ['mary']], 'nlj']
[['A/', ['pete']], 'nlj']
[['A/', ['heather']], 'om']
[['A/', ['admires']], 'ai']
[['A/', ['gavin']], 'f']
[['A/', ['detests']], 'r']
[['A/', ['likes']], 'cm']
[['A/', ['hates']], 'x']
[['S/', ['p', 'x', 'y']], ['/x/', '/p/', '/y/']]
[['S/', ['p', 'x', 'y']], ['/y/', '/p/', '/x/']]
```

Population 2

```
[['A/', ['loves']], 'gtm']
[['A/', ['john']], 'h']
[['A/', ['mary']], 'fe']
[['A/', ['pete']], 'l']
[['A/', ['heather']], 'cr']
[['A/', ['admires']], 'h']
[['A/', ['gavin']], 'l']
[['A/', ['detests']], 'hv']
[['A/', ['likes']], 'bot']
[['A/', ['hates']], 'h']
[['S/', ['p', 'x', 'y']], ['/y/', '/p/', '/x/']]
[['S/', ['p', 'x', 'y']], ['/y/', '/x/', '/p/']]
[['S/', ['p', 'x', 'y']], ['/p/', '/y/', '/x/']]
[['S/', ['p', 'x', 'y']], ['/x/', '/p/', '/y/']]
```

Creole

```
[['A/', ['loves']], 'xb']
[['A/', ['john']], 'l']
[['A/', ['mary']], 'ru']
[['A/', ['pete']], 'l']
[['A/', ['heather']], 'xbh']
[['A/', ['admires']], 'xb']
[['A/', ['gavin']], 'l']
[['A/', ['detests']], 'xb']
[['A/', ['likes']], 'xb']
[['A/', ['hates']], 's']
[['S/', ['p', 'x', 'y']], ['/p/', '/y/', '/x/']]
[['S/', ['p', 'x', 'y']], ['/p/', '/x/', '/y/']]
```

FIGURE 7.14: List of each of the different grammar rules within a random mature agent from population 2 (left) and population 3 (middle) during the final generation before the agents begin to play the pidgin phase. The list on the right shows the same for a randomly selected agent's creole language at generation 210, with tertiary hybridisation not playing a role in the emergence of the creole. For P_{CMM}=1 with P_{CMI}=10, with M and I=10 and B=50 and D=50.

Population 1

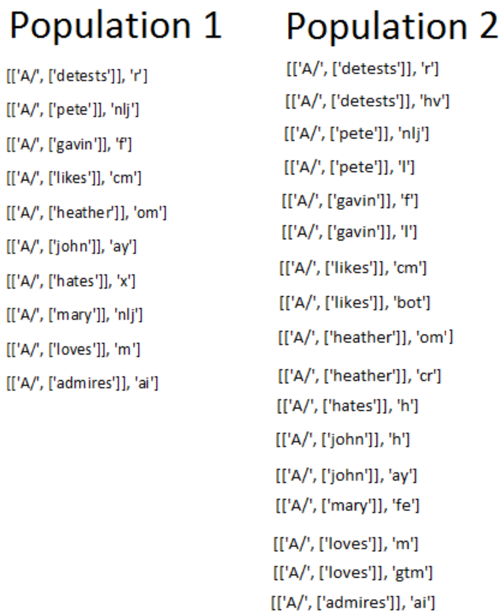
```
[['A/', ['mary']], 'y']
[['A/', ['john']], 'gj']
[['A/', ['admires']], 'hy']
[['A/', ['heather']], 'hzc']
[['A/', ['gavin']], 'gy']
[['A/', ['loves']], 'st']
[['A/', ['pete']], 'd']
[['A/', ['hates']], 'msh']
[['A/', ['likes']], 'kww']
[['A/', ['mary']], 'd']
[['A/', ['gavin']], 'da']
[['A/', ['detests']], 'bcv']
[['A/', ['john']], 'm']
[['A/', ['hates']], 'vq']
```

Population 2

```
[['A/', ['mary']], 'y']
[['A/', ['john']], 'gj']
[['A/', ['admires']], 'y']
[['A/', ['heather']], 'lka']
[['A/', ['gavin']], 'da']
[['A/', ['loves']], 'jnc']
[['A/', ['pete']], 'ord']
[['A/', ['hates']], 'msh']
[['A/', ['likes']], 't']
[['A/', ['detests']], 'tl']
```

FIGURE 7.15: List of each of the different grammar rules within a random mature agent's pidgin grammar for population 1 (left) and population 2 (right) during the final generation before the agents begin to play the creole phase, with tertiary hybridisation playing a role. For P_{CMM}=1 with P_{CMI}=10, with M and I=10 and B=50 and D=50.

Secondly, it has been shown here that tertiary hybridisation results in the simulation giving rise to a creole that is a mix of the languages of both populations two and three. While without tertiary hybridisation, the resulting language is essentially a whole new linguistic system.



Population 1

```
[['A/', ['detests']], 'r']
[['A/', ['pete']], 'nlj']
[['A/', ['gavin']], 'f']
[['A/', ['likes']], 'cm']
[['A/', ['heather']], 'om']
[['A/', ['john']], 'ay']
[['A/', ['hates']], 'x']
[['A/', ['mary']], 'nlj']
[['A/', ['loves']], 'm']
[['A/', ['admires']], 'ai']
```

Population 2

```
[['A/', ['detests']], 'r']
[['A/', ['detests']], 'hv']
[['A/', ['pete']], 'nlj']
[['A/', ['pete']], 'l']
[['A/', ['gavin']], 'f']
[['A/', ['gavin']], 'l']
[['A/', ['likes']], 'cm']
[['A/', ['likes']], 'bot']
[['A/', ['heather']], 'om']
[['A/', ['heather']], 'cr']
[['A/', ['hates']], 'h']
[['A/', ['john']], 'h']
[['A/', ['john']], 'ay']
[['A/', ['mary']], 'fe']
[['A/', ['loves']], 'm']
[['A/', ['loves']], 'gtm']
[['A/', ['admires']], 'ai']
```

FIGURE 7.16: List of each of the different grammar rules within a random mature agent's pidgin grammar for population 1 (left) and population 2 (right) during the final generation before the agents begin to play the creole phase, with tertiary hybridisation not playing any role. It is shown here that, without tertiary hyrbridisation, the agents are still able to establish a pidgin that can achieve a success rate of roughly 90%. For P_{CMM}=1 with P_{CMI}=10, with M and I=10 and B=50 and D=50.

We see that the presence of tertiary hybridisation results in one of the languages acting as the lexifer and another as the substrate language. It is arguable that the lexifer/substrate effect, which has typically been associated with social power, only plays a significant role in cases where tertiary hybridisation takes place.

This would make logical sense given that, in a tertiary hybridisation scenario, the third social group would have no understanding of the language of either population one or population two. This would result in them taking on more of the pidgin vocabulary being presented to them by population two and contributing less vocabulary from their own language. Indeed, we have seen the argument made that contact languages will only develop with stable norms, grammar, pronunciation, and vocabulary, once they are taken up by a third party (Muhlhausler 1986). This argument is supported by the fact that most of the successful pidgins, those with the longest histories and spoken most widely, are also *lingua francas* used by speakers of large numbers of mutually unintelligible languages.

Finally, it can also be seen that, while communication between higher numbers of mature language users allow for faster establishment of semantic conventions, such as seen in cases of pidgin emergence, it is really the inter-generational transmission of language through a bottleneck that fosters language development through its exposure to different learner biases. This explains why the transformation from pidgin to creole tends to

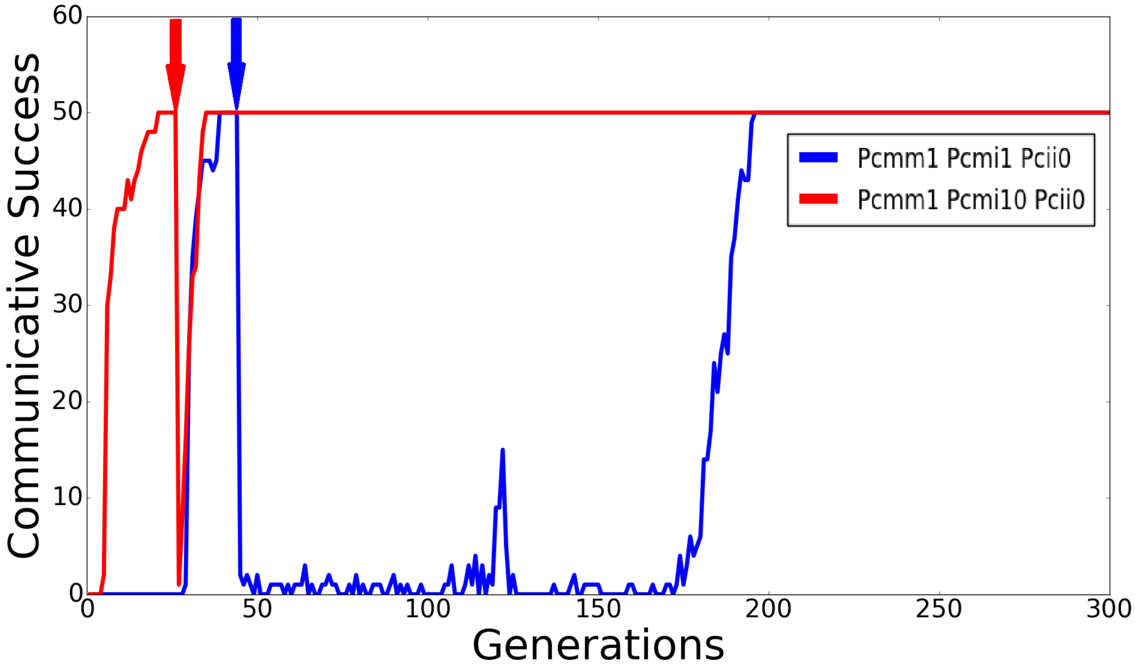


Figure 7.17: Graph depicting the amount of communicative success achieved by a single run of the model for P_{CMM}=1 and P_{CMI}=1 (blue line) and 10 (red line) without tertiary hybrisation playing and role (blue line) and D = 100, with M and I = 10 and B = 50. For each data set, the first sharp decrease in success rate is the pidgin phase generation, every subsequent generation then sees the agents playing the creole phase. The coloured arrows indicate when the version of the simulation represented by the corresponding coloured line enters phase two and plays the pidgin game, before entering phase three and playing the creole game in all subsequent generations.

