

Social Networks and Cultural Transmission

Justin Quillinan

**Linguistics and English Language
School of Philosophy, Psychology and Language Sciences
University of Edinburgh**

August 2006

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS OF THE DEGREE OF
MASTER OF SCIENCE

Declaration

I hereby declare that this thesis is of my own composition, and that it contains no material previously submitted for the award of any other degree. The work reported in this thesis has been executed by myself, except where due acknowledgement is made in the text.

Justin Quillinan

Abstract

Language is a complex dynamical system that is shaped not just through biological evolution but by the way it is used in a social context. Sociolinguists have long understood that the structure of a society strongly affects the nature of the languages that emerge. Computational models of language evolution, however, generally neglect the effect of social structure by modelling extremely simple population dynamics. This study explores the coevolution of language and social structure using a simple, abstract model of language learning and a plausible mechanism for network growth, namely homophily. Evolved networks are found to possess the characteristic measures of social networks: assortative mixing, transitivity and prominent community structure. The effect of embedding language-learners in the network is found to be significant. This model may also provide a platform on which existing theories and computational models of language evolution can be evaluated.

Acknowledgements

Thanks are due to my supervisors Simon Kirby and Kenny Smith for their advice and guidance. I would also like to thank Mary, Jim and Matthew Quillinan, for their support throughout the course. Finally I would like to thank the residents of the computer lab where the majority of this dissertation was carried out; in particular Karola Kreitmair, for her always informative and entertaining discussions, and Cyprian Laskowski for his technical assistance and frisbee.

Contents

Declaration	i
Abstract	ii
Acknowledgements	iii
Chapter 1 Introduction	1
Chapter 2 Literature Review	3
2.1 Introduction	3
2.1.1 Network theory	3
2.2 Language Change and Cultural Transmission	4
2.2.1 Introduction	4
2.2.2 Models of language/cultural transmission	5
2.2.3 Discussion	11
2.3 Language change in society	12
2.3.1 Introduction	12
2.3.2 Homophily	13
2.3.3 Antagonism	14

2.3.4	Norm enforcement	14
2.3.5	Social mobility	15
2.3.6	Discussion	15
2.4	Social Networks	16
2.4.1	Introduction	16
2.4.2	Characteristic measures of social networks	17
2.4.3	Evolving social networks	18
	Evolving networks for transitivity	19
	Evolving networks for assortativity	22
	Evolving networks for community structure	23
2.4.4	Discussion	24
Chapter 3	Methodology	25
3.1	Introduction	25
3.2	Aims	25
3.3	Design choices for the model	26
3.4	The model defined	27
3.5	Measures	29
3.5.1	Introduction	29
3.5.2	Measures of social network growth	29
	Assortativity	29
	Clustering Coefficient	30
	Prominent Community Structure	31

<i>CONTENTS</i>	vii
3.5.3 Language entropy	31
3.5.4 Social Difference vs. Language Distance	31
3.6 Testing the model	32
3.6.1 The Null model - the random network	32
3.6.2 Test 1: The effect of preferential attachment by trait similarity	33
3.6.3 Test 2: The effect of preferential attachment by degree sim- ilarity	33
3.6.4 Test 3: The effect of learning	33
Chapter 4 Results	35
4.1 Introduction	35
4.2 Evolved networks	35
4.2.1 The Null model - the random network	35
4.2.2 Test 1: The effect of preferential attachment by trait similarity	36
4.2.3 Test 2: The effect of preferential attachment by degree sim- ilarity	37
4.2.4 Test 3: The effect of learning	38
4.3 Evolving networks	39
4.3.1 Clustering coefficient over time	40
4.3.2 Degree correlation over time	40
4.3.3 Language entropy	41
Chapter 5 Discussion	43
5.1 Introduction	43
5.2 Discussion of the results	43

5.3	Further work	44
5.3.1	Exploring the parameters of the model	45
5.3.2	Language evolution	45
5.3.3	Extension to the preferential attachment model	45
5.4	Conclusions	47
Appendix A Diagrams of evolving networks		49
Appendix B Statistical results		57
B.1	Tests of Between-Subjects Effects for Dependant Variable Clustering Coefficient	57
B.2	Tests of Between-Subjects Effects for Dependant Variable Degree Correlation	58
B.3	Tests of Between-Subjects Effects for Dependant Variable Language Distance Correlation	59
References		60

List of Tables

2.1	Summary of population models of language transmission.	6
2.2	Summary of computational models for growing social networks. .	19
3.1	Degree correlation coefficients r_d for a number of different networks of size n (from Newman 2002)	30
B.1	Clustering coefficient statistics	57
B.2	Degree correlation statistics	58
B.3	Language distance and social distance correlation	59

List of Figures

2.1	A small undirected, unweighted network with 8 vertices and 10 edges.	4
2.2	Network structures used in language evolution simulations	8
2.3	Social network grown through attachment of neighbouring vertices (Jin et al. 2001)	21
4.1	Random network	36
4.2	Network built with preferential attachment based on trait similarity	37
4.3	Preferential attachment by degree and trait similarity	38
4.4	The effect of learning and preferential attachment based on trait similarity	39
4.5	Clustering coefficient $C(p)$ over time	40
4.6	Degree correlation r_d over time	41
4.7	Language entropy H_e over time	42
A.1	Evolving network with degree preference	50
A.2	Evolving network with trait preference	51
A.3	Evolving network with trait and degree preference	52
A.4	Evolving network with learning	53
A.5	Evolving network with learning and degree preference	54

A.6	Evolving network with learning and trait preference	55
A.7	Evolving network with learning, trait and degree preference	56

CHAPTER 1

Introduction

Language is a complex dynamical system that is shaped not just through biological evolution but by the way it is used in a social context. It is best understood as a product of society and culture (Sapir 1929). This view has been widely held in sociolinguistics, and has recently challenged the established view in linguistics of language being primarily innately determined (Chomsky 1986, Pinker 1994). Computational models of language evolution have demonstrated that many universal features of language previously thought to be given by some kind of innate grammar can in fact be explained by the constraints on the transmission of language by populations of language learners (Kirby & Hurford 2002). These computational models are necessarily much simpler than real-world populations, but one of the areas which seems to have been most neglected is that of the population dynamic. Computational models generally model extremely simplified or unrealistic social structure, yet the structure of social networks is much more complicated, and has a marked effect on the transmission of language (Chambers 2003).

Network theory has been applied to the study of social networks, and through this large scale statistical properties of networks have been discovered that distinguish them from other natural or artificial networks. Namely: assortative mixing, transitivity and prominent community structure (Newman & Park 2003). Several authors have put forward methods for growing social networks to achieve these particular measures, but most appear to be geared toward achieving these measures rather than proving a plausible explanation of how the networks may evolve (Toivonen et al. 2006, Jin et al. 2001, Newman 2002, Freeman 1996). This study explores the coevolution of social structure and language using a simple,

abstract model of language and learning, with the goal of achieving the characteristic measures of social networks using a plausible mechanism. The model could provide the basis for more complex algorithms of language evolution over an evolving social network.

CHAPTER 2

Literature Review

2.1 Introduction

This chapter explores language change and cultural transmission from both a computational and sociolinguistic perspective. Computational models are described that are based on simplified population dynamics that neglect many of the aspects of real-world social interaction. Language, however, is used in a complex social network structure in which individual language users are embedded. The topology of the network structure affects the nature of the language that emerges. The chapter concludes with a review of the computational models of evolving social networks, which may be used to model the coevolution of language and social structure.

First, some terminology from the field of network theory that will be used throughout this document.

2.1.1 *Network theory*

Newman (2003) gives an overview of the definitions and methods for analysing complex networks. A network consists of a set of *vertices* (singular: *vertex*), denoted by v_i , where the subscript is the unique number of the vertex; and a set of *edges* connecting them, denoted by a vertex pair (v_i, v_j) . In the mathematical literature, networks are also known as *graphs*; vertices as *nodes*; and edges as *connections*. In the sociolinguistic literature networks are referred to as *sociograms*; vertices as *points*; and edges as *paths*. For consistency, this document will favour

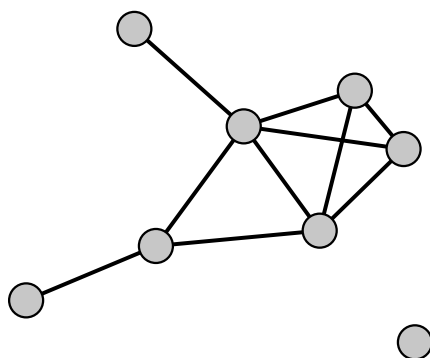


Figure 2.1: A small undirected, unweighted network with 8 vertices and 10 edges.

the terminology of networks, vertices and edges. Figure 2.1 shows a small network of 8 vertices and 10 edges. Vertices are represented by circles, and edges by the lines between them. Networks may be more complicated than this, however. Vertices and edges may be of different types, or have different properties. In particular, edges may be directed, whereby the relationship is only in one direction, or undirected. Edges may have weights, representing the strength of the connection, such as how well two people know each other. One crucial basic network measure is the *degree* of a vertex, occasionally referred to as the connectivity of the vertex, which is the number of edges connected to that vertex. The vertices that are connected by edges to a vertex v are said to be *adjacent* to v . In an undirected unweighted network that allows only single connections, the degree of a vertex is simply the number of adjacent vertices. A *component* to which a vertex v belongs includes all the vertices that can be reached from v , by travelling along edges in the network. A *geodesic path* is the shortest path(s) that connects two vertices in the network. Importantly, networks may also evolve over time, through the addition or removal of vertices or edges, or the changing of their properties. These networks that evolve over time are known as *longitudinal*.

2.2 Language Change and Cultural Transmission

2.2.1 Introduction

The emergence of language in a population involves the interaction of three complex systems: biological evolution, learning and cultural evolution (Kirby & Hur-

ford 2002). Biological evolution, operating over *phylogenetic* timescales, involves the adaptation of the learning and processing mechanisms of language in response to selection pressures for survival and reproduction. Learning, operating over *ontogenetic* timescales, involves individuals' adaptation of the knowledge of a language in order to optimise comprehension and production. Finally, cultural evolution involves the change in languages over historical (*glossogenetic*) timescales. Traditionally, the structure of language has been explained by the existence of an innate language faculty, placing the burden of explanation for the existence and nature of the universal features of language primarily on biological evolution (Chomsky 1986, Pinker 1994). Pinker & Bloom (1990) contend that language is no different from other abilities such as echolocation or stereopsis. It follows then that such a complex ability can only be explained by orthodox theory of evolution, and the origin of same is through biological natural selection. Recently however, much emphasis has been placed on explaining the universal properties of language in terms of learning and cultural transmission of languages by language-learners. Kirby & Hurford (2002) argue that it is through the cultural transmission of learned behaviour that the most basic features of language structure can be explained. For a language or pattern within a language to persist it must be transmitted from one generation of language users to the next, and is therefore shaped by the constraints on such transmission. In this way, languages are best explained, not just as the products of biological natural selection, but as "social and cultural entities that have evolved with respect to the forces of selection imposed by human users" (Deacon 1997, p.110). Deacon argues that linguistic universals emerge independently in languages in response to the selection processes that affect the transmission of language. Thus, it is not just the language users that are subject to natural selection, but the language itself. Furthermore, the rate of linguistic change is far greater than the rate of biological change, such that the selection pressures operating on languages are on a much faster timescale than those acting on species. Language then might be seen as an organism in itself, adapted through natural selection to fit the human brain: its particular ecological niche (Christiansen 1994). Language can thus be seen as coevolving with language-learners.

2.2.2 *Models of language/cultural transmission*

Computational models have been used to explore complex dynamical systems such as language learning and transmission, where natural experiments are difficult or impossible to perform. In a complex dynamical system such as a popula-

Reference	Population model
Kirby (2002)	Vertical transmission; 1 adult, 1 child
Hurford (2000)	Vertical transmission; 4 adults, 1 child; child learns from 1 adult only
Smith & Hurford (2003)	Vertical transmission; $N = 10$; child entering population has $1 \geq p \leq 10$ cultural parents to learn from
Keller (1994)	Regular lattice network (55×55 vertices);
Nettle (1999)	Regular lattice network (7×7 vertices representing groups of 20 individuals); migration; social selection
Kirby (2000)	Ring network model; $N = 10$
Baronchelli et al. (2006)	Random network and Barabasi-Albert Network; $N = 2000$
Gong et al. (2004)	Local-world model
Minett & Wang (2004)	Local-world model

Table 2.1: Summary of population models of language transmission.

tion of language-users, it is difficult to determine the effect of any one variable, or to test particular theories relating to the interaction of many individual components. Computational simulation of simplified models of these systems provide a way in which these theories can be developed and tested. In addition, the simulation of such complex dynamical systems can reveal unpredictable emergent consequences of the behaviour of populations from interactions at an individual level. Various computational models of the evolution of language have been proposed to account for the universals of language without placing the burden of explanation purely on biological evolution. These models explore the interactions of populations of language learners, using various different population models. Table 2.1 gives a sample of various population models of language transmission, described below.

Kirby (2002) describes the Iterated Learning Model (ILM), which explores the dynamics of the mapping between the I-Language (the language internal to an individual, i.e. the grammar) and the E-Language (the language external to an individual, i.e. the set of utterances produced or comprehended) (Chomsky 1986). A language or pattern within a language persists only if it is used (I-language to E-language) and if it is learned (E-language to I-language). This model has been used to show that, given the selection pressures on the language of a finite number of utterances that a learner is exposed to - the language *bottleneck* - a structured grammar emerges. The population in this model consists of a single

adult speaker and a single learner. Much of the work on the ILM trades off complex population dynamics for speed of simulation (Kirby & Hurford 2002).

Most models of the ILM, as in the simple model just described, model transmission as exclusively from fully competent individuals to naïve learners (*vertical transmission*), and the populations generally consist of a single agent. There are however, implementations of the ILM that move away from such a simplified population dynamic. Hurford (2000) demonstrates how linguistic generalisations can emerge and be preserved in a larger population of agents. However, even in this model, a population consists of just four adult speakers and one child learner, with interaction only between an adult and a child. After a few hundred cycles of learning, the child becomes an adult, a new child enters the population and an adult is removed. Experiments were carried out with larger populations, but the main analysis was carried out on such smaller populations, to shorten the computational time. Furthermore the learners were more strongly biased than in the standard ILM. Smith & Hurford (2003) apply the ILM to a larger population of agents, with non-overlapping generational turnover. Each individual has a certain number of cultural parents p drawn from the population from which they learn. When individuals have more than one cultural parent and there is no change to the fundamental ILM learning algorithm, the variability in the input to the learner results in rapid increase in the size of the rules and a failure to construct a suitable grammar. The learning algorithm is then augmented with a bias in favour of shorter utterances. Simulations were carried out with a population size of $N = 10$, and a variable number of cultural parents ($1 \leq p \leq 10$). The change to the learning algorithm eliminated the problem. Agents were then (in most cases) able to converge on compressed expressive grammars and thus high communicative accuracy. The number of cultural parents p does affect the coverage and compression of the grammar, and the speed at which populations converge. Furthermore, even in this model the population size is only ten individuals. It is telling that even with such a small population size, the dynamics of the population had such an effect on the evolution of the language.