happen at an incredibly fast rate; typically within a single generation. Indeed, it was mentioned above that, at any stage during the pidgin-creole developmental process, pidgins become creoles once immature language learners are born into a social milieu with a pidgin that they end up acquiring as their primary language.

This notion is largely supported by the way in which increasing P_{CMM} from 1 to 10 had an obvious beneficial impact upon creole emergence, regardless of the value of P_{CMI}. This was due to the fact that each of the mature agents in the first generation of the creole phase were immature agents during the previous generation, which was also the generation where the agents played the pidgin game. This resulted in this particular batch of agents receiving linguistic training data in the form of pidgin-based signals that originate from a larger number of mature agents playing the pidgin game. This in turn resulted in greater signal diversity. Thus, making the agents less idiosyncratic in their communicative system.

This idea, supported by the results presented here, also relates back to the 'linguistic blueprint' argument. Specifically, the results presented here make it arguable that, instead of there being any specific 'linguistic blueprint' for creoles, these languages possess the structure they do as a result of being exposed to learner biases that shape the languages in order to be more learnable. This results in immature language learners essentially levelling out variability and producing more regular grammars, while not actually innovating. This is supported by the way in which the children of plantation labourers, who lived within a diverse socio-linguistic milieu, developed a linguistic system based upon the mixed and inconsistent linguistic input that they received, and not based upon that of their parents (Roberge 2013).

Indeed, as will be seen in the following chapter, this kind of linguistic evolutionary behaviour has been present in all of the models presented in this thesis.

Chapter 8

Conclusions

8.1 Introduction

This thesis was motivated by the two umbrella questions that arguably underpin all research into human language evolution (Tallerman & Gibson 2013*b*). The first concerning how much of the human language faculty can be attributed to a genetic endowment for language learning, and to what extent can it be accounted for by other mechanisms; such as the nature of observational learning and social interactions? The second concerns how much of linguistic structure can be attributed to natural selection, such as which linguistic features were adaptive to early hominins, and to what extent does linguistic structure arise from self-organisational processes; such as languages themselves adapting to be easier to learn? Indeed, it was demonstrated in the early parts of this thesis that, while many researchers would agree that human language is the result of both innateness and ontogenetic development, there are vast differences of opinion in regards to what, and where, lines should be drawn.

However, it is clear that human language is strongly influenced by biology, and the case of the Nicaraguan school children, who developed a novel sign language in the absence of any pre-existing linguistic system (Senghas et al. 2005) is testament to this. However, given the rate at which lexical items change, it is clear that languages themselves are social-cultural phenomena. Interestingly, we saw evidence in chapter 2 that contemporary data also demonstrates that context-specific learning behaviours, akin to those seen in humans, have evolved in a range of taxa. In addition, the species that exhibit such behaviours also possess communication systems that are heavily influenced by social interactions and cultural learning.

This indicates that the development of such social learning mechanisms can be a response to specific evolutionary selection pressures; likely to be social in nature. Additionally, given the way in which mental representations and awareness of self and others are also well documented within the primate lineage, indicating that such traits are phylogenetically much older than *Homo sapiens* and were likely to have been in place prior to the Pan/Homo genera split, it could be argued that individual-based social learning plays a significant role in language learning in our species.

Overall, these observations suggest that any study of the evolution of human language should view it as being the result of three interacting dynamical systems; biology, ontogenetic growth, and social-cultural factors (Kirby & Hurford 2002, Kirby 2002*b*). Indeed, while it is clear that the biological and neurological apparatus that enables linguistic evolution is what endowed early homo with certain linguistic abilities, the earliest forms of protolanguage must also have been social-cultural phenomena learned by individuals during their lifetimes. Further, it is entirely plausible that the biological and social aspects of language eventually co-evolved to result in the biological/neurological apparatus and linguistic structures we see in the world today.

This is the theoretical foundation that this thesis adopted when exploring the above two questions. Once combined with contemporary data indicating a relationship between population make-up and linguistic structure (Johansson 1997, Wichmann & Holman 2009, Lupyan & Dale 2010, Milroy 2013, Trudgill 2013, Atkinson et al. 2015) and a shift in social dynamics being related to the evolutionary origins of human language (Dunbar 1993, Aiello & Dunbar 1993, Kudo & Dunbar 2001, Fitch 2010), it makes logical sense to frame these issues in terms of social co-ordination.

As such, the primary focus of this thesis was on expression/induction (E/I) models that adhered to the *iterated learning* framework. This particular flavour of agent-based model was selected for two reasons. Firstly, iterated learning models (ILMs) are specifically tailored to explore how linguistic structures evolve through self-organisational processes; how languages are affected by learner bias. Secondly, they have traditionally been characterised by both their particularly small population sizes, often just one mature and one immature agent per generation, and the way in which they aim to explore the nature of the emergence and evolution of mappings between meaning-signal pairs.

This second characteristic constituted a clear research gap; one that this thesis was well-positioned to explore. Although a large body of work existed that demonstrate how agents within such models force languages to evolve due to the way in which their individual learner biases impose selection pressures on it, such behaviour had not been extensively explored in larger, structured, populations with varying dynamics. Indeed, E/I models have traditionally not been concerned with matters of social structure or social dynamics, and their potential to impact linguistic form. Instead, the agent populations within such models tend to be both constant and small in size (Hurford 2000, 2002, Kirby 2002*c*, Kirby & Hurford 2002). This, from a computational research perspective, was a particularly important aspect of E/I models to explore. Particularly given the way in which early A-life research demonstrated the importance of co-ordinated behaviour in the emergence of communication systems (Werner & Dyer 1992, MacLennan & Burghardt 1994).

8.2 Contributions

This thesis had four specific aims to explore. Firstly, what impact, if any at all, does population size have upon the dynamics of iterated learning? Secondly, if population size does affected iterated learning dynamics, then in what manner does the make-up of said population impact linguistic structure? Thirdly, what impact does a fluctuating population size have upon the rate and nature of linguistic change within an iterated learning model? Finally, could iterated learning models be used in order to explore the nature of contact language emergence? The models presented here demonstrated the following for each of these four research questions.

8.2.1 Does population size have an impact upon the dynamics of iterated learning over generational time?

As a starting point, a well-documented iterated learning model, originally by Kirby & Hurford (2002), was re-implemented and extended to include multiple mature and immature agents per generation. Despite intuition dictating that the presence of more speaker agents within the population would make it more difficult for a stable linguistic system to emerge was, this model demonstrated that this is incorrect.

It did hold true that the various mature agents in the initial generation used different signals for the same meaning, and that this created inconsistent training input data for immature language learners. However, exposing these immature agents to a number of different linguistic tutors actually resulted in them coalescing this varied input. This, in turn, resulted in these immature agents having similar languages upon being promoted to mature agent status. These agents thus produced similar linguistic output to one another, which was then used to train the second generation of immature language users; and so on. In other words, while the initial language learning task was made more problematic for each individual language learner by the presence of multiple language trainers, the language itself benefitted from the rapid regularisation that this resulted in. This is due to the way in which language learners in the model benefitted from learning signals from multiple trainers, due to the way in which they see that multiple signals can be used for a particular meaning. However, due to the way in which the agent networks and obverter procedure operate, they are quickly able to pick the most commonly used signal for such a meaning.

8.2.2 In what manner does the make-up of a population of agents impact linguistic structure within iterated learning models?

In chapter 4, an ILM featuring an alternative learning mechanism was explored in order to discover whether the results in chapter 4 would generalise to a different setting. This model built upon previous work in exploring how simple signals can emerge as successful communicative devices as long as there is a stable and conventionalised system of usage (Skyrms 2004, 2010, Barrett 2006, 2009). This second model supported many of the findings reported chapter 4.

First, although it was initially harder for the agents in this latter model to establish a convention of usage for their relatively simplistic signals, once they had successfully done so, the resulting language was far more successful at being transmitted to subsequent generations of agents. This was shown to be the result of the linguistic training input being presented to immature agents. It is true that this training input is initially more diverse in situations where said agents are learning their version of the language from higher numbers of different mature agents. However, learning from multiple mature agents resulted in these immature agents developing a far less idiosyncratic language. This greatly increased their chances of successfully communicating with other agents once they were promoted to mature agent status. This increased chance of success, in turn, resulted in a particular convention being adopted by a wider number of agents, and in turn, resulted in the communicative system of the agents being far more stable across generations.

Second, this model also demonstrated that this increased success rate was even more efficient in situations where mature agents only interacted with a small number of other mature agents, while immature agents are tutored by a higher number of different mature agents. This is due to the way in which the mature agents that communicate with only a small number of other mature agents give rise to a well-established usage convention due to them having less variety in their potential outputs, compared to mature agents who interact with a large number of other mature agents. This, in turn, resulted in the immature agents having fewer potential signals to choose from when attempting to convey a meaning to their peers. A benefit that was even greater to the immature agents in subsequent generations.

These findings suggest that a key factor in the emergence and development of an expressive and stable communicative system was the number of trainers from which immature language users learned their language.

8.2.3 Does a fluctuating population size impact the rate of linguistic change within an iterated learning model?

There were two main conclusions to be drawn from the model presented in chapter 6. First, it yielded data that supported results gained from the first two models, in that the amount of diversity within an immature language learner's linguistic training data played a crucial role in how quickly, if at all, an expressive and stable linguistic system emerged within a population of agents.

This observation was most notable in this third model, when the agent population was contracting or expanding. Here, it was found that if the number of assigned communicative partners is held constant at a low amount of population interconnectivity, then a stable and successful linguistic system not only emerges, but does so at a much faster rate than in smaller populations. In short, this model demonstrated how a higher percentage of population interconnectivity aids a linguistic group in giving rise to a stable and successful language.

Secondly, and following on from the first finding, the model in chapter 6 produced data that was in-line with much of the contemporary work in the literature. Work that argues how it is not the size of the population *per se* that influences language emergence and change. Moreover, given that varying the social network structure yielded different results, linguistic changes within this model are linked to the connectivity of a linguistic social group. In other words, an individual is less likely to exhibit linguistic variations when they have a higher level of integration within the social group (Ke et al. 2008, Wichmann & Holman 2009); i.e. connected to a higher number of other language users.

8.2.4 Can iterated learning models be adapted in order to explore the development of contact languages, and if so, what inferences can be made from the data yielded by such simulations?

It has been argued elsewhere (Botha 2003, 2006, Roberge 2009, Tallerman & Gibson 2013*a*) that the nature of the emergence and evolution of contact languages could offer a lot to researchers in regards to understanding the nature of linguistic change. This is a reasonable assertion, given the evidence in favour of languages themselves changing and adapting in relation to the biases of their users. After all, despite some of the issues surrounding the nature of second language acquisition, it is highly unlikely that we would possess a completely different, and secondary, set of language learning mechanisms; either social or biological. As such, it was deemed interesting to explore these issues using the iterated learning framework, which was now a possibility given the nature of the model presented in chapter 6.

As such, the model presented in chapter 7 adapted the model from the preceding chapter in order to explore some of the open questions encountered in the contact language literature. When tailored to mimic the social circumstances under which pidgins and creoles come into existence, this model displayed a number of interesting behaviours.

First, it suggested that the idea of there needing to be an in-balance in terms of social power in order for a pidgin to arise may not hold true. In the model, pidgin formation had to be bilateral, involving multiple communicative efforts by both language communities involved. This finding is supported by the way in which such contact languages come into being through communicative episodes between social groups of equal power. As is the case with trading partners and sea-fearers (Holm 2000).

Secondly, this model was also used in order to explore the enigmatic role of tertiary hybridisation in pidginisation and creole origins. Specifically, when social circumstances akin to tertiary hybridisation were forced upon the agents during the creole phase of the simulation, these agents gave rise to a linguistic system that was a blend of that used by both groups in order to communicate with members of their respective social group. In comparison, without tertiary hybridisation, the agents had to establish a completely novel linguistic system in order to communicate between social groups. Thus, in the context of the model in chapter 7, tertiary hybridisation was a crucial component in the establishment of a successful creole. A conclusion that is supported by the empirical evidence that indicates how users of such a contact language will only establish a between-groups language, with stable norms and grammar, once a third party is included in the equation (Whinnom 1968, Sebba 1997).

Finally, this model adds more weight to the overall argument of this thesis. Namely, that it is the inter-generational linguistic transmission between mature and immature agents, through a linguistic bottleneck, that truly fosters linguistic change as a result of the language being exposed to learner biases in E/I models. Although it had been demonstrated before by previous ILM-based work that the linguistic bottleneck is a key factor in regards to whether or not languages can be successfully passed from one generation, and how the bottleneck can dictate the nature of the resulting linguistic structure (Kirby 2002*b*,*a*, Kirby & Hurford 2002, Kirby et al. 2014, Smith 2002, Smith et al. 2003), the data presented in this thesis clearly demonstrates that another crucial factor is the amount of signal variety that is passed through the linguistic bottleneck.

Furthermore, by using contact languages as a case study, where it has been observed that children of plantation labourers were quickly able to form a coherent and expressive linguistic systems from varied and inconsistent input (Roberge 2013), it is possible to argue that, what has traditionally been mistaken for humans creating languages in accordance with some form of innate ‘linguistic blueprint’ may have actually been the result of linguistic systems themselves being refined through usage in accordance with learner bias.

8.3 Future Work

Although the models presented here were novel in a number of ways and produced a number of interesting findings, there is still a lot of further work that could be done in this area. Specifically, in regards to two key areas.

First, a particular characteristic of E/I models that could provide further interesting results is the way in which they, rather unrealistically, create a noise-free environment (Hurford 2002). In other words, every utterance produced by a speaker is perfectly observed by the hearer, and language learners are assumed to have perfect access to the meanings expressed by the speakers that they are exposed to. As such, E/I models arguably fail to capture, and take into account, a potentially significant source of linguistic change.

Secondly, given the aforementioned literature concerning how social network structures impact linguistic change, valuable results could be gained from taking the models presented here and altering the population structure to mimic different network structures that are discussed in the network theory literature (Newman 2010, Milroy 2013); such as random and scale-free networks, etc.

8.4 Summary Of Thesis

This thesis aimed to explore how much of linguistic structure could be accounted for by self-organisational processes and ontogenetic social learning, by utilising a specific form of agent-based model. In doing so, the models it presented enable a number of conclusions to be drawn.

Firstly, once a species has productive and interpretive mechanisms in place, as the available data indicates was the case long before the Pan/Homo genera split, then it is relatively easy for communicative systems to emerge through a social process of conventionalisation. Communicative systems may then be refined by way of social learning on the part of future generations of the species.

Secondly, a higher level of interconnectivity between different individuals within a social group leads to linguistic conventions becoming more stable at a much faster rate. This is due to the way in which such a language will, over time, be increasingly refined by learner bias. Thus resulting in said language adapting at a much faster rate.

In answer to the discussion that began this chapter, while biological evolution clearly played a role in endowing us with the mechanisms necessary to produce and interpret signals, the work presented in this thesis has built upon a larger body of research, and demonstrated how learner bias acts as a selection pressure upon languages. One that results in linguistic structures themselves evolving so as to be easier to learn.

In short, social-cultural factors in the form of individual learning over generations can have a significant impact upon linguistic structure; as a language is refined in order to be easier for its speakers to learn and use. This refinement process is accelerated when individuals are presented with a more varied linguistic sample, due to the way in which it better enables them to use the most optimal communicative option of the language.

Bibliography

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

Arbib, M. (2005), The Mirror System Hypothesis: How Did Protolanguage Evolve, *in* M. Tallerman, ed., 'Language Origins: Perspectives on Evoultion', Oxford University Press, Oxford.

Arbib, M. A. (2010), The Evolving Mirror System: A Neural Basis for Language Readiness, *in* M. Christiansen & S. Kirby, eds, 'Language Evolution', Oxford University Press, Oxford.

Arbib, M., Liebal, K. & Pika, S. (2008), 'Primate Vocalization, Gesture, and the Evolution of Human Language', *Current Anthropology* **49**(6), 1053–1076.

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, M., Kirby, S. & Smith, K. (2015), 'Speaker Input Variability Does Not Explain Why Larger Populations Have Simpler Languages', *PloS one* **10**(6), 937–962.

Balari, S., Benitez-Burraco, A., Longa, V. & Lorenzo, G. (2013), The Fossils of Language: What Are They? Who Has Them? How Did They Evolve?, *in* K. Grohmann & C. Boeckx, eds, 'The Cambridge Handbook of Biolinguistics', Cambridge University Press, Cambridge.

Barr, D. (2004), 'Establishing Conventional Communication Systems: Is Common Knowledge Necessary?', *Cognitive Science* **28**(6), 937–962.

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

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

Batali, J. (1998), Computational Simulation of the Emergence of Grammar, *in* J. Hurford, M. Studdert-Kennedy & C. Knight, eds, ‘Approaches to the Evolution of Language’, Cambridge University Press, Cambridge.