There have been many models exploring the evolution of language through cultural transmission, but such models as those described above do not take into account the dynamics of large scale populations. Spatial structure of a population has been used to demonstrate the emergence and maintenance of diversity. Keller (1994) reports on a model of a population arranged in a grid of 55x55 cells. This is defined as a *regular lattice* in graph theory, whereby the vertices

are positioned at each intersection and the gridlines correspond to the edges. Figure 2.2 shows the structure of a regular lattice. Each vertex in this model is initially assigned one of two linguistic variants, either A or A' . During the simulation, cells change their variant to match the majority of their neighbours. The simulation eventually stabilises with homogeneous regions of A and A' , similar to a map of isoglosses (a geographical boundary of a certain linguistic feature). Through very simple rules governing the evolution of the linguistic variants give rise to familiar sociolinguistic structures. With a large population structure, the microscopic rules and properties governing the learning algorithms and linguistic traits may prove less important. The tradeoff in this research would be between simplicity of models of learning and complexity of the population model. In this way, models of language evolution in populations may benefit from research into existing A-Life techniques, such as Conway's Game of Life model (Gardner 1970).

Nettle (1999) describes another spatial network model of language evolution,

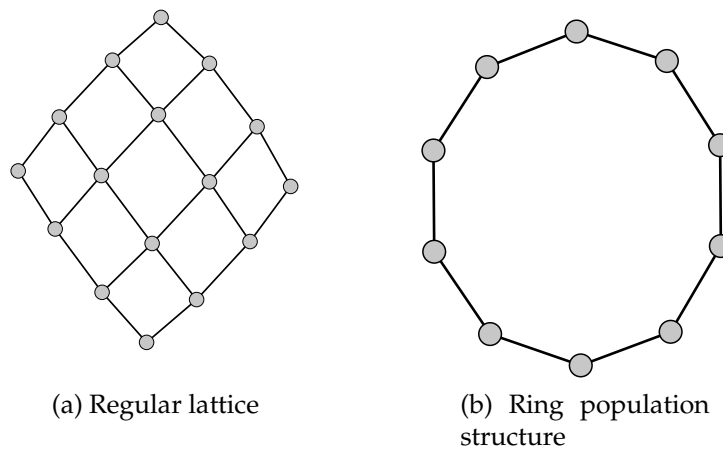


Figure 2.2: Network structures used in language evolution simulations

whereby individuals are embedded in a 7×7 regular lattice. Each intersection in this regular lattice contains 20 vertices, defined as a social group. The population is a generational turnover model, with naïve learners entering the system and enculturated adults leaving, such that the population of each social group remains constant. The linguistic system in this case consists of eight vowels, modelled as continuous numerical values. An individual learns the numerical value of their vowels by sampling all the adults in its social group and adopting the average of the sample. When individuals sample only from those in their own social group, diversity can be created (with noise in learning) and is amplified

and maintained through transmission, learning and population turnover. Nettle explores the effect of migration on the evolution of the vowel system in such a social structure. The effect of the permanent transfer of individuals from one social group to another is analysed by vertices moving with a certain probability to another position on the lattice. In this case, even low migration rates are sufficient to prevent local diversity in the social groups. A more advanced model of social selection is then explored, whereby certain adult individuals have higher status, and learners only learn from these high status individuals. This model highlights the effect of social selection on cultural transmission, in that with such social selection, diversity is significantly more robust than without.

Further to spatially organised populations (or regular lattice networks), there have been models of language evolution taking place with individuals embedded in other kinds of networks. Kirby (2000) describes an ILM model with a simple network population structure based on Oliphant (1996). The population consists of ten individuals, organised in a ring, such that each member of the population has two neighbours with which it interacts. Figure 2.2 shows the network structure of such a population. Each individual only learns from the neighbours to which it is connected. From the simulation emerges a simple, language-like syntax. At first, however, the language goes through radical and unpredictable changes before the population converges on a simple system. The initial chaotic stages are when the language is brittle. It is tempting to see this brittleness as related to the population structure in some way, as in Smith & Hurford (2003), whereby the number of cultural parents (here represented as the two adjacent vertices), may provide conflicting information. In spite of this, transmission may only percolate in two directions, limiting the amount of variability in the input. The language eventually converges, though it would be interesting to see under what conditions this would happen given a less restricted population structure.

Baronchelli et al. (2006) report on a population model of language evolution whereby agents interact using the Naming Game (Steels 2000) to negotiate conventions. The population consists of $N = 2000$ agents, and interactions take place - in the first model - with pairs of randomly chosen agents. The population is large, but in this case completely unstructured. This assumption that any agent can interact with any other agent becomes unrealistic when the population size becomes this large, however. They remedy this by embedding agents in a static

quenched spatial structure, namely a regular lattice. The other network structure they consider is that of the Barabasi-Albert Network, an artificial network that has more in common with models of Internet topology than social structure, whereby new vertices are added to the network with probability proportional to their degree (Barabási & Albert 1999a). Baronchelli et al. (2006) show that with such a network structure present, the population takes considerably longer to converge on a shared language, and the maximum number of words in the language is smaller. When agents are structured in a population, shared conventions emerge locally among clusters of agents and take longer to spread through the network. Furthermore, in such a network structure the order of the choice of speaker and hearer is important. The first choice is likely to be a vertex of low degree, since these form the majority in a scale-free network. Choosing an adjacent vertex to the initial random vertex is then likely to be a vertex of high degree due to the dissortative mixing. This has implications for cultural transmission, since an ordered choice of speakers and hearers will favour either high or low degree vertices as the inventors or transmitters of language. Thus, the convergence of the population on shared conventions is strongly influenced by the topology of the network.

The previous models of networks generally assume the network topology to be static. Gong et al. (2004) present a model of the coevolution of language and social structure. They postulate that mutual understanding based on an evolving language may be able to trigger social structure. This mutual understanding is generated by language games between agents, as a measure of whether these agents will interact in the future. The social structure is represented by a fully connected weighted network based on the local-world model of Li & Chen (2003), used to represent an individual's lack of global knowledge of the social structure or the properties of the population as a whole. In the model edge weight between two vertices is adjusted based on the degree of success of communication. Gong *et al.* show that factors such as friendship and popularity (modelled as edge weight and vertex degree respectively) have an influence on the structure of the network as well as the resulting language. However, the structure generated is based on a local-world model of evolving networks that has not been shown to possess the same distinct characteristics of social networks such as assortative mixing (though this is an effect that arises from Gong *et al.*'s model or high values of network transitivity). Minett & Wang (2004) also use the local-world model to model the effects of bilingualism on languages competing

for speakers in a population. They show that social structure has no significant effect on which language is maintained unless the population intervenes to attempt to maintain both languages. Their conclusion is that information can diffuse more rapidly across a local-world network than a scale-free network, allowing for easier maintenance of an endangered language.

2.2.3 Discussion

Many models of cultural transmission suffer from an oversimplification of the dynamics of a population. Populations are modelled as discrete, non-overlapping generations of single agents, unordered collections of agents in a single generation, or collections of agents ordered spatially or in simplified or unrealistic network structures. Often such impoverished treatment of the dynamics of a population is a useful idealisation both for computational time and testing specific theories. However, as the models themselves demonstrate, the use and change of language depends in some way on the nature and structure of the population of language-learners. Neglecting the real-world dynamics of a population may mean the theories that were developed using simplified population structure are not so robust. In these models, the analogy of the evolution of language through natural selection neglects the influence of social selection. Extending the evolution metaphor, this can be seen as playing a role similar to that of sexual selection (Nettle 1999). Individuals do not pick up all the linguistic activity that occurs around them, but instead show preference for the traits of particular target groups. Nettle describes the *amplifiers of variation* of language transmission in populations - factors that affect the transmission of variation through populations. These amplifiers take small differences between individuals and potentially enlarge them to become stable differences between populations. Factors such as functional selection (the preferential adoption of linguistic traits based on functional distinctions) and geographical isolation affect the variation in language transmission, but still neglect the internal structure of a single population. However, as is described in section 2.3 below, diversity is created and maintained even in the absence of geographical isolation. Nettle also defines *social selection* as an amplifier of variation within such populations. He quotes Lepage (1968) who defines a general model of language acquisition based on social selection, whereby individuals create systems for their verbal behaviour for the purpose of group identification. The acquisition of language by an individual is affected by the structure of the social group to which he/she belongs. Labov (1972) calls this *social embedding*, whereby language change mirrors social structure. Thus the

structure and dynamics of the social groups should be taken into account when considering the evolution of language as a whole. The field of sociolinguistics can provide clues as the problem of language evolution and social structure.

2.3 Language change in society

2.3.1 *Introduction*

Sociolinguists have long emphasised the importance of social factors in the transmission of language. Sapir (1929) stressed that “language is primarily a cultural or social product, and must be understood as such” (p. 76-77). Individuals in society do not interact randomly or uniformly, even in extremely close-knit groups. Rather, most people carry on interactions with people who are sufficiently similar to themselves. In modern society, three social characteristics have been identified as the primary determiner of social roles: class, sex and age (Chambers 2003). Of these, social class has been identified as the most important. These factors give rise to linguistic variation by, in one sense, acting as barriers to communication. Chambers (2003) reports on a study by Bogart (1951) in which the communication of a major event in a small community was sharply divided along class lines. The event was reported in the local newspaper, radio and discussed in common meeting places, circumstances which would appear to facilitate rapid diffusion of news. After interviewing townspeople several weeks after the event, Bogart found that age or sex did not prove a barrier to transmission. Among the lower classes however, knowledge of the event was significantly lower than those of the upper classes.

“Partly by choice and partly by chance, then, the social classes are not in constant or close contact. Their segregation allows differences to take root. In their speech as in other attributes, these differences may come about in the first place unwittingly, simply because one group is unaware of changes taking hold in the other group. Once established, the differences may take on status as emblems or markers of a particular class”. (Chambers 2003, p.56)

Thus, even in a small, dense community where rapid diffusion may be possible, the structure of the population has a major impact on transmission. Individuals formed into groups consisting of individuals with similar traits, in this case a particular social class. Furthermore, even within tightly structured, relatively

homogeneous communities not subject to such class differentiation, still linguistic variation persists. Micro-level clusters such as these are known as *networks* (Chambers 2003). Kerswill & Williams (2000) describes the study of a process of koineization, whereby a new mixed variety dialect is formed following social contact. They found that the adoption of features by a speaker depends on their characteristics in the network. A close knit network resists the adoption of changes unless they come from an insider who also has weak ties elsewhere. But once a change has been accepted into a close knit network, they will be accelerated due to the density of connections. The methods by which similarity and diversity are maintained in these networks are the focus of the following sections.

2.3.2 *Homophily*

The principle of *homophily* is the tendency for individuals to establish social bonds with those to whom they are similar, expressed by the adage “birds of a feather flock together”. Contact between people that are in some way similar occurs at a higher rate than among dissimilar people, such that “similarity breeds connection” (McPherson et al. 2001, p.415). These social bonds then influence the linguistic behaviour of the individuals, since it is within the group that the majority of social contact takes place. Eckert (1988) describes two groups of individuals - known by their peers as the Jocks and the Burnouts - in several high schools in Detroit. They are characterised primarily by their social activities: the former participate in sport and other school-related activities, while the latter carried on their social lives outside of the school. They thus formed separate micro cultures, which were reflected in significant positive linguistic correlations within groups and negative correlations between them. Importantly, though there was a loose correlation between the class and the group an individual belonged to, there were many crossovers, indicating that social network and social class can be independent. There was also a high proportion of In-Betweens, individuals who were affiliated with neither of the groups but whose linguistic behaviour was somewhere between, due to social contact with both groups. Whether consciously or subconsciously, individuals in each group matched their linguistic behaviour to those with whom they have the most contact. The Burnouts seek connection with their community, and hence spoke more like those in the community. The Jocks on the other hand, sought to transcend the community and thus resisted such changes. Chambers (2003) describes the underlying cause of

sociolinguistic differences as the instinct to establish and maintain a social identity:

“[I]t is not enough to mark our territory as belonging to us by name tags, mailboxes, fences, hedges and wall. We must also mark ourselves as belonging to the territory, and one of the most convincing markers is by speaking like the people who live there.” (p.274)

2.3.3 *Antagonism*

In the same way that speakers who want to signal membership in a group will match their speech to the members of the group, Bourhis & Giles (1977) found that speakers who want to signal their social distance will increase the distinctiveness of their speech (cited in Nettle 1999). This has been demonstrated in the case of the 1961 study of dialects in Martha’s Vineyard in the United States (Labov 1972). On the island there was a resident population and a transient population of tourists. Labov was able to observe, in certain individuals in the resident population, a linguistic change in progress. This change was *away* from the standard realisations of certain vowels (i.e. as they were spoken by the tourists). Individuals adopted these changes depending not only on whether they wanted to be identified as Vineyarders, but also how much they resented the outsiders. By contrast, it was observed that other individuals that felt favourably about the encroachment of tourism shifted their linguistic behaviour *toward* the linguistic norms of the tourists. Thus, not only can linguistic change be seen as a mark of solidarity to a certain social group, but can also be seen as antagonistic. Adolescents typically differentiate themselves from adults through different linguistic markers, and may even diverge further if adults appear to be converging with them (Chambers 2003).

2.3.4 *Norm enforcement*

All of these markers of identity reflect the way in which individuals conform to particular social norms. Individuals do not simply choose to be with those similar to them, but actively or subconsciously alter their behaviour toward or away from groups or social norms. These norms, including linguistic behaviour, are reinforced by the strength of the example of the individuals in the social network to which a particular individual belongs (Chambers 2003). Milroy (1987) describes the function of social networks as a “norm-enforcement mechanism”.

An individual learns particular norms from its social group, which are reinforced with continued contact. Importantly, an individual's degree of integration into a social network affects the level of linguistic conformity that he/she displays. Thus, "[t]he closer an individual's network ties are with his local community, the closer his language approximates to localized vernacular norms" (Milroy 1987). This demonstrates that an individual's position in the social network is correlated with their linguistic behaviour. It follows then that the topology of the social network as a whole affects the language that emerges. It is not just functional pressures on the language that drive its evolution. Indeed, Chambers (2003) conjectures that "[t]he more deeply we inquire into the social meaning of language, the more clearly we see how arbitrary are the values that are commonly attached to it" (p.277). Of course there are many elements of language that are not arbitrary - the presence of linguistic universals rules this out - but many, such as particular realisations of phonemes could simply be negotiated within the social exchange.

2.3.5 *Social mobility*

Social mobility is the extent to which individuals can move between social groups, such as in migration or rising above a class. Mobility tends to promote linguistic conformity, especially in extreme cases, as individuals have to negotiate a shared language. Chambers (2003) describes the most extreme case of social mobility where communities are formed from a diverse group of people where before there was nothing. This kind of situation goes hand-in-hand with conquest and colonisation, whereby hundreds of communities may come into being, comprising individuals of different origin, social class or dialect. In all these communities rapid linguistic homogenisation occurs within the first generation. For instance, the English spoken in former colonies such as Canada, the United States, Australia, New Zealand have much less diversity than England, from which they were colonised.

2.3.6 *Discussion*

Language is a cultural artefact that is not only shaped by transmission from one individual to another, but shows emergent properties correlated with the structure of the social network in which these interactions take place. Individuals tend to interact with those similar to themselves and actively or subconsciously alter their linguistic behaviour toward or away from those they interact with in

order to signify solidarity or antagonism with particular social groups. Simply being integrated in a social group enforces these particular norms, and affects the transmission of information through the network. The topology of the network thus affects the properties of linguistic variations that emerge. Language use is inexplicably tied to social interaction, and takes place on social networks. Computational models of language evolution could therefore benefit from a more realistic treatment of population dynamics.