Belpaeme, T. & Bleys, J. (2007), Language, Perceptual Categories and their Interaction: Insights from Computational Modelling, *in* C. Lyon, C. Nehaniv & A. Cangelosi, eds, ‘Emergence of Communication and Language’, Springer-Verlag, London.

Bickerton, D. (1981), *Roots of Language*, Karoma, Ann Arbor, USA.

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

Bickerton, D. (1995), *Language and Human Behaviour*, UCL Press., London.

Bickerton, D. (1998), Catastrophic Evolution: The Case for a Single Step from Protolanguage to Full Human Language, *in* J. Hurford, M. Studdert-Kennedy & C. Knight, eds, ‘Approaches to the Evolution of Language’, Cambridge University Press, Cambridge.

Bickerton, D. (2000), How Protolanguage Became Language, *in* C. Knight, M. Studdert-Kennedy & J. Hurford, eds, ‘The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form’, Cambridge University Press, Cambridge.

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

Bickerton, D. (2003), Symbol and Structure: A Comprehensive Framework for Language Evolution., *in* M. Christiansen & S. Kirby, eds, ‘Language Evolution’, Oxford University Press, Oxford.

Bickerton, D. (2013), Language and Natural Selection, *in* ‘The Cambridge Handbook of Biolinguistics’, Cambridge University Press, Cambridge.

Boe, L., Maeda, S. & Heaim, J. (1999), ‘Neanderthal Man was not Morphologically Handicapped for Speech’, *Evolution of Communication* **3**(1), 49–77.

Bolhuis, J., Okanoya, K. & Scharff, C. (2010), ‘Twitter Evolution: Converging Mechanisms in Birdsong and Human Speech.’, *Nature Reviews. Neuroscience* **11**(11), 747–59.

Botha, R. (2003), *Unravelling the Evolution of Language*, Elsevier Ltd, Amsterdam, The Netherlands.

Botha, R. (2006), ‘Pidgin Languages as a Putative Window on Language Evolution’, *Language and Communication* **26**(1), 1–14.

Botha, R. (2009), Theoretical Underpinnings of Inferences About Language Evolution: The Syntax Used at Blombos Cave, *in* R. Botha & C. Knight, eds, ‘The Cradle of Language’, Oxford University Press, Oxford.

Boyland, J. (1996), 'Conditional Attribute Grammars', *ACM Transactions on Programming Languages and Systems* **18**(1), 73–108.

Brace, L. & Bullock, S. (2016), Achieving Communicative Success in a Population of Agents Through Reinforcement Learning, *in* C. Gershenson, T. Froese, W. Siqueiros, W. Aguilar, E. Izquierdo & H. Sayama, eds, 'Artificial Life XV: Proceedings of the Fifteenth International Conference on the Synthesis and Simulation of Living Systems', The MIT Press, pp. 492–499.

Brace, L., Bullock, S. & Noble, J. (2015), Achieving Compositional Language in a Population of Iterated Learners, *in* P. Andrews, L. Caves, S. Doursat, S. Hickinbotham, F. Polack, S. Stepney, T. Taylor & J. Timmis, eds, 'Advances in Artificial Life: Proceedings of the Thirteenth European Conference of Artificial Life (ECAL 2015)', The MIT Press, pp. 349–356.

Bradbury, J. (2003), Vocal Communication in Wild Parrots, *in* F. de Waal & P. Tyack, eds, 'Animal Social Complexity: Intelligence, Culture, and Individualized Society', Harvard University Press, London.

Brighton, H. (2002), 'Compositional Syntax from Cultural Transmission', *Artificial Life* **8**(1), 25–54.

Brighton, H. & Kirby, S. (2001), The Survival of the Smallest : Stability Conditions for the Cultural Evolution of Compositional Language, *in* J. Kelemen & Sosk. P., eds, 'Advances in Artificial Life: 6th European Conference', Springer-Verlag, pp. 592–601.

Brighton, H., Kirby, S. & Smith, K. (2005), Cultural Selection for Learnability: Three Principles Underlying the View that Language Adapts to be Learnable, *in* M. Tallerman, ed., 'Language Origins: Perspectives on Evoultion', Oxford University Press, Oxford.

Briscoe, T. (1998), Language as a Complex Adaptive System: Co-evolution of Language and of the Language Acquisition Device, *in* 'Proceedings of the 8th Meeting Of Computational Linguistics In The Netherlands', Rodopi, Amsterdam, The Netherlands, pp. 3–40.

Briscoe, T. (1999), 'The Acquisition of Grammar in an Evolving Population of Language Agents', *Electronic Transactions of Artificial Intelligence (Special Issue: Machine Intelligence, 16)* **3**(B), 44–77.

Bromham, L., Hua, X., Fitzpatrick, T. & Greenhill, S. (2015), 'Rate of Language Evolution is Affected by Population Size', *Proceedings of the National Academy of Sciences* **112**(7), 1–6.

Burkett, D. & Griffiths, T. (2010), Iterated Learning of Multiple Language from Multiple Teachers, *in* A. Smith, M. Schouwstra, B. de Boer & K. Smith, eds, 'The Evolution of

Language: Proceedings of the 8th International Conference.', Vol. 8, World Scientific, London, pp. 58–65.

Byrne, R. & Corp, N. (2004), 'Neocortex Size Predicts Deception Rate in Primates', *Proceedings from the Royal Society of London, Series B* **271**(1549), 1693–1699.

Call, J. & Tomasello, M. (2008), 'Does the Chimpanzee have a Theory of Mind? 30 Years Later', *Trends in Cognitive Science* **12**(5), 187–192.

Calvin, W. (2004), *A Brief History of the Mind*, Oxford University Press, Oxford.

Calvin, W. & Bickerton, D. (2000), *Lingua ex machina: Reconciling Darwin and Chomsky with the Human Brain*, The MIT Press, Cambridge, USA.

Cangelosi, A. (2001), 'Evolution of Communication and Language Using Signals, Symbols, and Words', *IEEE Transactions on Evolutionary Computation* **5**(2), 93–101.

Cangelosi, A., Greco, A. & Harnad, S. (2002), Symbol Grounding and the Symbolic Theft Hypothesis, *in* A. Cangelosi & D. Parisi, eds, 'Simulating the Evolution of Language', Springer-Verlag, London.

Cann, R. (2013), Molecular Perspectives on Human Evolution, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Carre, R. & Mrayati, M. (1995), Vowel Transitions, Vowel Systems, and the Distinctive Region Model, *in* C. Dobrovie-Sorin, ed., 'Levels in Speech Communication: Relations and Interactions: A Tribute to Max Wajskop', Elsevier Science Ltd, London.

Cartmill, E. & Byrne, R. (2010), 'Semantics of Primate Gestures: Intentional Meanings of Orangutan Gestures', *Animal Cognition* **13**(6), 793–804.

Catchpole, C. & Slater, P. (1995), *Bird Song: Biological Themes and Variations*, Cambridge University Press, Cambridge.

Cheney, D. & Seyfarth, R. (1990), *How Monkeys See the World: Inside the Mind of Another Species*, Chicago University Press., Chicago, USA.

Chomsky, N. (1957), *Syntactic Structures*, Mouton, The Hague.

Chomsky, N. (1980), *Rules and Representations*, Columbia University Press, New York, USA.

Chomsky, N. (1986), *Knowledge of Language: Its Nature, Origin, and Use*, Praeger Publishers, Westport, USA.

Christiansen, M. (1990), 'A Survey of Adaptable Grammars', *ACM SIGPLAN Notices* **25**(11), 35–44.

Christiansen, M. & Chater, N. (2008), 'Language as Shaped by the Brain', *Behavioural and Brain Sciences* **31**(5), 489–509.

Christiansen, M. & Kirby, S. (2003), Language Evolution: The Hardest Problem in Science?, *in* M. Christiansen & S. Kirby, eds, 'Language Evolution', Oxford University Press, Oxford.

Clark, R. (1979), *In Search of Beach-La-Mar: Historical Relations Among Pacific Pidgins and Creoles*, Auckland, N.Z. : Dept. of Anthropology, University of Auckland.

Clarke, E., Reichard, U. & Zuberbuhler, K. (2006), 'The Syntax and Meaning of Wild Gibbon Songs', *PloS ONE2* **1**(1), e73.

Clay, Z. & Zuberbuhler, K. (2009), 'Food-associated Calling Sequences in Bonobos', *Animal Behaviour* **77**(6), 1387–1396.

Corballis, M. (2002), *From Hand to Mouth: the Origins of Language*, Princeton University Press, Princeton, USA.

Corballis, M. (2003), From Hand to Mouth: The Gestural Origins of Language, *in* M. Christiansen & S. Kirby, eds, 'Language Evolution', Oxford University Press, Oxford, pp. 201–218.

Corballis, M. (2013), The Origins of Language in Manual Gestures, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Crockford, C. & Boesch, C. (2005), 'Call Combinations in Wild Chimpanzees', *Behaviour* **142**(4), 397–421.

Crockford, C., Herbinger, I., Vigilant, L. & Boesch, C. (2004), 'Wild Chimpanzees Produce Group-specific Calls: A Case for Vocal Learning?', *Ethology* **110**(3), 221–243.

Cysouw, J. (2009), Quantitative Explorations of the Worldwide Distribution of Rare Characteristics, or the Exceptionality of Northwestern European Languages, *in* H. Simon & H. Wiese, eds, 'Expecting the Unexpected: Exceptions in Grammar', Mouton de Gruyter, Berlin, Germany.

Darwin, C. (1871), *The Descent of Man and Selection in Relation to Sex*, John Murray, London.

de Boer, B. (1997), Generating Vowel Systems in a Population of Agents, *in* P. Husbands & I. Harvey, eds, 'Fourth European Conference on Artificial Life', The MIT Press, London, pp. 503–510.

de Bot, K. & Stoessel, S. (2002), Introduction: Language change and social networks, *in* O. Garcia Otheguy, ed., 'International Journal of the Sociology of Language', Vol. 2002, De gruyter, London, pp. 1–24.

De Pauw, G. (2003), Evolutionary Computing as a Tool for Grammar Development, *in* 'Genetic and Evolutionary Computation Conference Proceedings, Part 1.', Springer, New York, USA.

de Waal, F. (1996), *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*, Harvard University Press, Cambridge, USA.

de Waal, F. & Pollick, A. (2013), Gesture as the Most Flexible Modality of Primate Communication, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Deacon, T. (1997), *The Symbolic Species: The Co-evolution of Language and the Brain.*, Norton, New York, USA.

D'Errico, F. & Vanhaeran, M. (2013), Linguistic Implications of the Earliest Ornaments, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Diller, K. & Cann, R. (2013), Genetic Influences on Language Evolution: An Evaluation of the Evidence, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Donald, M. (1991), *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*, Harvard University Press, Cambridge, USA.

Doupe, A. & Kuhl, P. (1999), 'Birdsong and Human Speech: Common Themes and Mechanisms', *Annual Review of Neuroscience* **22**, 567–631.

Dunbar, R. (1993), 'The Co-evolution of Neocortical Size, Group Size and Language in Humans', *Behavioral and Brain Sciences* **16**(4), 681–735.

Dunbar, R. (1996*a*), Determinants of Group Size in Primates: A General Model, *in* W. Runciman, J. Smith & R. Dunbar, eds, 'Evolution of Social Behaviour Patterns in Primates and Man', Oxford University Press, Oxford.

Dunbar, R. (1996*b*), *Grooming, Gossip and the Evolution of Language*, Faber & Faber, London.

Dunbar, R. (2003*a*), The Origin and Subsequent Evolution of Language, *in* M. Christiansen & S. Kirby, eds, 'Language Evolution', Oxford University Press, Oxford.

Dunbar, R. (2003*b*), 'The Social Brain: Mind, Language, and Society in Evolutionary Perspective', *Annual Review of Anthropology* **32**(1), 163–181.

Elman, J. (1990), 'Finding Structure in Time', *Cognitive Science* **14**(2), 179–211.

Enard, W., Przeworski, M., Fisher, S. & Lai, C. (2002), 'Molecular Evolution of FOXP2, A Gene Involved in Speech and Language', **418**, 869–872.

Errico, F., Henshilwood, C., Lawson, G., Vanhaeren, M., Tillier, A., Soressi, M., Maureille, B., Nowell, A. & Lakarra, J. (2003), 'Archaeological Evidence for the Emergence of Language, Symbolism, and Music: An Alternative Multidisciplinary Perspective', *Journal of World Prehistory* **17**(1), 1–70.

Fauconnier, G. & Turner, M. (2008), The Origin of Language as a Product of the Evolution of Modern Cognition, *in* B. Laks, S. Cleuziou, J. Demoule & P. Encreve, eds, 'Origin and Evolution of Languages: Approaches, Models, Paradigms', Equinox, London.

Feher, O., Wang, H., Saar, P., Mitra, P. & Tchernichovski, O. (2009), 'De Novo Establishment of Wild-type Song Culture in the Zebra Finch', *Nature* **459**(7246), 564–568.

Feldman, M. & Laland, K. (1996), 'Gene-culture Coevolutionary Theory', *Trends in Ecology and Evolution* **11**(11), 453–457.

Fisher, S. & Scharff, C. (2009), 'FOXP2 as a Molecular Window into Speech and Language', *Trends in Genetics* **25**(4), 166–177.

Fitch, W. (2000), 'The Evolution of Speech: A Comparative Review', *Trends in Cognitive Science* **4**(7), 258–267.

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

Franke, M. (2015), 'The Evolution of Compositionality and Proto-syntax in Signaling Games', *Journal of Logic, Language and Information* **25**, 355–377.

Gardner, H. (1985), *The Mind's New Science: A History of the Cognitive Revolution.*, Basic Books., New York, USA.

Gibson, K. (2002), 'Evolution of Human Intelligence: The Roles of Brain Size and Mental Construction', *Brain, Behavior and Evolution* **59**(1), 10–20.

Gibson, K. (2013), Are Other Animals as Smart as Great Apes? Do Others Provide Better Models for the Evolution of Speech or Language?, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Ginzburg, J. & Macura, Z. (2007), Lexical Acquisition with and without Metacommuniation, *in* C. Lyon, C. Nehaniv & A. Cangelosi, eds, 'Emergence of Communication and Language', Springer-Verlag, London.

Gong, T., Shuai, L. & Zhang, M. (2014), 'Modelling Language Evolution: Examples and Predictions', *Physics Of Life Reviews* **11**(2), 280–302.

Gould, S. & Lewontin, R. (1979), 'The Sprandels of San Marco and the Panglossian Program: A Critique of the Adaptionist Programme', *Proceedings of the Royal Society of London* **205**(1161), 281–288.

Griffiths, T. (2007), 'Language Evolution by Iterated Learning with Bayesian Agents', *Cognitive Science* **31**(3), 441–480.

Grifoni, P., D' Ulizia, A. & Ferri, F. (2015), 'Computational Methods and Grammars in Language Evolution: A Survey', *Artificial Intelligence Review* **45**(369), 1–35.

Hall, R. (1966), *Pidgin and Creole Languages*, Cornell University Press, London.

Hansen, E. (1976), 'Selective Responding by Recently Separated Juvenile Rhesus Monkeys to the Calls to their Mothers', *Developmental Psychology* **9**(1), 83–88.

Hardus, M., Lameira, A., Singleton, I., Morrogh-Bernard, H., Knott, C., Ancrenaz, M. & Wich, S. (2009), A Description of the Orangutan's Vocal and Sound Repertoire, with a Focus on Geographical Variation, *in* S. Wich, S. Utami Atmoko, T. Mitra Setia & C. van Schaik, eds, 'Orangutans: Geographical Variation in Behavioural Ecology and Conservation', Oxford University Press, Oxford.

Hardus, M., Lameira, A., Van Schaik, C. & Wich, S. (2009), 'Tool Use in Wild Orangutans Modifies Sound Production: A Functionally Deceptive Innovation?', *Proceedings of the Royal Society of London B2* **276**(1673), 3689–3694.

Hare, M. & Elman, J. (1995), 'Learning and Morphological Change', *Cognition1* **56**(1), 61–98.

Harms, W. (2004), Primitive Content, Translation, and the Emergence of Meaning in Animal Communication, *in* K. O. & U. Griebel, eds, 'Evolution of Communication Systems: A Comparative Approach', The MIT Press, London.

Hartshorne, C. (1973), *Born to Sing*, Indiana University Press, Bloomington, USA.

Hashimoto, T. (1997), Usage-based Structuralization of Relationships Between Words, *in* P. Husbands & I. Harvey, eds, 'Fourth European Conference on Artificial Life', The MIT Press, Cambridge, USA.

Hashimoto, T. & Ikegami, T. (1996), 'Emergence of Net-grammar in Communicating Agents', *BioSystems* **38**(1), 1–14.

Hauser, M. (1996), *The Evolution of Communication*, The MIT Press, London.

Herbinger, I., Papworth, S., Boesch, C. & Zuberbuhler, K. (2009), 'Vocal, Gestural and Locomotor Responses of Wild Chimpanzees to Familiar and Unfamiliar Intruders: A Playback Study', *Animal Behaviour* **78**(6), 1389–1396.

Herman, L. (2002), Vocal, Social, and Self-imitation by Bottlenosed Dolphins, *in* K. Dautenhahn & C. Nehaniv, eds, 'Imitation in Animals and Artifacts: Complex Adaptive Systems', MIT Press, Cambridge, USA.

Herman, L., Abichandani, S., Elhaji, A., Herman, E., Sanchez, J. & Pack, A. (1999), 'Dolphins (Tursiop trucatus) Comprehend the Referential Character of the Human Pointing Gesture', *Journal of Comparative Psychology* **113**(4), 347–364.

Herman, L. & Forestell, P. (1985), 'Reporting Presence or Absence of Named Objects by a Language-trained Dolphin', *Neuroscience and Biobehavioural Reviews* **9**(4), 667–681.

Herman, L., Pack, A. & Wood, A. (1994), 'Bottlenose Dolphins can Generalise Rules and Develop Abstract Concepts', *Marine Mammal Science* **10**(1), 70–80.

Hinnenkamp, V. (1982), *Foreigner Talk und Tarzanisch*, Helmut Buske Verlag, Hamburg.