2.4 Social Networks

2.4.1 Introduction

In social network analysis the environment in which individuals interact is expressed in the patterns or regularities among interacting individuals (Wasserman & Faust 1994). Network theory has been utilised for the study of many real-world networks. The application of network theory to social networks involves, at the outset, identifying *actors* and *relational ties* (Scott 1991). Actors are defined as discrete entities, either individuals or social units, such as people in a neighbourhood, departments in a company or nations in the world. The vertices in the network represent such entities. A relational tie is the social link between two actors, such as friendship, kinship, trade or other behavioural interactions. The edges in a network represent these relational ties.

There are many measures of social network analysis. The focus of social network analysis is on issues of centrality and connectivity. Centrality addressed which individuals have the most influence on a network, i.e. which individuals are best connected to others. Connectivity addresses whether individuals are connected to each other throughout the network (Wasserman & Faust 1994). These measures are focussed primarily on individual vertices (or actors) in the network, or on the properties of small networks. Recently in network research, however, the focus has shifted to large-scale statistical properties of larger networks. When examining networks as a whole, as is the case in computational modelling of social networks, other measures become important. Indeed, when characterising a social network as distinct from other real-world networks, these general properties of the network are more relevant. Though there are many similarities between social networks and other types of networks, studies of large social networks have revealed several unique properties that distinguish social structures

from other types of networks. The three distinguishing properties that have been identified are *clustering* or *network transitivity*, *assortativity* and *community structure* (Newman & Park 2003).

2.4.2 Characteristic measures of social networks

Social networks exhibit *assortative mixing*, showing positive correlations between the degrees of adjacent vertices. In most other networks, a vertex with high degree tends to be connected to vertices of low degree, and vice versa. This forms *hubs* - vertices with many connections - and *spokes* - vertices with very few connections. This type of network displaying such *dissortative mixing*, of which the Internet is a classic example, is referred to as *scale-free*. By contrast, social networks are *assortative*, whereby vertices of high degree tend to be connected to other vertices of high degree (Newman 2003). Informally, an individual with many friends is likely to be connected to individuals who have many friends themselves. Newman (2002) describes two important implications for transmission in social networks related to this assortativity. Firstly, if high degree vertices connect preferentially with other high degree vertices, the network percolates more easily, creating a large subnetwork of higher average degree than the rest of the network. Transmission in this core group would thus be more rapid. Secondly, in a dissortatively mixed network, the removal of a vertex of high degree would severely disrupt the transmission, since these vertices provide connections for many sparsely connected vertices. In an assortatively mixed social network, however, the network is resilient against the removal of vertices of high degree, since they are likely to be connected themselves to other vertices of high degree. Thus, transmission in social networks is more robust than other networks to the removal of vertices, at least to removal of the high-degree vertices.

The second distinguishing property of social networks is high clustering or network transitivity (Newman & Park 2003). This is the phenomena of mutual acquaintances, whereby, given three vertices v_A , v_B and v_C , the presence of an edge between vertices v_A and v_B and another between v_B and v_C makes it likely that there will be an edge between v_A and v_C . This feature of social networks makes their community structure very robust, as the formation of such 'triangles', known as *triadic closure* (edges between vertices v_A , v_B and v_C) makes communication through the community rapid without the presence of hubs (Newman & Park 2003).

Thirdly, community structure has been identified as a defining feature of social networks. Individuals belongs to groups or communities, such that there is a high density of connections within groups and low density between them (Newman & Park 2003). Individuals may belong to many groups, giving rise to a nested hierarchy of nested social communities that in many cases exhibit a self-similar structure (Boguna et al. 2003, Girvan & Newman 2002). Analysis of real-world social networks within organisations or small communities has revealed that there often exists a central group of closely connected individuals and a larger proportion of individuals who are less densely connected, both to the core and to each other; a so called *core periphery pattern* (Brass 1985).

Lastly - though it is not unique to social networks - it is worth mentioning an interesting and oft cited property of many natural networks: the “small-world effect” - such networks have a high density of structured local connections and a few random remote connections (Watts & Strogatz 1998). Travers & Milgram (1969) conducted an experiment that gave us the phrase “six degrees of separation”. Individuals were selected randomly from Boston and Omaha in the United states, and given the task of directing letters to a target person in Boston, not acquainted with the subjects. The participants forwarded their letters to an acquaintance of theirs that they believed to be closer to the target than themselves, and these recipients did the same. Of those letters that reached their target, the average number of acquaintances that the letters passed through was around 6. So not only do short paths exist in social networks, but more importantly, individuals in a network can find these paths, not through an individuals awareness of the network as a whole, but simply through local knowledge of their acquaintances. Kleinberg (2000) demonstrates that local knowledge and network structure is a crucial addition to the small world property in social networks. If connections are made uniformly and at random, the resulting network has a homogeneity of structure, but much variance between the properties of adjacent vertices. Short paths exist, but individuals cannot find them due to their many, extremely heterogeneous social contacts.

2.4.3 *Evolving social networks*

One of the main aims in network analysis is to create models of networks that can aid in understanding the meaning of the properties of the resulting network (Newman 2003). Computational models of growing networks can give us clues as to how the statistical properties of the network originated, and the nature

Network model	Growth method
Toivonen et al. (2006)	Attachment to random vertex and vertex's neighbours
Jin et al. (2001)	Random attachment, attachment of adjacent vertices
Newman (2002)	Edges are removed and added based on the similarity of their degree distributions
Boguna et al. (2003)	Preference for social similarity; Fixed, continuous traits
Watts et al. (2002)	Preference for social similarity; Fixed, continuous traits; Groups defined <i>a priori</i>

Table 2.2: Summary of computational models for growing social networks.

of their interactions. In social network analysis one goal is to explain how the unique properties of social networks - assortativity, clustering and community structure - came to be as they are. There have been several methods put forward for growing computational models of social networks. Table 2.2 gives some examples of models of social network growth. All are governed primarily by preferential attachment, whether explicit, through attachment to “similar” vertices, or implicit through attachment to the neighbourhood of connected vertices. The algorithms that generate these models are geared towards satisfying one of the unique properties of social networks: assortativity, clustering or community structure. The other properties tend to result from explicitly building one of these into the evolution process. These models are described below.

Evolving networks for transitivity

Toivonen et al. (2006) describe a model of growing social networks. The network is a grown model in the sense that new vertices are added to the network during the simulation. The network begins with a seed network of N_0 vertices. At each timestep, an average $m_r \geq 1$ vertices are chosen as initial contacts and $m_s \geq 0$ neighbours of each initial contact as secondary contacts. A new vertex is connected to each of the initial and secondary contacts, and the algorithm is repeated until the network reaches the desired size. For any non-negative distributions of m_r and m_s , the network demonstrates assortative degree correlations, clustering coefficient comparable to real-world social networks and community structure. The first two properties can be attributed to the neighbourhood connections. In the first case, when a new vertex is added to a neighbourhood, the degrees of all the neighbours are increased by one, maintaining the existing degree-degree

correlation. The high level of clustering is a trivial outcome that is built into the algorithm, whereby a new vertex will connect with adjacent vertices, creating the triadic connections measured by the clustering coefficient. Community structure is explained partly by this mechanism also, whereby a new vertex tends to connect to members of the same community, but the random connection also allows new vertices to connect to more than one community, thus acting as “bridges” between communities. An important point about this model is that it is a growing social network model, whereby new vertices are continually added to the network. This may be a realistic assumption for the growth of networks such as the World-Wide-Web (Faloutsos et al. 1999), but social networks differ significantly in this respect. In a social network, the number of vertices changes at a greatly lower rate than the number of edges. A more realistic social model would have a fixed number of vertices and vary the configuration (and perhaps quantity) of edges exclusively.

A similar model of the evolution of a social network by neighbourhood attachment is given by Jin et al. (2001). They propose a minimal set of features that a model of social network evolution should have:

1. Fixed number of vertices: we consider a closed population of fixed size.
2. Limited degree: the probability of a person developing a new acquaintance should fall off sharply once their current number of friends reaches a certain level.
3. Clustering: the probability of two people becoming acquainted should be significantly higher if they have one or more mutual friends.
4. Decay of friendships: given that the number of vertices is fixed, and the degree is limited, friendships must be broken as well as made if the evolution of the network is not to stagnate.

Jin et al. (2001) propose two different models of evolving social networks using the above features. Both are based, as is the previous model, on making connections between mutual friends. The first model attempts a more realistic numerical portrayal of the evolution of a social network, with weighted edges, and meetings, maximum degree and decay all variable on probability distributions. The large number of free parameters may make the model more realistic,

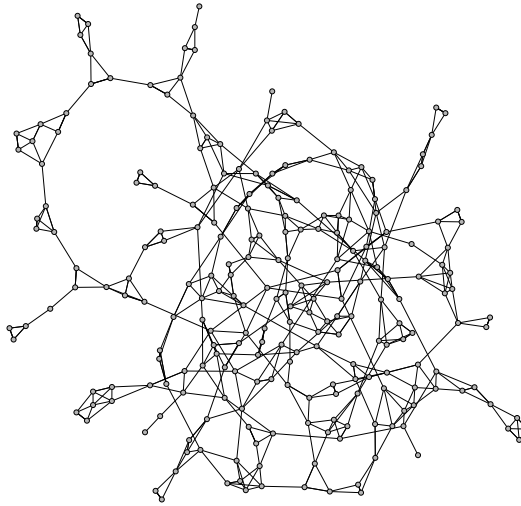


Figure 2.3: Social network grown through attachment of neighbouring vertices (Jin et al. 2001)

but it makes the evolution of the network computationally intensive and difficult to analyse. The second model is a much simplified version of the first, but importantly it also reproduces the characteristic features of social networks. The model consists of three steps: random connection, connection to neighbourhood and decay. The network is initialised with N vertices and no edges, then steps 1 and 2 below are run at each timestep without the third until the network has reached a certain mean degree, at which point all three steps are run. Meetings occur between vertices at a rate r linear to their number of mutual friends: $r = r_0 + r_1 m$ Formally:

1. $n_p r_0$ pairs of vertices are chosen at random, where n_p is the number of pairs of vertices in the network. If a pair meet that do not do not have an existing edge between them and neither has the maximum degree z^* , then an edge is established between them.
2. $n_p r_0$ vertices are chosen at random with probability proportional to their number of mutual neighbours. For each vertex, two of its adjacent vertices are chosen, and there is not an existing edge between them and if neither has degree z^* , then an edge is established between them.
3. Choose $n_e \gamma$ vertices with probability proportional to their degree z_i , where n_e is the number of existing edges in the network and γ is a factor representing the probability per unit time that an existing edge will be removed. For each vertex choose one adjacent vertex and random and remove the edge between them.

Figure 2.3 shows a network evolved using the above steps, with $N = 250$, $r_0 = 0.0005$, $r_1 = 2$, $\gamma = 0.005$ and $z^* = 5$. Values for mutual introductions r_1 and decay γ were chosen so that the former have some stability, such that even when edges are broken it is probably that they will be remade quickly. In this way, the triadic closure of vertices is a self-sustaining structure in this model. The model produces clustering coefficients comparable with those of real social networks and visible community structure. Interestingly, however, degree correlation is not mentioned. The first measure - the clustering coefficient - is not such a surprising result, since, as in the previous model, the algorithm was geared to produce transitive structures. Community structure occurs in the simulation when a region forms that has a higher than average density, such that there are more pairs of vertices in a region. New edges will be added preferentially between these pairs, further increasing the density of the region. In this way, fluctuations in network density may cause the formation of highly connected communities (Jin et al. 2001). The measure of degree correlation remains open.

Evolving networks for assortativity

The above models explicitly build network transitivity into models of network evolution through attachment to neighbourhood vertices. Newman (2002) describe a model of evolving social networks that explicitly builds assortativity into the algorithm. Preferential attachment processes, whereby the probability of a source vertex connecting to a target vertex is some function of the degree of the target vertex, provide explanations for the dissortative degree distributions found in many classes of networks (Barabasi & Albert 1999b). To achieve the assortative mixing seen in social networks, however, another mechanism must be at work. Newman (2002) suggests that the probability of attachment should not only depend on the degree of the target vertex, but also on the degree of the source vertex. He describes a model of generating a network specifically to demonstrate assortative mixing. Starting with a random network (Molloy & Reed 1995), two edges are randomly chosen, denoted by the vertex pairs (v_1, w_1) and (v_2, w_2) . The degrees of the vertex pairs (j_1, k_1) and (j_2, k_2) are measured, minus the edge in question. These two edges are then replaced with edges (v_1, v_2) and (w_1, w_2) with a probability that depends on whether the joint degree distribution e_{jk} of each potential edge is greater than the original edges. Formally, the edges are changed with probability $\min(1, (e_{j_1 j_2} e_{k_1 k_2}) / (e_{j_1 k_1} e_{k_1 k_1}))$. In this way, edges are replaced to connect vertices of high degree, thus building assortativity into the network. High degree vertices stick together in a subnetwork that

displays higher degree than the network as a whole. While this model achieves a level of assortativity comparable with that of observed social networks, the mechanism for growth of the network does not reveal, or is not informed by how real-world social networks evolve.

Evolving networks for community structure

In the above models, the evolution of the network is governed by the structural properties of the network. Thus, an individual chooses a community to belong to by virtue of their own position in the network. However, individuals do not exclusively form relationships based on, for instance, mutual acquaintances. This section describes models of social network evolution whereby individuals connect preferentially based not on properties of the network, but on a preference for social similarity. This principle of homophily was one of the first features of social networks that was noted in structural analysis of social networks, whereby there is a positive relationship between the similarity of two vertices in a network and the probability of an edge between them (Freeman 1996). Boguna et al. (2003) describes a model of social network evolution based on the principle of homophily. Vertices in the network possess identities: characteristics that assign them to various social groups. These characteristics define an individual's social position relative to the rest of the population. In a previous paper, Watts et al. (2002) define such social identities, and give individuals the task of partitioning their world into a series of layers, the top layer representing the entire network and lower layers representing divisions into increasingly small social groups. Where Watts et al. (2002) define the hierarchies *a priori*, here hierarchies emerge as a result of the construction process. Simply put, a set of N unconnected individuals are randomly placed in a social space H of different social features (intended to represent profession, religion and so on) according to a density $p(\vec{h})$. Thus, each individual i has a position in the social space given by a vector $h_i \equiv (h_i^1, \dots, h_i^{d_H})$, where d_H is the dimension of H . Each element of the vector is given by a continuous variable growing with the size of the population, such that each individual is unique. Edges are constructed between vertices with a probability decreasing with their relative social distance. When individuals are assigned a random, uniformly distributed position in the social space, networks can be generated that exhibit high clustering coefficients, assortative degree correlation and hierarchical community structure. Though the network is homogeneous in the social structure, the stochastic process of assigning individuals an identity means that there will be small fluctuations of densities of individuals in

the social space. Individuals from denser clusters will then form communities of close individuals. Thus, the distinctive properties of social networks can be generated by the very presence of communities in the social space.

2.4.4 Discussion

Many network growing models have explicit neighbourhood rules, and many emphasise that the rules governing the evolution of the network are not in themselves important, focusing instead on the structure of the resulting network. There appears to be a gap in the literature for more realistic models of rules governing the creation of a social network. Algorithms for evolving social networks appear to build the characteristic features of social networks into the model, rendering the results, at least for the particular measure to which the network is built, less than surprising. Admittedly, most models of social network evolution are not intended to shed light on the actual mechanisms for growth of the network, but rather have the evolved social network as the goal. Nevertheless, a more interesting approach might be to see which social practices identified in the sociolinguistic literature lead to the creation of characteristic social networks.