Hinnenkamp, V. (1984), Eye-witnessing Pidginization? Structural and Sociolinguistic Aspects of German and Turkish Foreigner Talk, *in* M. Sebba & L. Todd, eds, 'Papers from the York Creole Conference, September 24-27 1983', 11 edn, York Papers in Linguistics, York.

Hockett, C. (1960), 'The Origin of Speech', *Scientific American* **203**, 88–96.

Holm, J. (2000), *An Introduction to Pidgins and Creoles*, Cambridge University Press, Cambridge.

Hopkins, W., Taglialatela, J. & Leavens, D. (2007), 'Chimpanzees Differentially Produce Novel Vocalisations to Capture the Attention of a Human', *Animal Behaviour* **73**(2), 281–286.

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

Hurford, J. (2000), Social Transmission Favours Linguistic Generalisation, *in* 'The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form', Cambridge University Press., Cambridge.

Hurford, J. (2007), *Origins of Meaning: Language in the Light of Evolution*, Oxford University Press, Oxford.

Hurford, J. (2014), What is Wrong, and What is Right, about Current Theories of Language, in the Light of Evolution?, *in* F. Ferretti & I. Adornetti, eds, 'Humana Mente: Journal of Philosophical Studies', Vol. 27, pp. 123–133.

Hurford, J. R. (2002), Expression/Indction Models of Languae Evolution: Dimensions and Issues, *in* T. Briscoe, ed., 'Linguistic Evolution Through Language Acquisition: Formal and Computational Models', Cambridge University Press, Cambridge.

Huttegger, S. (2007), 'Evolutionary Explanations of Meaning', *Philosophy of Science* **74**(1), 1–27.

International Human Genome Sequencing Consortium (2004), 'Finishing the Euchromatic Sequence of the Human Genome', *Nature* **431**, 931–945.

Jackendoff, R. (2002), *Foundations of Language: Brain, Meaning, Grammar, Evolution*, Oxford University Press, Oxford.

Jaeger, H., Baronchelli, A., Briscoe, E., Christiansen, M., Griffiths, T., Jager, G., Kirby, S., Komarova, N., Richerson, P., Steels, L. & Triesch, J. (2009), What Can Mathematical, Computational and Robotic Models Tell Us About the Origins of Syntax?, *in* D. Bickerton & E. Szathmary, eds, 'Biological Foundations and Origins of Syntax', MIT Press, Cambridge, USA.

Jager, G. (2007), 'Evolutionary Game Theory and Typology: A Case Study', *Language* **83**(1), 74–109.

Janik, V. (2009), 'Acoustic Communication in Delphinids', *Advances in the Study of Behavior* **40**, 123–157.

Janik, V. (2013), Vocal Communication and Cognition in Cetaceans, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Janik, V. & Slater, P. (1997), 'Vocal Learning in Mammals', *Advances in the Study of Behavior* **26**, 59–99.

Jenkins, L. (2000), *Biolinguistics: Exploring the Biology of Language*, Cambridge University Press, Cambridge.

Johansson, C. (1997), *A View From Language: Growth of Language in Individuals and Populations*, Lund University Press, Lund.

Juergen, E. & Pizka, M. (2006), 'The Language Evolver Lever: Tool Demonstration', *Electronic Notes in Theoretical Computer Science* .

Kako, E. (1999), 'Elements of Syntax in the Systems of Three Language-trained Animals', *Animal Learning and Behaviour* **27**(1), 1–14.

Kaplan, F., Oudeyer, P. & Bergen, B. (2008), 'Computational Models in the Debate Over Language Learnability', *Infant and Child Development* **17**(1), 55–80.

Karov, Y. & Edelman, S. (1996), 'Similarity-based Word Sense Disambiguation', *Technical Report of Weizmann Institute* (96-06).

Ke, J., Hong, T. & Wang, W. (2008), 'Language Change and Social Networks', *Communications in Computational Physics* **3**(4), 935–949.

Kirby, S. (2002*a*), Learning, Bottlenecks and the Evolution of Recursive Syntax, *in* T. Briscoe, ed., 'Linguistic Evolution Through Language Acquisition: Formal and Computational Models', Cambridge University Press., Cambridge.

Kirby, S. (2002*b*), 'Natural Language from Artificial Life.', *Artificial life* **8**(2), 185–215.

Kirby, S. (2002*c*), Syntax Without Natural Selection: How Compositionality Emerges From Vocabulary In A Population Of Learners, *in* C. Knight, J. Hurford & M. Studdert-Kennedy, eds, 'The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form', Cambridge University Press, Cambridge.

Kirby, S. (2007), The Evolution of Meaning-Space Structure Through Iterated Learning, *in* C. Lyon, C. Nehaniv & A. Cangelosi, eds, 'Emergence of Communication and Language', Springer-Verlag, London.

Kirby, S. (2012), Language is an Adaptive System: The Role of Cultural Evolution in the Origins of Language, *in* 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Kirby, S. (2013), Language, Culture, and Computation: An Adaptive Systems Approach to Biolinguistics, *in* C. Boeckx, ed., 'The Cambridge Handbook of Biolinguistics', Cambridge University Press, Cambridge.

Kirby, S., Cornish, H. & Smith, K. (2008), 'Cumulative Cultural Evolution in the Laboratory: An Experimental Approach to the Origins of Structure in Human Language.', *Proceedings of the National Academy of Sciences of the United States of America* **105**(31), 10681–6.

Kirby, S., Griffiths, T. & Smith, K. (2014), 'Iterated Learning and the Evolution of Language', *Current Opinion in Neurobiology* **28**(C), 108–114.

Kirby, S. & Hurford, J. (2002), The Emergence of Linguistic Structure an Overview of the Iterated Learning Model, *in* A. Cangelosi & D. Parisi, eds, 'Simulating the Evolution of Language', Springer, London.

Knight, C. & Power, C. (2013), Social Conditions for the Evolutionary Emergence of Language, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Kohler, W. (1927), *The Mentality of Apes*, Routledge & Kegan Paul, London.

Konopka, G., Bomar, J. & Winden, K. (2009), 'Human-specific Transcriptional Regulation of CNS Development Genes by FOXP2', *Nature* **462**(7270), 213–218.

Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R., Burbano, H., Hublin, J., Hanni, C., Fortea, J., de la Rasilla, M., Bertranpetit, J., Rosas, A. & Paabo, S. (2007), 'The Derived FOXP2 Variant of Modern Humans was Shared by Neanderthals', *Current Biology* **17**(21), 1908–1912.

Kudo, H. & Dunbar, R. (2001), ‘Neocortex Size and Social Network Size in Primates’, *Animal Behaviour* **62**(4), 711–722.

Labov, W. (2001), *Principles of Linguistic Change (II): Social Factors*, Blackwell, Oxford.

Lai, C., Fisher, S., Hurst, J., Vargha-Khadem, F. & Monaco, A. (2001), ‘A Forkhead-domain Gene is Mutated in Severe Speech and Language Disorder’, *Nature* **413**, 519–523.

Laporte, M. & Zuberbuhler, K. (2010), ‘Vocal Greeting Behaviour in Wild Chimpanzee Females’, *Animal Behaviour* **80**(3), 467–473.

Lewis, D. (1969), *Convention: A Philosophical Study*, Harvard University Press, Harvard, USA.

Lieberman, P. (1968), ‘Primate Vocalizations and Human Linguistic Ability’, *Journal of the Acoustical Society of America* **44**(6), 1574–1584.

Lieberman, P. (1984), *The Biology and Evolution of Language*, Harvard University Press., London.

Lieberman, P. (2006), *Toward an Evolutionary Biology of Language*, Harvard University Press, Cambridge, USA.

Lilijencrants, L. & Lindbolm, B. (1972), ‘Numerical Simulations of Vowel Quality Systems: The Role of Perceptual Contrast’, *Language* **48**(4), 89–862.

Lupyan, G. & Dale, R. (2010), ‘Language Structure is Partly Determined by Social Structure.’, *PloS one* **5**(1), e8559.

MacLarnon, A. (2013), The Anatomical and Physiological Basis of Human Speech Production: Adaptations and Exaptations, *in* M. Tallerman & K. Gibson, eds, ‘The Oxford Handbook of Language Evolution’, Oxford University Press, Oxford.

MacLennan, B. J. & Burghardt, M. (1994), ‘Synthetic Ethology and the Evolution of Cooperative Communication’, *Adaptive Behaviour* **2**, 151–188.

MacNeilage, P. (2013), The Evolution of Phonology, *in* M. Tallerman & K. Gibson, eds, ‘The Oxford Handbook of Language Evolution’, Oxford University Press, Oxford.

Marler, P. (1957), ‘Specific Distictiveness in the Communication Signals of Birds’, *Behaviour* **11**(1), 13–19.

Marler, P. (1970), ‘Birdsong and Speech Development: Could there be Parallels?’, *American Scientist* **58**(6), 669–673.

Marshack, A. (1991), *The Roots of Civilization*, Moyer Bell, New York, USA.

Marshall, A., Wrangham, R. & Arcadi, A. (1999), 'Does Learning Affect the Structure of Vocalizations in Chimpanzees', *Animal Behaviour* **58**(4), 825–830.

Martinez, I., Quam, R., Rosa, M., Jarabo, P., Lorenzo, C. & Arsuaga, J. (2008), 'Auditory Capacities of Human Fossils: A New Approach to the Origin of Speech', *Proceeedings of the Conference Acoustics 08* pp. 4177–4182.

McComb, K. & Semple, S. (2005), 'Co-evolution of Vocal Communication and Sociality in Primates', *Biology Letters* **1**(4), 381–385.

McGrew, W. (2004), *The Cultural Chimpanzee: Reflections on Cultural Primatology*, Cambridge University Press, Cambridge.

Meguerditchian, A., Cochet, H. & Vauclair, J. (2011), From Gesture to Language: Ontogenetic and Phylogenetic Perspectives on Gestural Communication and its Cerebral Lateralization, *in* K. Dautenhahn & A. Cangelosi, eds, 'Primate Communication and Human Language: Vocalisation, Gestures, Imitiation, and Deixis in Humans and Non-humans', John Benjamins Publishing Company, Amsterdam, The Netherlands.

Mesthrie, R. (1989), 'The Origins of Fanakalo', *Journal of Pidgin and Creole Languages* **4**(2), 211–240.

Milroy, L. (1987), *Language and Social Networks*, Blackwell, Oxford.

Milroy, L. (2013), Social Networks, *in* J. Chambers & N. Schilling, eds, 'The Handbook Of Language Variation and Change', Wiley-Blackwell, Oxford.

Mitchener, W. (2007), 'Game Dynamics with Learning and Evolution of Universal Grammar', *Bulletin Mathematical Biology* **69**(3), 1093–1118.

Mithen, S. (2000), Mind, Brain, and Material Culture: An Archaeological Perspective, *in* P. Carruthers & A. Chamberlin, eds, 'Evolution and the Human Mind: Modularity, Langugae, and Meta-cognition', Cambridge University Press, Cambridge.

Mufwene, S. S. (2008), What do Creoles and Pidgins Tell Us About the Evolution of Language?, *in* B. Laks, ed., 'Origin and Evolution of Languages: Approaches, Models, Paradigms', Equinox, London.

Muhlhausler, P. (1986), *Pidgin and Creole Linguistics*, Blackwell, Oxford.

Negus, V. (1949), *The Comparative Anatomy and Physiology of the Larynx*, Hafner, New York.

Nettle, D. (1999*a*), 'Is the Rate of Linguistic Change Constant?', *Lingua* **108**, 119–136.

Nettle, D. (1999*b*), 'Linguistic Diversity of the Americas can be Reconciled with a Recent Colonization', *Proceedings of the National Academy of Sciences of the United States of America* **96**(6), 3325–3329.

Nettle, D. (1999*c*), 'Using Social Impact Theory to Simulate Language Change', *Lingua1* **108**, 95–117.

Newman, M. (2010), *Networks: An Introduction*, Oxford University Press, Oxford.

Niyogi, P. & Berwick, R. (1997), 'A Dynamical Systems Model for Language Change', *Complex Systems* **11**, 161–204.

Nowak, M. (2000), 'Evolutionary Biology Of Language', *Philosophical Transactions: Biological Sciences* **335**(1403), 1615–1622.

Nowak, M., Komarova, N. & Niyogi, P. (2001), 'Evolution of Universal Grammar', *Science* **291**(5501), 114–118.

Okanoya, K. (2013), Birdsong for Biolinguistics, *in* 'The Cambridge Handbook of Biolinguistics', Cambridge University Press, Cambridge.

Oliphant, M. (2002), Rethinking the Language Bottleneck: Why don't Animals Learn to Communicate?, *in* K. Dautenhahn & C. Nehaniv, eds, 'Imitation in Animals and Artifacts', The MIT Press, London.

Oliphant, M. & Batali, J. (1997), 'Learning and the Emergence of Coordinated Communication', *The Newsletter of the Center for Research in Language* **11**(1), 1–46.

Ortega, A., de la Cruz, M. & Alfonseca, M. (2007), 'Christiansen Grammar Evolution: Grammatical Evolution with Semantics', *IEEE Transactions on Evolutionary Computation* **11**(1).

Payne, K. (2003), Sources of Social Complexity in Three Elephant Species, *in* F. de Waal & P. Tyack, eds, 'Animal Social Complexity: Intelligence, Culture, and Individualized Socieites', Harvard University Press, London.

Pepperberg, I. (2013), Evolution of Communication and Language: Insights from Parrots and Songbirds, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Perfors, A. (2002), 'Simulated Evolution of Language : A Review of the Field', *Journal of Artificial Societies and Social Simulation* **5**(2), online.

Pinker, S. (1994), *The Language Instinct*, Penguin, London.

Pinker, S. (2003), Language as an Adaptation to the Cognitive Niche, *in* M. Christiansen & S. Kirby, eds, 'Language Evolution', Oxford University Press, Oxford.

Pinker, S. & Bloom, P. (1990), 'Natural Language and Natural Selection', *Behavioral and Brain Sciences* **13**(4), 707–784.

Plotnik, J., de Waal, F. & Reiss, D. (2006), ‘Self-recognition in an Asian Elephant.’, *Proceedings of the National Academy of Sciences of the United States of America* **103**(45), 17053–7.

Povinelli, D., Bering, J. & Giambrone, S. (2000), ‘Toward a Science of Other Minds: Escaping the Argument by Analogy’, *Cognitive Science* **24**(3), 509–542.

Reader, S. & Laland, K. (2002), ‘Social Intelligence, Innovation, and Enhanced Brain Size in Primates.’, *Proceedings of the National Academy of Sciences of the United States of America* **99**(7), 4436–41.

Reiss, D. & Marino, L. (2001), ‘Mirror Self-recognition in the Bottlenose Dolphin: A Case of Cognitive Convergence.’, *Proceedings of the National Academy of Sciences of the United States of America* **98**(10), 5937–42.

Rendell, L., Fogarty, L., Hoppitt, W., Morgan, T., Webster, M. & Laland, K. (2011), ‘Cognitive Culture: Theoretical and Empirical Insights into Social Learning Strategies’, *Trends in Cognitive Science* **15**(2), 68–76.

Rendell, L. & Whitehead, H. (2001), ‘Culture in Whales and Dolphins’, *The Behavioral and Brain Sciences* **24**(2), 309–24.

Rizzolatti, G. & Arbib, M. (1998), ‘Language within our Grasp.’, *Trends in Neurosciences* **21**(5), 188–94.

Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. (1995), ‘Premotor Cortex and the Recognition of Motor Actions’, *Cognitive Brain Research* **3**(2), 131–141.

Roberge, P. T. (2009), The Creation of Pidgins as a Possible Window on Language Evolution, *in* R. Botha & H. de Swart, eds, ‘Language Evolution: The View from Restricted Linguistic Systems’, LOT - Netherlands Graduate School of Linguistics, The Netherlands.

Roberge, P. T. (2013), Pidgins, Creoles and the Creation of Language, *in* M. Tallerman & K. Gibson, eds, ‘The Oxford Handbook of Language Evolution’, Oxford University Press, Oxford.

Roth, A. & Erev, I. (1995), ‘Learning in Extensive form Games: Experimental Data and Simple Dynamical Models in the Intermediate Term’, *Games and Economic Behaviour* **8**(1), 164–212.

Rumelhart, D., Hinton, G. & Williams, R. (1986), ‘Learning Representations by Backpropogation Errors’, *Nature* **323**(6088), 533–536.

Sadedin, S. & Paperin, G. (2011), ‘Implications of the Social Brain Hypothesis for Evolving Human-like Cognition in Digital Organisms’, *Advances in Artificial Life. Darwin Meets von Neumann* **103**(43), 61–68.

Sawaguchi, T. (1992), ‘The Size of the Neocortex in Relation to Ecology and Social Struture in Monkeys and Apes’, *Folia Primatologica* **58**(3), 130–145.

Scharff, C. & Haesler, S. (2005), ‘An Evolutionary Perspective on FOXP2: Strictly for the Birds’, *Current Opinion in Neurobiology* **15**(6), 694–703.

Sebba, M. (1997), *Contact Languages: Pidgins and Creoles*, Macmillan Press LTD., London.

Senghas, R., Senghas, A. & Pyers, J. (2005), The Emergence of Nicaraguan Sign Language: Questions of Development, Acquistion, and Evolution, *in* S. Parker, J. Langer & C. Milbrath, eds, ‘Biology and Knowledge Revisited: From Neurogenesis to Psychogenesis’, Lawrence Erlbaum Associates, London.

Seyfarth, R. (1987), Vocal Communication and its Relation to Language, *in* D. Smuts, D. Cheney, R. Seyfarth, R. Wrangham & T. Struthsaker, eds, ‘Primate Societies’, Chicago, USA, pp. 440–451.

Seyfarth, R. & Cheney, D. (1984), ‘The Acoustic Features of Vervet Monkey Grunts’, *Journal of the Acoustical Society of America* **75**(5), 129–134.

Seyfarth, R. & Cheney, D. (2013), Primate Social Cognition as a Precursor to Language, *in* M. Tallerman & K. Gibson, eds, ‘The Oxford Handbook of Language Evolution’, Oxford University Press, Oxford.