CHAPTER 3

Methodology

3.1 Introduction

This chapter sets forth the methodology for creating and evaluating a computational model of the evolution of language and social network structure. A justification for the study is given, followed by a series of aims to adhere to in designing and evaluating the model. Some significant choices made in the design of the model are then given, followed by a definition of the model itself. Finally, a series of measures that will be used to evaluate the resulting network and language are defined.

3.2 Aims

There are two purposes of the study: the first is to attempt to replace the building of social networks that take explicit measures to produce the characteristic measures of social networks with more realistic, implicit methods. The second is to examine the coevolution of social network structure and language by means of learning and interactions between agents. The purpose of this study is to produce a model of network evolution and language evolution that:

1. Uses a plausible algorithm for the evolution of social ties
2. Displays the features of real world social networks
 - Assortative degree correlations
 - High clustering or network transitivity
 - Community Structure
3. Demonstrates the coevolution of social and language structure.

Some of the mechanisms by which these are negotiated are discussed in chapter 2.3. This study focuses on the principle of homophily, detailed in 2.3.2.

3.3 Design choices for the model

Though the structure of the resulting network is important - it must display the unique properties of a social network - the growth mechanism is also important. In other social network evolution algorithms, the aim is to generate a social network in order to, for instance, study sociodynamic phenomena (Toivonen et al. 2006). As such, the method of generating the network is less important, it is merely sufficient that the appropriate network is built. Whereas these studies emphasise the properties of the resulting network for study and often describe the mechanism as unimportant, here the mechanism needs to be plausible. The principle of homophily is commonly cited in the sociolinguistic literature (see review chapter 2.3.2), and will be the primary mechanism for the model. Secondly, learning will be added to see what effect it has on the network growth. Other models described in chapter 2.4.3 also use this principle of homophily to generate social network structure, but have the agents construct their own hierarchies of group membership *a priori* (Boguna et al. 2003), or have the agents traits remain static (Boguna et al. 2003, Watts et al. 2002), thus building in community structure at the outset of the simulation whether the agents are aware of it or not. The fact that agents will dynamically adjust their traits during the simulation distinguishes this model from other models of social network evolution.

The principles we can distil from the literature reviewed in chapter 2 on computational models of the evolution of social networks are:

- A model of the evolution of social networks need not attempt to capture the microscopic details of social dynamics where a simple general model will suffice;
- Due to the disparity in the rate of change of vertices to edges, the addition and removal of vertices should be at a much smaller rate than the addition and removal of edges. Given that a simple model may be just as effective as a more complex one, the number of vertices can remain constant;
- Vertices should have a limited degree;
- Edges need to be removed as well as added if the network is not to stagnate.

In order for an agent to learn it must possess a changeable internal state. Nettle (1999) introduces the notion of a *human linguistic pool*, an abstract entity analogous to the human gene pool containing all the bits of linguistic structure that are found in human languages. The elements of are not languages but linguistic items. These linguistic items are defined as any piece of structure that can be independantly learned and transmitted from one speaker or language to another. The important point is that the items are potential *replicators*, that can be passed from one speaker to another. In this model the elements will be discrete, representing, for instance, choices between lexical or grammatical alternatives. There will be no functional selection: each linguistic item is as equally likely in this respect, and not relative to any of the others. The set of traits that an agent possesses can be seen as their vector position in the social space, just as they have a position in the network.

When choosing a learning algorithm the competing considerations are between endowing agents with simple properties where the focus is on the global behaviour of the population, or with more complicated and realistic structures that may confuse the experimental output. This study follows the former method, focussing on the evolution of the network structure using simple, abstract learning mechanisms. The learning algorithm simply has an agent changing a random trait to the value of the corresponding trait in an agent that it meets. This is because, in this initial, exploratory model, any interactions and comparisons between agents merely need to provide a measure of similarity.

The model should start with an existing network to provide the basis for the network and language evolution. The choice here is to initialise the network with a fixed number of vertices with a random assignment of traits and edges. This initialisation of the model with random traits and edges could be seen to correspond with the most extreme case of social mobility described in chapter 2.3.5.

3.4 The model defined

The model consists of a population of N agents. Each agent is randomly assigned a position in the social space H_s given by the vector $h_i = (h_i^1, \dots, h_i^{d_H})$, where d_H is the dimension of H , i.e. the number of traits that an individual possesses. Each trait is a discrete characteristic represented as an integer, and one of a selection

of d_h , the dimension of each trait, such that $(0 \leq h_i^j \leq (d_h - 1))$. Importantly, though the traits are represented as numbers they are all independent and do not have any correlation between them, such that 5 is not greater than 4, nor is it closer to 6 than 7, for instance. These sets of traits could be seen as the parameters for a certain language, for instance, where each trait represents a linguistic item drawn from the human linguistic pool.

The vertices in the network are initialised with random values for their traits (i.e. a random position vector in the social space). These vectors are not necessarily unique. The network is initialised with a random distribution of edges with each vertex having a maximum degree z_* . The simulation then proceeds by iterating the following three steps:

1. **Preferential attachment:** Choose two vertices v_i and v_j with probability proportional to their degree, such that neither have the maximum z_* degree and they fulfil one of the following criteria:
 - both of their degrees are less than or equal to a minimum degree for connection $z_c \leq z_*$; or
 - they are more (or equally) similar to each other than they are to at least one of their respective adjacent vertices. Similarity is defined as the number of traits they have in common.

Form an edge (v_i, v_j) between them.

2. **Decay of friendships:** Choose a vertex v_i with probability proportional to its degree, and that has greater than or equal to a minimum degree for disconnection $z_d \leq z_*$ and disconnect one of its adjacent vertices v_j with which it has the least communicative success, i.e. remove the edge (v_i, v_j) .
3. **Learning:** Choose a vertex v_i with probability proportional to its degree that has at least one adjacent vertex. Select an adjacent vertex v_j and a trait h_i^k at random. Change the h_i^k to the value of the corresponding trait h_j^k in v_j .

The evolution proceeds for a certain number of iterations, or timesteps t .

3.5 Measures

3.5.1 Introduction

This section outlines the network measures used in the analysis of the network simulations. The first step is to identify that the networks being built possesses the unique characteristics of social networks - assortativity, high clustering and prominent community structure. Secondly, and most importantly, that the structure of the language is correlated with the structure of the network, such that the two are predictably coevolving. In order to determine the possibly nontrivial features in the evolving networks, they can be compared with the null model of purely random network growth with the same degree distribution.

3.5.2 Measures of social network growth

Assortativity

To evaluate the degree-degree correlation of a network and determine whether it possesses assortative degree mixing, the correlation coefficient of the network can be calculated. This gives a single number that gives the mean probability over the network that adjacent vertices have the same degree. In an undirected network, as is being studied here, it is given by the Pearson correlation coefficient r_d of the degrees of the vertices at either ends of an edge. Newman (2002) gives a straightforward method of calculating the Pearson correlation coefficient on an observed network, given by

$$r_d = \frac{M^{-1} \sum_i j_i k_i - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}{M^{-1} \sum_i \frac{1}{2}(j_i^2 + k_i^2) - [M^{-1} \sum_i \frac{1}{2}(j_i + k_i)]^2}, \quad (3.1)$$

where j_i and k_i are the degrees of the vertices at the end of the i th edge, with $i = 1 \dots M$. It will lie in the range $-1 \leq r \leq 1$. It has been demonstrated that the value of r is negative for essentially all networks *except* social networks, due to the assertion that disassortativity is the natural state for all networks (Park & Newman 2003). Thus, without additional mechanisms to achieve the assortativity seen in social networks, all networks will tend towards disassortativity, barring the random network, which is by definition neither assortative nor disassortative (Newman & Park 2003). Table B.2 shows a series of real-world networks, the first group being social networks all showing positive degree correlations and

the second being technological and biological networks, all of which show negative degree correlations. Thus, in order to demonstrate that the evolved network shows this property of dissortativity, it suffices to demonstrate that r is positive.

Network	n	r
physics coauthorship	52 909	0.363
biology coauthorship	1 520 251	0.127
mathematics coauthorship	253 339	0.120
film actor collaborations	449 913	0.208
company directors	7 673	0.276
Internet	10 697	-0.189
World-Wide-Web	269 504	-0.065
protein interactions	2 115	-0.156
neural network	307	-0.163
food web	92	-0.276
random graph		0

Table 3.1: Degree correlation coefficients r_d for a number of different networks of size n (from Newman 2002)

Clustering Coefficient

The clustering coefficient $C(p)$ is a measure of the topology of the network, which measures the density of *connected triples* in the network. Suppose that a vertex v has k adjacent vertices, then at most $k_v(k_v - 1)/2$ edges can exist between them, when every adjacent vertex is connected to every other adjacent vertex. C_v then is the fraction of these edges that actually exist, and $C(p)$ is the average over all v (Watts & Strogatz 1998). Thus $C(p)$ is the probability ($0 \leq C(p) \leq 1$) that two vertices adjacent to a given vertex are themselves adjacent. This is given by

$$C(p) = \frac{1}{N} \sum_i \frac{2E_i}{k_i(k_i - 1)}, \quad (3.2)$$

where E_i is the number of edges between the adjacent vertices of the i th vertex. The clustering coefficient decreases as N increases, since N^{-1} , and so is very small for large networks. Calculating whether the clustering coefficient is significant for a network of the same number of vertices and edges means comparing it with the random model, generated by random connections between vertices. The statistical ANOVA test will determine whether the values of the clustering coefficient observed in the model are significant.

Prominent Community Structure

Communities can be defined as sets of vertices with dense internal connections, and relatively sparse connections between them (Toivonen et al. 2006). The most basic analysis of the structure of a network is simply by visual analysis. Observation of the layout of the network can give strong clues as to the community structure of the network. Indeed, this is one of the primary methods of network analysis, and a visual analysis of the properties of a network is a useful way to gain an understanding of its structure (Newman 2003, Newman & Girvan 2003). Visualisation of the networks was performed with the graph drawing software Himmeli, which uses a simulating annealing algorithm to arrange vertices in a two dimensional plane according to the network topology (Mäkinen n.d.). In observing the network it should there should be an obvious community structure, showing distinct clusters of vertices. The core periphery pattern of Brass (1985) may also be observable.

3.5.3 *Language entropy*

The entropy of a system can be thought of as the degree of randomness in a single random event. In this simulation, the entropy of the language system is the average uncertainty that a selected vertex v_i will have a particular language (or vector h_i) in the social space H_s . The entropy H_e of the system is defined by the Shannon entropy equation, first described in Shannon (1948):

$$H_e(H_s) = - \sum_{h_i \in H_s} p(h_i) \log_b p(h_i) \quad (3.3)$$

with the convention that $0 \log 0 = 0$, and $p(h_i)$ is the probability of selecting h_i in H_s . The entropy H_e of the system will be high in a heterogeneous population of agents, decreasing as agents converge on languages. Of course, if the agents are unable to change their language in any way (through learning in this model), the entropy of the system will remain constant. If the agents converge on common traits, within communities, for instance, the entropy will decrease.

3.5.4 *Social Difference vs. Language Distance*

McPherson et al. (2001) asserts that “homophily implies that distance in terms of social characteristics translates into network distance”. Thus we would expect individuals that share certain characteristics to be close together in the network.

As in section 3.5.2 above, the first method of observing this property is simply visual. The traits that an individual possesses can be translated to a visual property such as colour, and we should observe clusters of similarly coloured individuals corresponding to clusters in the network. For a simulation of individuals with up to three traits this is relatively straightforward. Each trait corresponds to a value in the colour spectrum of either red, green or blue.

The second measure of social distance versus language distance is the correlation between the difference between the traits of two vertices and their social distance. The language distance of two vertices is defined as the Hamming distance of their traits. Thus, the quantitative difference between two vertices is the number of their traits that differ. The social distance is the length of the geodesic path(s) between the vertices. Only vertices that have at least one geodesic path are considered to prevent the measure being skewed for highly disconnected graphs. To test whether the relationship between the language distance and the social distance is predictable, the Pearson's correlation coefficient r_{ls} of the two measures can be determined. To determine whether this measure is significant it is tested against the random graph model with an ANOVA statistical test. Further, we can visually plot the three measures (social distance, language distance and language distance correlation) in the evolving network over time.

3.6 Testing the model

There are three primary mechanisms at work in this algorithm for social network evolution: preferential attachment and detachment to and from vertices by trait similarity; preferential attachment to and from vertices by degree similarity; and convergence on similar traits through learning. The effects of these mechanisms on the network structure will be explored.

3.6.1 *The Null model - the random network*

The null model is the starting model for the algorithm, whereby the vertices are initialised with random values for their traits, and the network is initialised with a random distribution of edges with each vertex having a maximum degree z_* . This network is built simply by selecting two random vertices v_i and v_j and making an edge between them if neither has the maximum degree z_* . The network is built when it reaches a certain mean degree. This null model is an ideal basis

for testing whether the mechanisms implemented in the following sections have an appreciable effect.

3.6.2 *Test 1: The effect of preferential attachment by trait similarity*

The effect of preferential attachment by trait similarity means that agents connect and disconnect either randomly or based on the similarity of their degrees, as detailed in 3.6.3 below.

3.6.3 *Test 2: The effect of preferential attachment by degree similarity*

Preferential attachment by degree similarity is implemented in the algorithm through the minimum degree for connection z_c disconnection z_d . This test will investigate whether this mechanism has an appreciable effect on the network and language evolution.

3.6.4 *Test 3: The effect of learning*

The learning mechanism means that agents will make their traits similar to those they are connected to. In the absence of learning there will be no change to the languages present in the network, and the model resembles those of Boguna et al. (2003) and Watts et al. (2002). With learning, however, in addition to seeking out those they are similar to, they are able to actively make themselves similar to their neighbours. This section will investigate whether learning has an appreciable effect on the network and language evolution.

CHAPTER 4

Results

4.1 Introduction

This chapter details the results obtained from an exploration of the algorithm detailed in chapter 3. The models are all built with the number of vertices $N = 250$, the number of traits $d_H = 3$, the possible values for each trait $d_h = 5$, maximum degree $z_* = 8$, maximum degree for random connection $z_c = 3$ minimum degree for disconnection $z_d = 4$ and $t = 50000$. Sample evolving networks at $t = 0, 10000, 20000, 30000, 40000$ and 50000 typical simulations for each configuration of mechanisms are presented in Appendix . Statistics referred to in the following sections are given in Appendix B, for clustering coefficient (B.1), degree correlation (B.2) and language difference vs. social distance correlation (B.3). The study explores the minimum mechanisms necessary for building a characteristic social network, and the effect that learning/cultural transmission has on this process.

4.2 Evolved networks

4.2.1 *The Null model - the random network*

The random model is the baseline model for the following simulations. At each timestep a random edge is removed and replaced with another random edge between two vertices that do not already have the maximum z^* number of connections. Figure 4.1 shows such a random network. Results are consistent with other models of random networks. There is no correlation between the degrees of adjacent vertices. The clustering coefficient is negligible, between a minimum

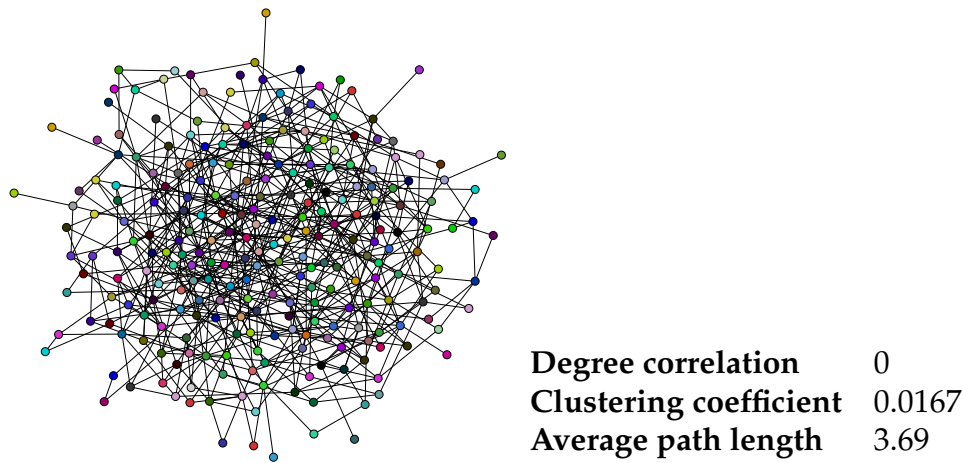


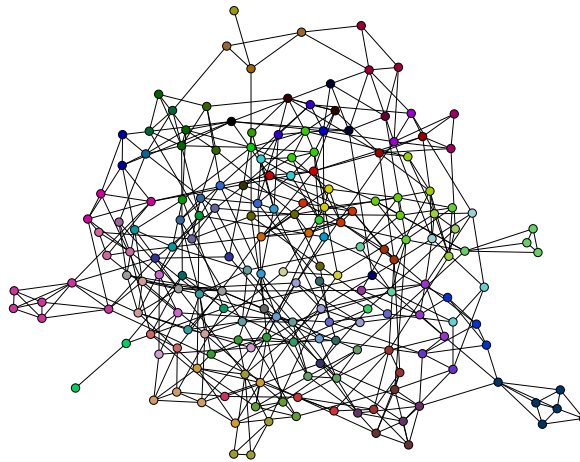
Figure 4.1: Random network

of 0.01 and 0.03 with a standard deviation $\sigma = 0.005$. This demonstrates that with no other mechanisms in place, given a vertex v , there is only a negligible chance that two vertices adjacent to v are themselves connected. It is worth noting that the average path length is quite small, however, though not surprising given that the network is clustered in a single group, with each vertex as likely to connect to any other. This demonstrates that the small-world property of social networks is not necessarily unique. There is also a notable lack of any kind of community structure in the network, nor any order to the languages (given by the vertex colours). All these factors demonstrate that additional mechanisms must be in place to create the characteristic measures of a social network

4.2.2 Test 1: The effect of preferential attachment by trait similarity

After the baseline random model has been established, it is now possible to test the effects of the various mechanisms of network evolution on the social network and trait distribution. First among these is the effect of preferential attachment. An agent will connect to another agent based on the similarity of their traits. In this respect, the model is similar to Freeman (1996) and Boguna et al. (2003), discussed in chapter 2.4.3. There is no learning involved, nor are agents selected on the basis of vertex degree. Figure 4.2. shows the main component of a network grown from such a model (the evolving structure of the network is shown in Appendix A.2). Smaller groups of one or two vertices made up of vertices with identical traits are split from the main component, and not shown here. The mean clustering coefficient ($\langle C(p) \rangle = 0.4099$ with $\sigma = 0.0231$) is extremely high for a network of this size, and significant with respect to the null model.

Similarly, the mean degree correlation ($r_d = 0.3691$ with $\sigma = 0.0510$) is also significantly higher than the null model. A visual analysis of the network reveals small communities of similar clustering together, with bridge vertices connecting them that appear to have similar traits to both groups. Finally, and most importantly, grown network displays a significant level of correlation between the language difference and social distance ($r_{ls} = 0.6034$ with $\sigma = 0.0431$). It is therefore possible to evolve a reasonable social network displaying their characteristic properties with preferential attachment for similarity alone.



Clustering coefficient	0.4099
Degree correlation	0.3691

Figure 4.2: Network built with preferential attachment based on trait similarity

4.2.3 Test 2: The effect of preferential attachment by degree similarity

Next we turn to the effect of preferential attachment by degree similarity. Surprisingly, this has no significant effect on the behaviour of the model. When preference for degree alone is used to evolve the network, there is no significant variation from the null model: with clustering coefficient ($C(p) = 0.0141$ with $\sigma = 0.0057$); degree correlation ($r_d = -0.0099$ with $\sigma = 0.0282$) and language difference and social distance correlation ($r_{ls} = 0.0020$ with $\sigma = 0.0080$) all explainable by random processes of attachment and detachment. Previously, Molloy & Reed (1995) have demonstrated a model that uses degree correlation to build a social network (outlined in chapter 2.4.3), though their model is more strict in this respect, exchanging edges by preference for those that have the correct degree distribution. It seems that the simple method of connecting two vertices that are both either below or above a certain degree is not sufficient to evolve social

networks. In addition, since there is no preferential neither learning nor preferential attachment for trait similarity, there is no use of the language at all, and the population unsurprisingly remains heterogenous. Lastly, combining preferential attachment for degree with preferential attachment for traits does not produce any behaviour that differs significantly from the preference for traits alone: ($C(p) = 0.4377$ with $\sigma = 0.0402$), ($r_d = 0.3786$ with $\sigma = 0.0678$) and ($r_{ls} = 0.6131$ with $\sigma = 0.02181$). Figure 4.3 shows the differences between the two models.

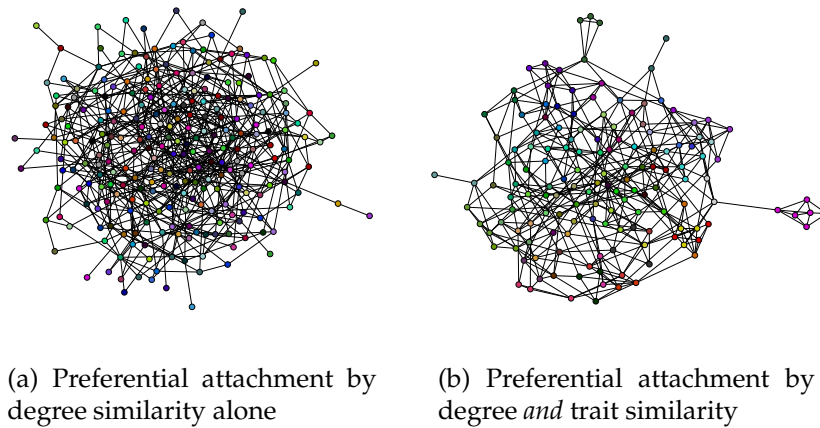


Figure 4.3: Preferential attachment by degree and trait similarity

4.2.4 Test 3: The effect of learning

Up to this point the test have not altered the distribution of the language in any way. Each agent's position remains static in the social space H , and it is only the network that evolves. Learning means that agents may move about the social space. They may actively alter their position toward (or indeed, away from) other agents through changing the value of their traits. The effect of learning with random attachment is similar to the effect of an unordered population of agents. Agents may eventually converge on one or more main languages, but since the agents show no preference toward or away from the traits of other agents, learning has no effect on the evolution of the network. This can be seen in the first network in figure 4.4, which shows a more homogeneous population than the purely random model, but is nonetheless effectively a random network. There is even no correlation between language difference and social distance, presumably due to the completely random connection that makes an agent as likely to connect to any other agent in the network and disconnect from any adjacent agent. Thus, there is no chance for separate distinct populations to form.

When learning is combined with preferential attachment, however, there is a dramatic change, clearly visible in the right network in figure 4.4. Now, distinct communities have formed that share a common language within each group but differ between them. Both the mean clustering coefficient ($C(p) = 0.2918$ with $\sigma = 0.093$) and the degree correlation ($r_d = 0.3133$ with $\sigma = 0.0447$) are significantly higher than that of the null model, and both contribute to the final scores. Surprisingly in this series of simulations, though the correlation between language difference and social distance is significant ($r_{ls} = 0.5891$ with $\sigma = 0.0107$), the effect of learning on this correlation appears not to be significant. This could be attributed to the fact that learning increases the size of the language groups that form, thus making longer paths within groups of identical languages. Further tests will be required to determine whether this is the case (see chapter 5.3.1).

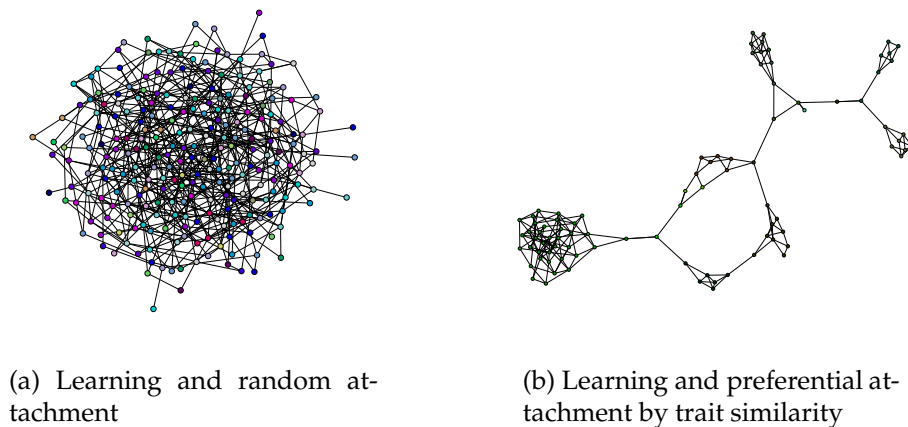


Figure 4.4: The effect of learning and preferential attachment based on trait similarity

4.3 Evolving networks

The above analysis of the behaviour of the built networks has revealed that, while learning does have a significant effect on the evolution of the social network, it is also true that the characteristic properties of social networks can be constructed with simply preferential attachment and detachment depending on similarity. In order to investigate further the effect that learning has on the network structure we can examine longitudinal data of the language as it evolves. As has been shown in section 4.2, the effect of degree preference on this model of social network evolution does not appear to be significant, and will therefore not be explored further.

4.3.1 Clustering coefficient over time

Figure 4.5 shows the clustering coefficient of two evolving networks over time in a typical run of the simulation. The green line is the simulation with preferential attachment and no learning, and the red line is with preferential attachment learning. Without learning, the transitivity of the network rapidly approaches a value of approximately 0.38, at which point it remains more or less constant for the rest of the simulation. With learning, however, the network transitivity gradually increases and does not appear to converge on a limit, at least within the time that the simulation is run. The qualitative explanation for this is that in a population where the values of h_i are fixed, a vertex v_i will only have a limited number of other vertices to which it can preferentially attach. It follows that its adjacent vertices will also have a limited number of vertices with which to attach to, so the system will tend to cap at a certain level depending on the size of the population N and the assignment of traits h_i at the outset. When agents are able to learn, however, they can change their traits to match other agents and form larger groups with the potential for more triadic connections.

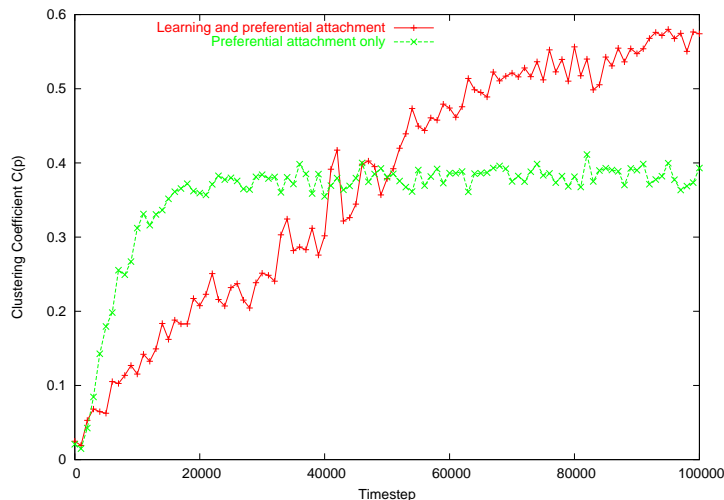


Figure 4.5: Clustering coefficient $C(p)$ over time

4.3.2 Degree correlation over time

Similar to the clustering coefficient over time discussed above, the degree correlation r_d for preferential attachment without learning rapidly approaches a certain value at which it appears to remain (approximately 0.3). This is shown in figure 4.6. With learning, however, the degree correlation of the network increases gradually and only appears to reach a limit towards the end of the simulation.

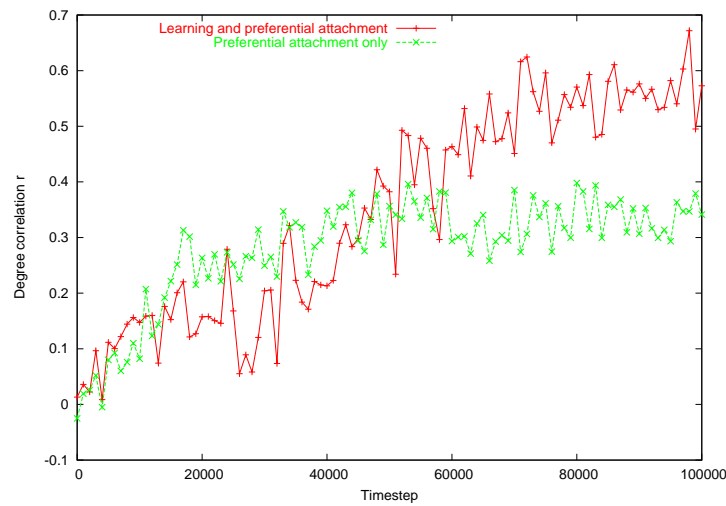


Figure 4.6: Degree correlation r_d over time

The explanation for this behaviour is similar to above also. In a population that does not learn there will be a fixed number of vertices to which an agent will preferentially connect. There is still a possibility of random connection to dissimilar vertices if the vertex has a small degree however. This will happen throughout the simulation after the agents have converged in communities of similar agents, thus connecting agents of different degrees and capping the degree correlation r_d . By contrast, a population of learners can alter their traits and form larger communities, connecting preferentially to other similar degree vertices in within their community, resulting in increased assortative mixing in the network.

4.3.3 Language entropy

Where a population of static agents unable to change their traits the entropy of the system will remain constant. When the agents are able to change their traits the entropy is variable. Figure 4.7 shows the entropy H_e of the language system in a population of learners over time in a typical run of the simulation, given by equation 3.3. As the population of agents settle into a community structures and negotiate one or more shared languages between them, the entropy of the system is reduced substantially. As agents change their traits to match their neighbours the number of languages in the population is reduced, and the entropy rapidly decreases. This measure of the nature of the traits in the population gives a solid indication that the agents are converging on shared traits.