Skyrms, B. (2004), *The Stag Hunt and the Evolution of Social Structure*, Cambridge University Press., Cambridge.

Skyrms, B. (2009), ‘Evolution of Signalling Systems with Multiple Senders and Receivers’, *Philosophical Transactions of the Royal Society of London. Series B* **364**(1518), 771–779.

Skyrms, B. (2010), *Signals: Evolution, Learning, and Information*, Oxford University Press, Oxford.

Slater, P. (2013), Bird Song and Language, *in* M. Tallerman & K. Gibson, eds, ‘The Oxford Handbook of Language Evolution’, Oxford University Press, Oxford.

Slocombe, K. (2013), Have we Underestimated Great Ape Vocal Capacities?, *in* M. Tallerman & K. Gibson, eds, ‘The Oxford Handbook of Language Evolution’, Oxford University Press, Oxford.

Slocombe, K. & Zuberbuhler, K. (2007), ‘Chimpanzees Modify Recruitment Screams as a Function of Audience Composition’, *Proceedings of the National Academy of Sciences of the United States of America* **104**(43), 441–449.

Smith, A. (2007), Language Change and the Inference of Meaning, *in* C. Lyon, C. Nehaniv & A. Cangelosi, eds, 'Emergence of Communication and Language', Springer-Verlag, London.

Smith, H., Newman, J. & Symmes, D. (1982), Vocal Concomitants of Affiliative Behaviour in Squirrel Monkeys, *in* C. Snowden, C. Brown & M. Petersen, eds, 'Primate Communication', Cambridge University Press, New York, USA.

Smith, K. (2002), 'Natural Selection and Cultural Selection in the Evolution of Communication', *Adaptive Behavior* **10**(1), 25–45.

Smith, K. (2003), The Transmission of Language: Models of Biological and Cultural Evolution, PhD thesis, University of Edinburgh.

Smith, K. (2009), Iterated Learning in Populations of Bayesian Agents, *in* N. Taatgen & H. van Rijn, eds, 'Proceedings of the 31st Annual Conference of the Cognitive Science Society', Cognitive Science Society, Austin, USA, pp. 697–702.

Smith, K., Kirby, S. & Brighton, H. (2003), 'Iterated Learning: a Framework for the Emergence of Language', *Artificial Life* **9**(4), 371–86.

Steels, L. (1995), 'A Self-organizing Spatial Vocabulary', *Artificial Life* **2**(3), 319–332.

Steels, L. (2007), The Recruitment Theory of Language Origins, *in* C. Lyon, C. Nehaniv & A. Cangelosi, eds, 'Emergence of Communication and Language', Springer-Verlag, London.

Steels, L. (2011*a*), A First Encounter with Fluid Construction Grammar, *in* L. Steels, ed., 'Design Patterns in Fluid Construction Grammar', John Benjamins Publishing Company, Amsterdam, The Netherlands.

Steels, L. (2011*b*), Introducing Fluid Construction Grammar, *in* L. Steels, ed., 'Design Patterns in Fluid Construction Grammar', John Benjamins Publishing Company, Amsterdam, The Netherlands.

Steels, L. & Garcia-Casademont, E. (2015), How to Play the Syntax Game, *in* P. Andrews, L. Caves, S. Doursat, S. Hickinbotham, F. Polack, S. Stepney, T. Taylor & J. Timmis, eds, 'Advances in Artificial Life: Proceedings of the Thirteenth European Conference of Artifici al Life (ECAL 2015)', The MIT Press, Mountain View, USA, pp. 479–486.

Steiper, M. & Young, N. (2006), 'Primate Molecular Divergence Dates', *Molecular Phylogenetics and Evolution* **41**(2), 384–482.

Stephan, C. & Zuberbuhler, K. (2008), 'Predation Increases Acoustic Complexity in Primate Alarm Calls', *Biology Letters* **4**(6), 641–644.

Striedter, G. (2005), *Principles of Brain Evolution*, Sinauer Associates, Sunderland, USA.

Studdert-Kennedy, M. (2013), The Emergence of Phonetic Form, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Tallerman, M. (2013), Protolanguage, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, oxford.

Tallerman, M. & Gibson, K. (2013*a*), Introduction: The Evolution of Language, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Tallerman, M. & Gibson, K. (2013*b*), Language Change, Creation, and Transmission in Modern Humans, *in* M. Tallerman & K. Gibson, eds, 'The Oxford Handbook of Language Evolution', Oxford University Press, Oxford.

Tayler, C. & Saayman, G. (1973), 'Imitative Behaviour by Indian Ocean Bottlenose Dolphins (Tursiops aduncus) in Captivity', *Behaviour* **44**(3/4), 286–297.

Tenaza, R. (1976), 'Songs and Related Behaviour of Kloss' Gibbons (Hylobates Klossii) in Siberut Island, Indonesia', *Zeitschrift für Tierpsychologie* **40**(1), 37–52.

Todd, L. (1990), *Pidgins and Creoles*, Routledge, London.

Tomasello, M. & Call, J. (1997), *Primate Cognition*, Oxford University Press, Oxford.

Townsend, S., Deschner, T. & Zuberbuhler, K. (2008), 'Female Chimpanzees use Copulation Calls Flexibly to Prevent Social Competition', *PloS ONE2* **3**, e2431.

Trudgill, P. (2013), Linguistic and Social Typology, *in* J. Chambers & N. Schilling, eds, 'The Handbook of Language Variation and Change', Wiley-Blackwell, Oxford.

Turner, H. (2002), An Introduction to Methods for Simulating the Evolution of Language, *in* A. Cangelosi & D. Parisi, eds, 'Simulating the Evolution of Language', Springer-Verlag, London.

Tyack, P. (2003), Dolphins Communicate About Individual-Specific Social Relationships, *in* F. de Waal & P. Tyack, eds, 'Animal Social Complexity: Intelligence, Culture, and Individualized Socieites', Harvard University Press, London.

Vogt, P. (2009), 'Modeling Interactions Between Language Evolution and Demography', *Human Biology* **81**(2), 237–258.

Werner, G. & Dyer, M. (1992), Evolution of Communication in Artificial Organisms, *in* C. Langton, C. Taylor, J. Farmer & S. Rasmussen, eds, 'Artificial Life II', Addison-Wesley, Redwood City, USA, pp. 659–687.

Whinnom, K. (1968), Linguistic Hybridization and the 'Special Case' of Pidgins and Creoles, *in* D. Hymes, ed., 'Pidginization and Creolization of Languages', Cambridge University Press, Cambridge.

Whitehead, H. (2003), Society and Culture in the Deep and Open Ocean: the Sperm Whale and Other Cetaceans, *in* F. de Waal & P. Tyack, eds, 'Animal Social Complexity: Intelligence, Culture, and Individualized Socieites', Harvard University Press, London.

Whiten, A. (2000), 'Primate Culture and Social Learning', *Cognitive Science* **24**(3), 477–508.

Whiten, A. & van Schaik, C. (2007), 'The Evolution of Animal 'Cultures' and Social Intelligence.', *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* **362**(1480), 603–20.

Wich, S., Swartz, K., Hardus, M., Lameira, A., Stromberg, E. & Shumaker, R. (2009), 'A Case of Spontaneous Acquisition of a Human Sound by an Orangutan', *Primates* **50**(1), 56–64.

Wichmann, S. & Holman, E. (2009), 'Population Size and Rates of Language Change', *Human Biology* **81**(2), 259–274.

Wichmann, S., Stauffer, D., Schulze, C. & Holman, E. (2008), 'Do Language Change Rates Depend on Population Size?', *Advances in Complex Systems* **11**(3), 259–274.

Wilkins, W. & Wakefield, J. (1995), 'Brains Evolution and Neurolinguistic Preconditions', *Behavioral and Brain Sciences* **18**(11), 161–226.

Wolpoff, M., Mannheim, B., Mann, A., Hawks, J., Caspari, R., Rosenberg, K., Frayer, D., Gill, G. & Clark, G. (2004), 'Why Not the Neandertals?', *World Archaeology* **36**(4), 527–546.

Worden, R. (1998), The Evolution of Language from Social Intelligence, *in* J. Hurford, M. Studdert-Kennedy & C. Knight, eds, 'Approaches to the Evolution of Language: Social and Cognitive Bases', Cambridge University Press, Cambridge.

Zuberbuhler, K. (2000), 'Interspecies Semantic Communication in Two Forest Primates.', *Proceedings from the Royal Society B: Biological Sciences* **267**(1444), 713–718.

Zuberbuhler, K. (2002), 'A Syntactic Rule in Forest Monkey Communication', *Animal Behaviour* **63**(2), 293–299.

Zuberbuhler, K. (2003), Natural Semanticity in Wild Primates, *in* F. de Waal & P. Tyack, eds, 'Animal Social Complexity: Intelligence, Culture, and Individualized Socieites', Harvard University Press, London.

Zuberbuhler, K. (2012), ‘Primate Communication’, *Nature Education Knowledge* **3**(10), 83–93.

Zuberbuhler, K. (2013), The Cognitive Capacities of Non-Human Primates, *in* C. Boeckx, ed., ‘The Cambridge Handbook of Biolinguistics’, Cambridge University Press, Cambridge.