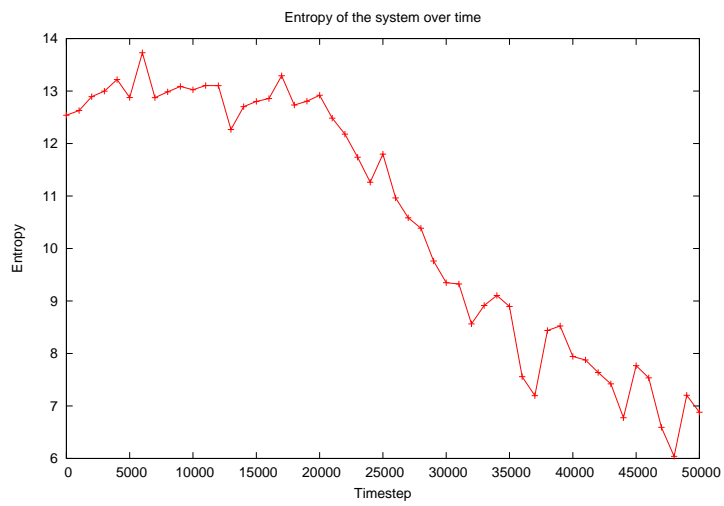


Figure 4.7: Language entropy H_e over time

CHAPTER 5

Discussion

5.1 Introduction

This chapter gives a discussion of the results obtained in chapter 4. There are several important results that emerge from the simulations run on the model, as well as verifications of existing results. These are discussed below. Secondly, suggestions for further work that might be undertaken on the model are outlined and justified. Lastly, the study is concluded with some closing remarks.

5.2 Discussion of the results

There are several results from simulations of the model that confirm existing results obtained elsewhere. Firstly, an obvious result is that the purely random attachment model is not sufficient to produce social network structure. Though it does produce short path lengths characteristic of a “small-world” network, this measure in itself is not sufficient. Secondly, preferential attachment based on similarity *is* sufficient to produce the characteristic measures of social networks. However, as with other models of social network growth, this is merely building one property of the desired network into the algorithm. If agents begin with a set of traits that define their membership in a community, then it should come as no surprise when such communities form. Adding learning to the model means that these communities are not defined at the outset. Instead they must be negotiated within the cultural exchange.

This study has focussed primarily on the evolution of the network structure, since the simple model of homophily relies only on a measure of how similar

agents are to each other. It requires only that an individual is able to determine the relative similarity of their own behaviour to another agents behaviour, and adjust it accordingly. This, combined with the literature on models of language evolution in populations, suggest that the relationship between the evolution of language and the evolution of social structure is not a balanced one. It appears that the structure of the social network would have more of an effect on the evolution of language than vice versa. Simple patterns of behaviour can account for the complex structure of social networks, yet embedding a language evolution model in a social network requires that it become more complex. The exact details of the learning algorithms are not crucial for the evolution of the social network, provided that the outcome of an interaction between agents is some measure of similarity. The structure of the network, however, would appear to heavily influence the languages that emerge.

The primary distinguishing properties of social networks - assortative mixing, transitivity and prominent community structure - need not be explicitly coded into the algorithm for them to reliably emerge. They arise as a result of individuals having a preference for those they are similar to, and a means to change their own traits to become more similar to those to whom they are connected.

An important result that language difference is being used as a proxy for social distance. Since agents with similar traits cluster together, the mechanism of meeting friends of friends used in other models to generate a transitive network is not necessary. In an evolved network, it is probable that two similar agents possessing similar traits belong to a similar group - i.e. are close together in the network. The reverse is true also - agents that are very dissimilar are probably further from each other.

5.3 Further work

There are three avenues of further work that may be undertaken - a deeper exploration of the parameters of the model; or an investigation into changing the mechanisms of the model of network evolution; or changing the learning algorithm for the purpose of exploring language evolution on a realistic social network.

5.3.1 *Exploring the parameters of the model*

There are several parameters of the model that were not specifically addressed in this study. Namely, what effect does population size N , maximum degree z^* and the number d_H and dimension d_h of agents' traits. These parameters may have non-trivial interactions, and may be worth exploring to further understand the behaviour of the model. It may also be worth exploring

5.3.2 *Language evolution*

Given that this model of evolving network structure is now available and that the evolution of a characteristic social network can proceed with simple parameters from the learning mechanism (i.e. measures of similarity), the next step is to consider the evolution of more complex language models under such conditions. For instance, an ILM of language learning and use could be implemented in place of the agents' current learning algorithm to determine whether it is still robust under such an evolving network structure. Results from Smith & Hurford (2003) suggest that this may not be the case, and the algorithms may need to be augmented with further mechanisms to be useful under such conditions.

The traits used in this model are discrete and independent of each other, as well as being equally functional. This is a reasonable assumption for many elements of language, illustrated by Chambers (2003): "The more deeply we inquire into the social meaning of language, the more clearly we see how arbitrary are the values that are commonly attached to it" (p.277). Of course there are many elements of language that are not arbitrary - the presence of linguistic universals rules this out - but many, such as particular realisations of phonemes are simply negotiated within the social exchange. It would be interesting to model differences in functionality of traits, and determine whether a functional language might emerge. It would also be interesting to evaluate the extent to which dysfunctional norms could arise and be maintained in an otherwise functional language through pressures arising from the social network structure. Also, continuous values could be used in place of discrete values for the traits.

5.3.3 *Extension to the preferential attachment model*

Obviously, homophily is not the only factor involved in the coevolution of language and social networks. Another factor involved in the formation of social

networks is the observation that an individual may actively try to be different than average, discussed in chapter 2.3.3. (Dittrich et al. 2000) describes the *seceder model*, whereby a communities of similar individuals evolve given a set of N individuals each with a random real number trait $s(i)$ and this preference to be different than the average. At each timestep, three individuals i_1, i_2 and i_3 are chosen at random. Of these individuals, the one whose s -value $s(i)$ is farthest from the average is chosen. The s -value of a fourth uniformly randomly chosen agent is replace with $s(i) + \eta$, where η is a random number between -1 and 1. (Gronlund & Holme 2004) adapts the model by embedding it in a network, whereby the s -value is discarded and replaced with implicit similarities based on an vertex's position in the network. The resulting evolved network exhibits high clustering and community structure, but interestingly only displays degree-correlation marginally above that of a random graph. Perhaps the desire to be different than the average needs to be balanced with the desire to maintain similarity with those to whom an individual is close to. The model of network evolution demonstrated here could be expanded to include a pressure to be different than the average, combined with the learning algorithm that maintains similarity with those to whom an individual is connected.

Another possible addition to the model based on homophily is the notion of *obligaton*. If there is some intrinsic value or cost to an interaction, whether either the speaker or hearer that benefits, this could create an obligation (or desire) for the other to return the signal. This creates an asymmetry in the network whereby pressures resulting from the obligations contracted from the network will influence individuals behaviour (Milroy 1987).

The timescale of the network could be increased to allow for the addition and removal of vertices, extending it to an intergenerational moel of language and network evolution. In addition, a genetic algorithm (GA) could be implemented to assess the fitness of the speakers and their probability of becoming cultural parents to a new generation.

The dynamics of the network could be altered to allow for other types of communication, such as having one speaker and multiple listeners (Gong et al. 2004). In addition, many models of language evolution distinguish between naïve learners and fully enculturated adults. The model could be extended to explore the effect this has on the network and language. Indeed, Kerswill & Williams (2000)

have found that adults, adolescents and children influence the outcome of language interactions differently.

5.4 Conclusions

This study has presented a model of the coevolution of language and social structure. It reproduces the characteristic measures of social networks - assortative mixing, transitivity and prominent community structure using an algorithm based on homophily. It provides a simple, though plausible mechanism for the evolution of social networks, and allows for expansion to implement and test existing language evolution theories such as the ILM in a more realistic social structure.

APPENDIX A

Diagrams of evolving networks

This appendix shows the main component of the evolving networks for the different implementations of the algorithm.

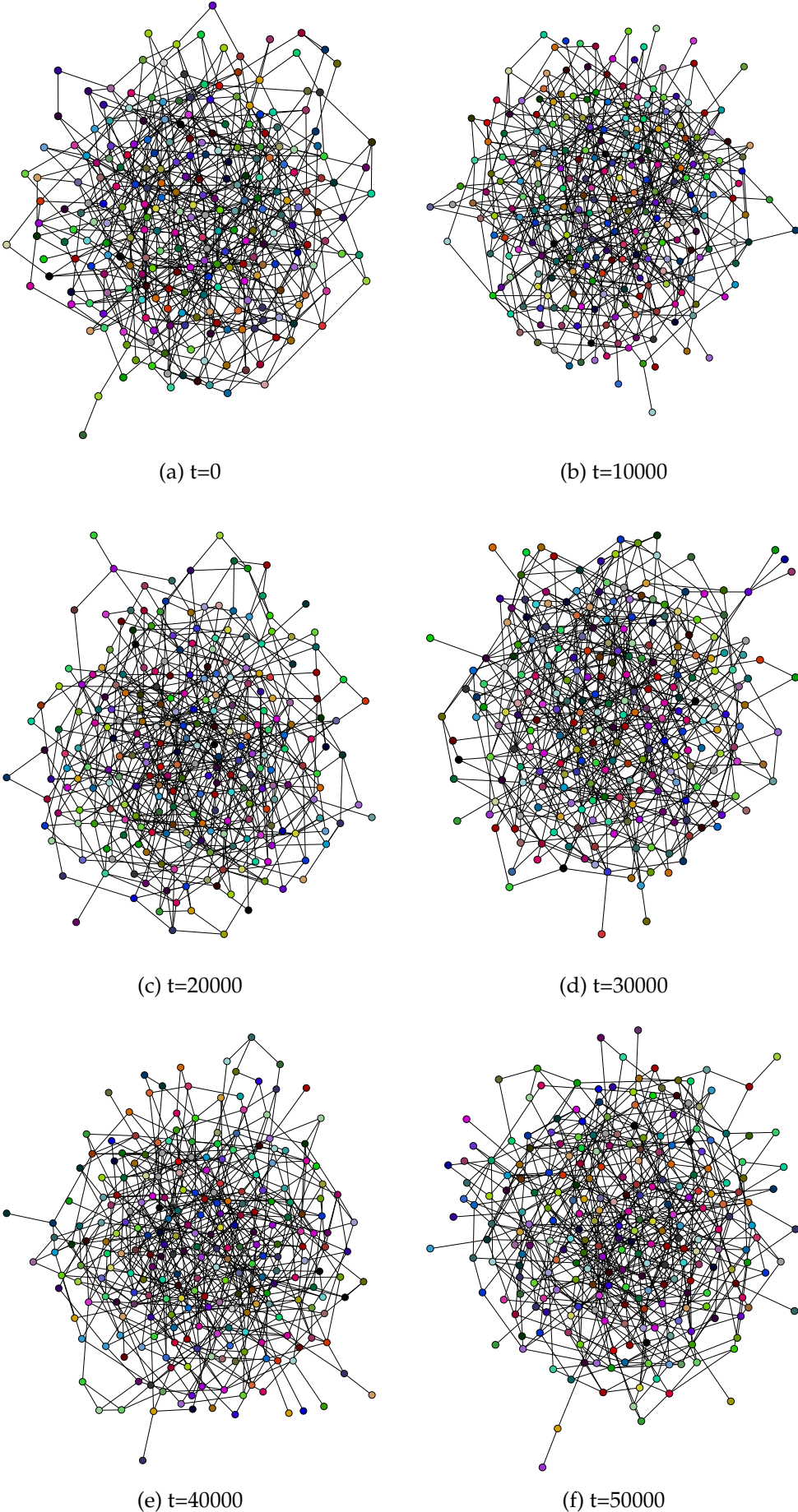


Figure A.1: Evolving network with degree preference

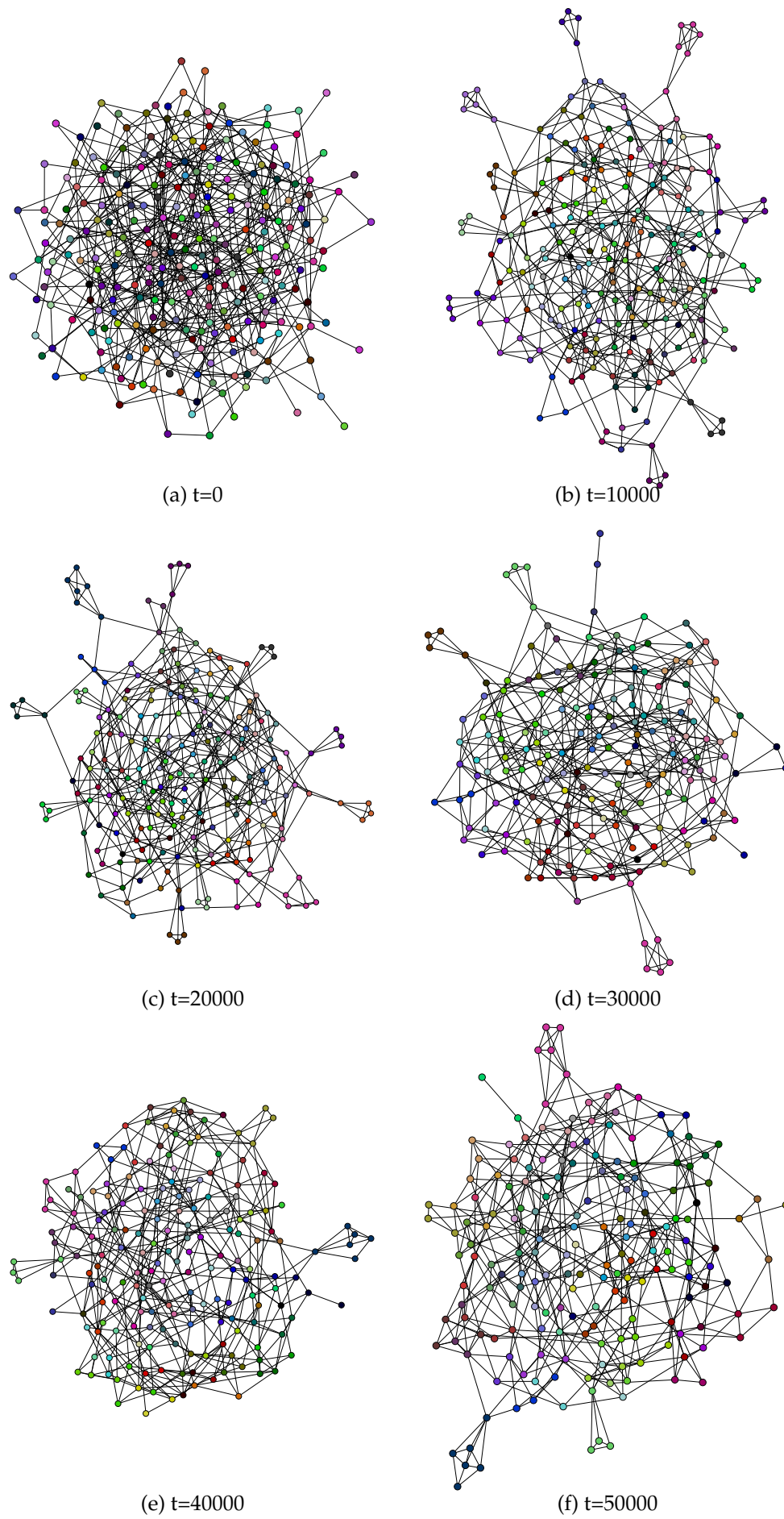


Figure A.2: Evolving network with trait preference

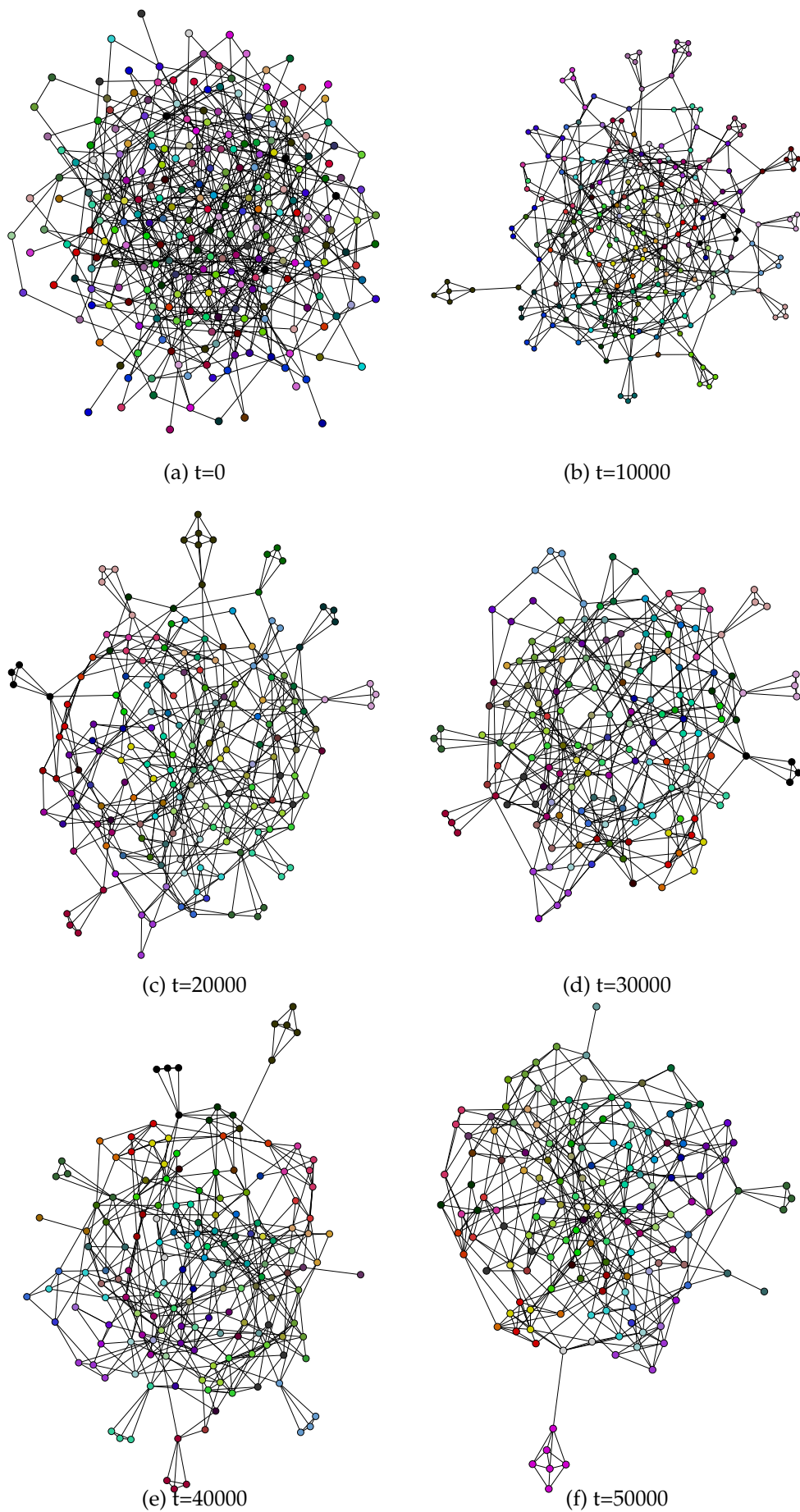


Figure A.3: Evolving network with trait and degree preference

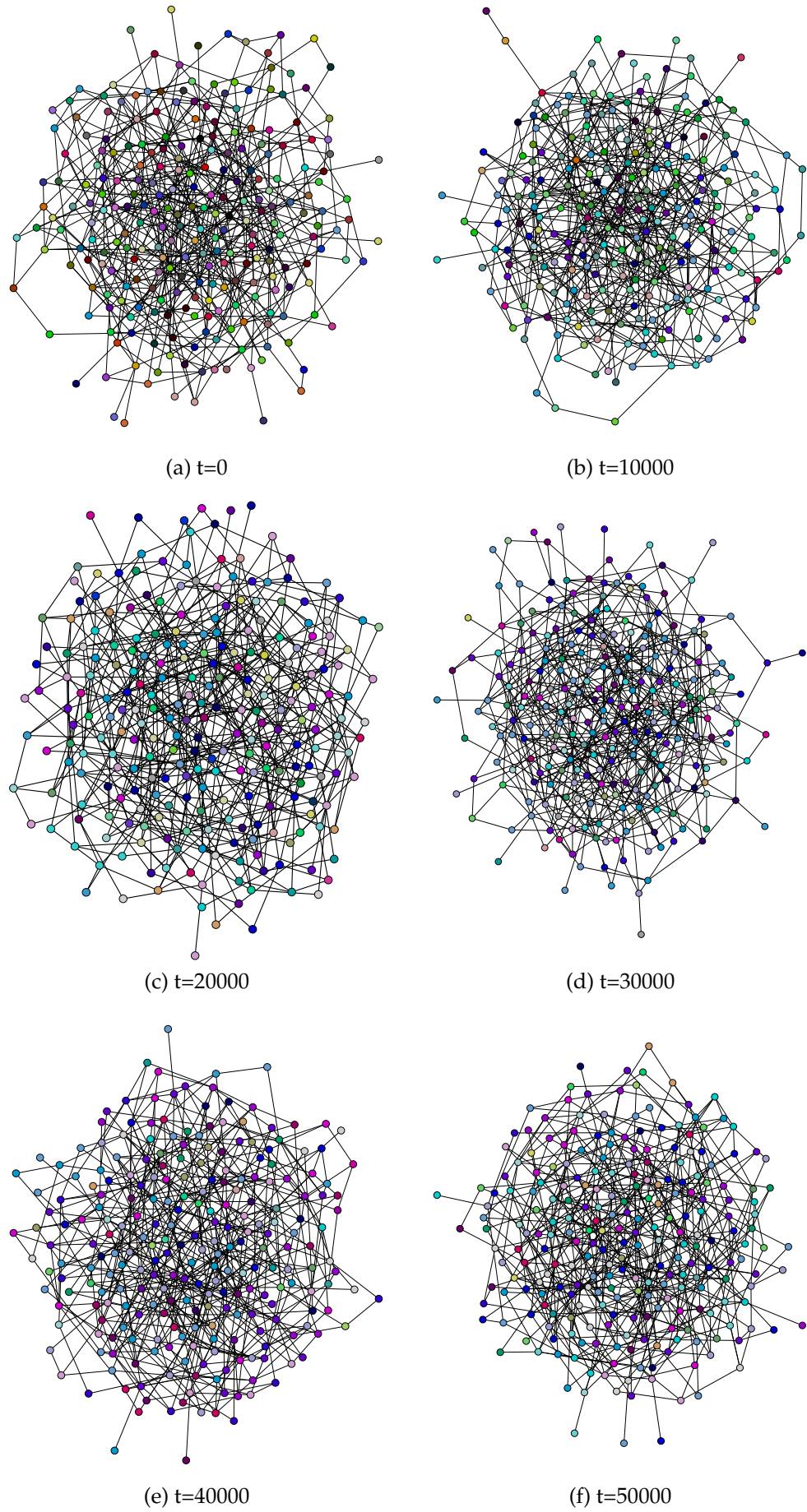


Figure A.4: Evolving network with learning

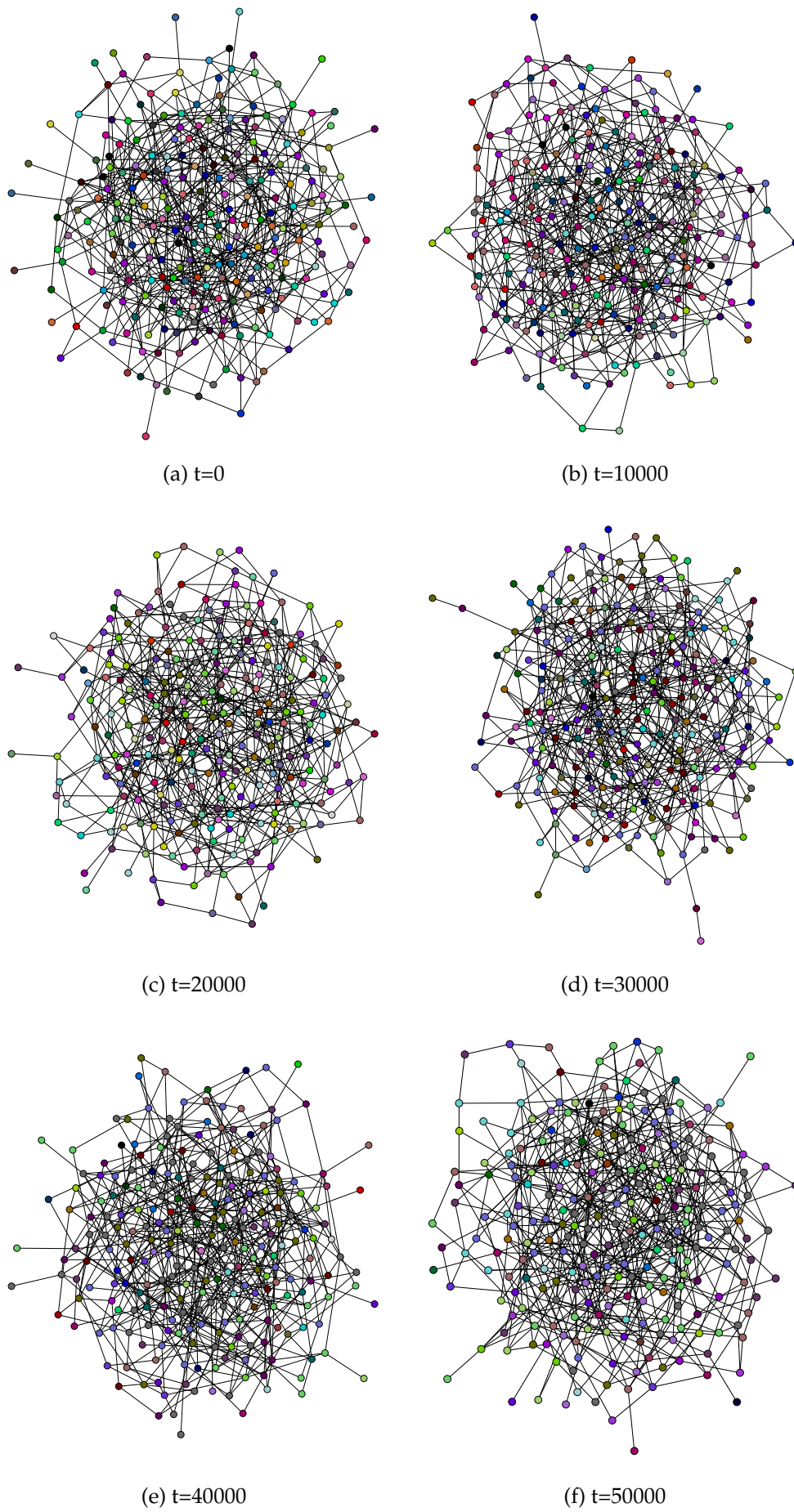


Figure A.5: Evolving network with learning and degree preference

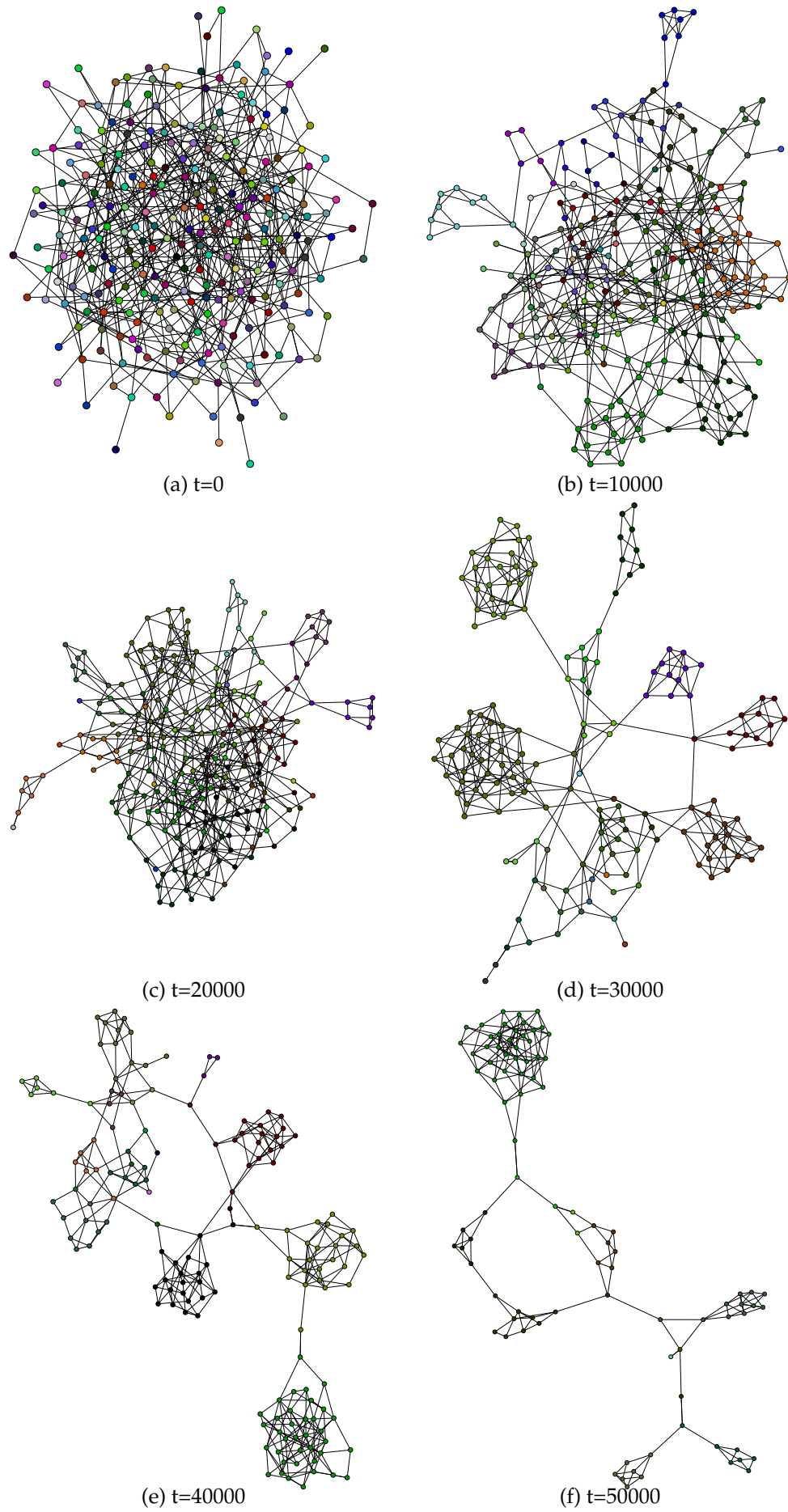


Figure A.6: Evolving network with learning and trait preference

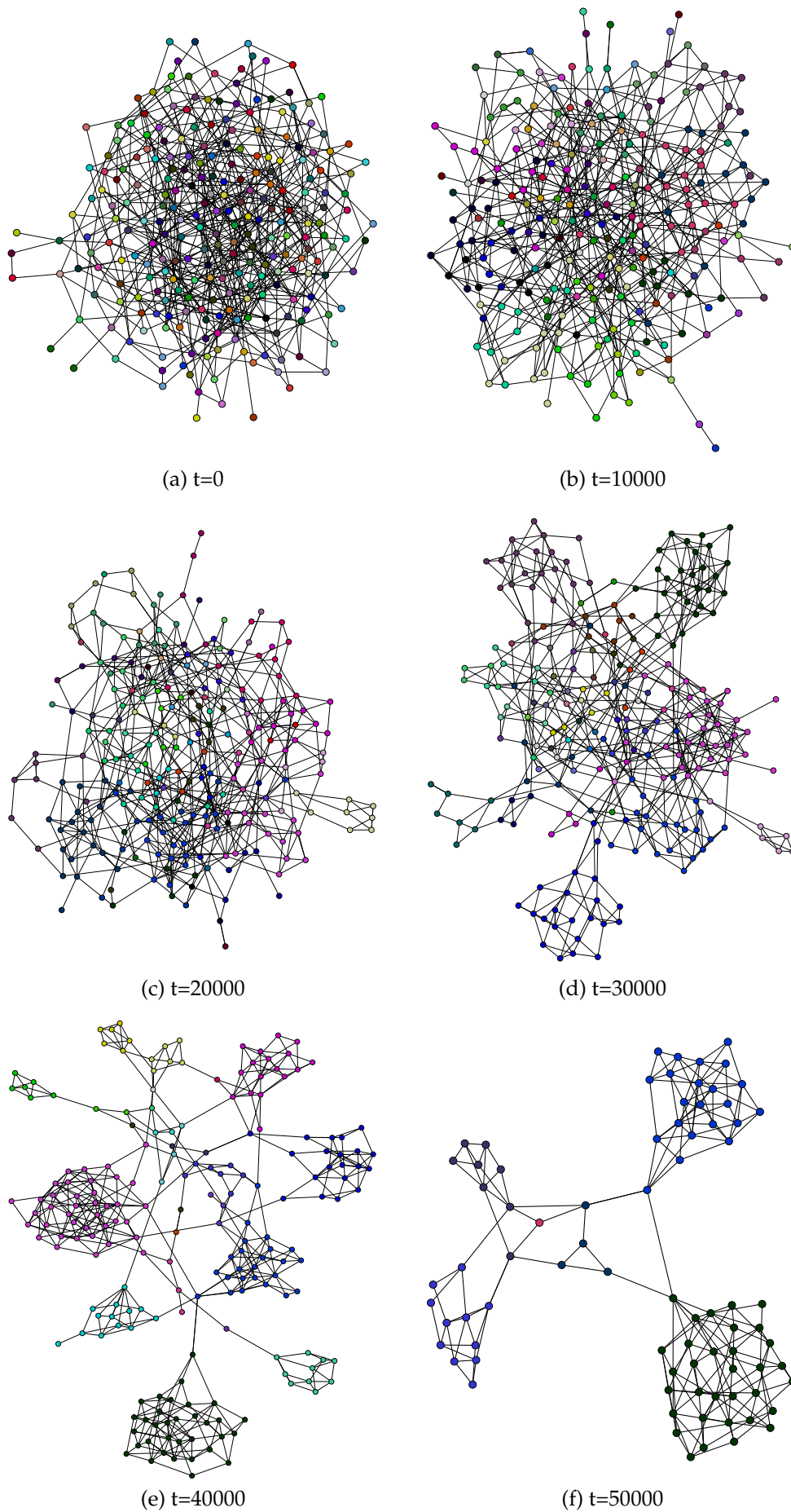


Figure A.7: Evolving network with learning, trait and degree preference

APPENDIX B

Statistical results

B.1 Tests of Between-Subjects Effects for Dependant Variable Clustering Coefficient

Source	Type III Sum of Sq's	df	Mean Square	F	Sig.
Corrected Model	1.450(a)	7	.207	101.888	.000
Intercept	1.529	1	1.529	751.921	.000
<i>LEARNING</i>	.079	1	.079	38.724	.000
<i>PREF</i>	1.271	1	1.271	624.929	.000
PREF_DEG	.001	1	.001	.634	.431
<i>LEARNING * PREF</i>	.085	1	.085	41.882	.000
LEARNING * PREF_DEG	.007	1	.007	3.463	.070
PREF * PREF_DEG	.001	1	.001	.479	.493
LEARNING * PREF * PREF_DEG	.006	1	.006	3.103	.086
Error	.081	40	.002		
Total	3.061	48			
Corrected Total	1.532	47			

a R Squared = .947 (Adjusted R Squared = .938) Significant effects are shown in italics.

Table B.1: Clustering coefficient statistics

B.2 Tests of Between-Subjects Effects for Dependant Variable Degree Correlation

Source	Type III Sum of Sq's	df	Mean Square	F	Sig.
Corrected Model	1.359(a)	7	.194	38.309	.000
Intercept	1.269	1	1.269	250.418	.000
LEARNING	.031	1	.031	6.203	.017
<i>PREF</i>	1.290	1	1.290	254.443	.000
PREF_DEG	.003	1	.003	.586	.449
LEARNING * PREF	.023	1	.023	4.483	.040
LEARNING * PREF_DEG	.002	1	.002	.386	.538
PREF * PREF_DEG	.002	1	.002	.444	.509
LEARNING * PREF * PREF_DEG	.008	1	.008	1.616	.211
Error	.203	40	.005		
Total	2.831	48			
Corrected Total	1.562	47			

a R Squared = .870 (Adjusted R Squared = .847)

Significant effects are shown in italics.

Table B.2: Degree correlation statistics

B.3 Tests of Between-Subjects Effects for Dependant Variable Language Distance Correlation

Source	Type III Sum of Sq's	df	Mean Square	F	Sig.
Corrected Model	4.209(a)	7	.601	154.953	.000
Intercept	4.299	1	4.299	1107.678	.000
LEARNING	.001	1	.001	.262	.612
<i>PREF</i>	4.203	1	4.203	1083.101	.000
PREF_DEG	.000	1	.000	.064	.801
LEARNING * PREF	.003	1	.003	.888	.352
LEARNING * PREF_DEG	.001	1	.001	.327	.571
PREF * PREF_DEG	7.227E-05	1	7.227E-05	.019	.892
LEARNING * PREF * PREF_DEG	2.795E-05	1	2.795E-05	.007	.933
Error	.155	40	.004		
Total	8.663	48			
Corrected Total	4.365	47			

a R Squared = .964 (Adjusted R Squared = .958) Significant effects are shown in italics.

Table B.3: Language distance and social distance correlation

References

- Barabási, A.-L. & Albert, R. (1999a), 'Emergence of scaling in random networks', *Science* **286**(509).
- Barabasi, A. L. & Albert, R. (1999b), 'Emergence of scaling in random networks', *Science* **286**, 509–512.
- Baronchelli, A., Loreto, V., Dall'Asta, L. & Barrat, A. (2006), Bootstrapping communication in language games. strategy, topology and all that, in A. Cangelosi, A. D. M. Smith & K. Smith, eds, 'The Evolution of Language: Proceedings of the 6th International Conference (EVOLANG6)', World Scientific Publishing Co. Pte. Ltd., Singapore, pp. 11–18.
- Bogart, L. (1951), 'The spread of news on a local event: A case history', *Public Opinion Quarterly* **14**(4), 769–772.
- Boguna, M., Pastor-Satorras, R., Diaz-Guilera, A. & Arenas, A. (2003), 'Emergence of clustering, correlations, and communities in a social network model'.
- Bourhis, R. Y. & Giles, H. (1977), The language of intergroup distinctiveness, in H. Giles, ed., 'Language, Ethnicity and Intergroup Relations', London: Academic Press, pp. 119–34.
- Brass, D. J. (1985), 'Mens and womens networks - a study of interaction patterns and influence in an organization', *Academy of Management Journal* **28**(2), 327–343.
- Chambers, J. K. (2003), *Sociolinguistic theory: linguistic variation and its social significance*, Blackwell, Oxford.
- Chomsky, N. (1986), *Knowledge of Language*, Praeger.
- Christiansen, M. H. (1994), *Infinite Languages, Finite Minds: Connectionism, Learning and Linguistic Structure*, PhD thesis, University of Edinburgh, Scotland.
- URL:** <http://www.isrl.uiuc.edu/amag/langev/paper/christiansen94phd.html>
- Deacon, T. (1997), *The Symbolic Species: The Co-Evolution of Language and the Brain*, Penguin.

- Dittrich, P., Liljeros, F., Soulier, A. & Banzhaf, W. (2000), 'Spontaneous group formation in the seceder model'.
URL: citeseer.ist.psu.edu/dittrich00spontaneous.html
- Eckert, P. (1988), 'Adolescent social structure and the spread of linguistic change', *Language and Society* **17**, 183–207.
- Faloutsos, M., Faloutsos, P. & Faloutsos, C. (1999), On power-law relationships of the internet topology, in 'SIGCOMM', pp. 251–262.
- Freeman, L. (1996), 'Some antecedents of social network analysis', *Connections* **19**, 39–42.
- Gardner, M. (1970), 'Mathematical games: The fantastic combinations of John Conway's new solitaire game "life"', *Scientific American* **223**, 120–123.
- Girvan, M. & Newman, M. E. J. (2002), 'Community structure in social and biological networks', *PROC.NATL.ACAD.SCI.USA* **99**, 7821.
URL: <http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0112110>
- Gong, T., Ke, J., Minett, J. W. & Wang, W. S.-Y. (2004), A computational framework to simulate the coevolution of language and social structure, in 'Ninth International Conference on the Simulation and Synthesis of Living Systems (ALIFE9)', Boston, pp. 158–163.
- Gronlund, A. & Holme, P. (2004), 'The networked seceder model: Group formation in social and economic systems', *Physical Review E* **70**, 036108.
URL: [doi:10.1103/PhysRevE.70.036108](https://doi.org/10.1103/PhysRevE.70.036108)
- Hurford, J. (2000), Social transmission favours linguistic generalization, in J. R. H. Chris Knight & M. Studdert-Kennedy, eds, 'The Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form', Cambridge University Press, Cambridge, pp. 324–352.
URL: <http://www.isrl.uiuc.edu/amag/langev/paper/hurford00socialTransmission.html>
- Jin, E. M., Girvan, M. & Newman, M. E. J. (2001), The structure of growing social networks, Working Papers 01-06-032, Santa Fe Institute.
- Keller, R. (1994), *On language change: The invisible hand in language*, Routledge, London.
URL: <http://www.isrl.uiuc.edu/amag/langev/paper/keller94book.html>
- Kerswill, P. & Williams, A. (2000), Creating a new town koine: Children and language change in Milton Keynes, in 'Language in Society', chapter 29, pp. 65–115.
- Kirby, S. (2000), Syntax without natural selection: How compositionality emerges from vocabulary in a population of learners, in C. Knight, ed., 'The

- Evolutionary Emergence of Language: Social Function and the Origins of Linguistic Form', Cambridge University Press, pp. 303–323.
URL: <http://www.isrl.uiuc.edu/amag/langev/paper/kirby00syntaxWithout.html>
- Kirby, S. (2002), Learning, bottlenecks and the evolution of recursive syntax, in T. Briscoe, ed., 'Linguistic Evolution through Language Acquisition: Formal and Computational Models', Cambridge University Press, chapter 6.
URL: <http://www.isrl.uiuc.edu/amag/langev/paper/kirby02learningBottlenecks.html>
- Kirby, S. & Hurford, J. R. (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, chapter 6, pp. 121–147.
- Kleinberg, J. M. (2000), 'Navigation in a small world', *Nature* **406**(845).
- Labov, W. (1972), *Sociolinguistic Patterns*, University of Pennsylvania Press, Philadelphia.
- Lepage, R. B. (1968), 'Problems of description in multilingual communities', *Transactions of the Philological Society* pp. 189–212.
- Li, X. & Chen, G. R. (2003), 'A local-world evolving network model', *Physica A* **328**(1-2), 274–286.
- Mäkinen, V. (n.d.), 'Himmeli: a free software package for visualising complex networks'. available at <http://www.artemis.kll.helsinki.fi/himmeli>.
- McPherson, M., Smith-Lovin, L. & Cook, J. M. (2001), 'Birds of a feather: Homophily in social networks', *Annual Review of Sociology* **27**, 415–444.
- Milroy, L. (1987), *Language and social networks*, Basil Blackwell, Oxford.
- Minett, J. W. & Wang, W. S. Y. (2004), Modeling endangered languages: the effects of bilingualism and social structure.
URL: <http://www.isrl.uiuc.edu/amag/langev/paper/minett04endangeredLanguage.html>
- Molloy, M. & Reed, B. (1995), 'A critical point for random graphs with a given degree sequence'.
- Nettle, D. (1999), *Linguistic Diversity*, Oxford University Press, New York, NY.
- Newman, M. (2003), 'The structure and function of complex networks'.
- Newman, M. E. J. (2002), 'Assortative mixing in networks', *Physical Review Letters* **89**, 208701.
URL: <http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0205405>
- Newman, M. E. J. & Girvan, M. (2003), 'Mixing patterns and community structure in networks'.
URL: <http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0210146>
- Newman, M. E. J. & Park, J. (2003), 'Why social networks are different from other types of networks', *Physical Review E* **68**.

- Oliphant, M. (1996), 'The dilemma of saussurean communication', *Biosystems* **37**(1-2), 31–38.
URL: <http://www.isrl.uiuc.edu/amag/langev/paper/oliphant96theDilemma.html>
- Park, J. & Newman, M. E. J. (2003), 'The origin of degree correlations in the internet and other networks', *Physical Review E* **68**, 026112.
URL: <http://www.citebase.org/abstract?id=oai:arXiv.org:cond-mat/0303327>
- Pinker, S. (1994), *The Language Instinct: How the Mind Creates Language*, Harper-Collins, New York.
- Pinker, S. & Bloom, P. (1990), 'Natural language and natural selection', *Behavioural and Brain Sciences* **13**(4), 707–784.
- Sapir, E. (1929), The status of linguistics as science, in 'Selected Writings of Edward Sapir', University of California Press, Berkeley, CA.
- Scott, J. (1991), *Social Network Analysis: A Handbook*, SAGE Publications, California.
- Shannon, C. (1948), 'A mathematical theory of communication', *Bell System Technical Journal* **27**, 379–423 and 623–656.
- Smith, K. & Hurford, J. (2003), Language evolution in populations: extending the Iterated Learning Model, in 'ECAL03', pp. 507–516.
- Steels, L. (2000), Language as a complex adaptive system, in M. Schoenauer, ed., 'Proceedings of PPSN VI', Lecture Notes in Computer Science, Springer-Verlag, Berlin, Germany.
URL: <http://www.isrl.uiuc.edu/amag/langev/paper/steels00languageAs.html>
- Toivonen, R., Onnela, J. P., Saramaki, J., Hyvonen, J. & Kaski, K. (2006), 'A model for social networks'.
- Travers, J. & Milgram, S. (1969), 'Identity and search in social networks', *Sociometry* **32**(4), 425–443.
- Wasserman, S. & Faust, K. (1994), *Social Network Analysis: Methods and Applications*, Cambridge University Press, Cambridge.
- Watts, D. J., Dodds, P. S. & Newman, M. E. J. (2002), 'Identity and search in social networks', *Science* **296**, 1302–1305.
- Watts, D. J. & Strogatz, S. H. (1998), 'Collective dynamics of 'small-world' networks.', *Nature* **393**(6684), 440–442.
URL: <http://dx.doi.org/10.1038/393440a0>