Social network structure, accommodation, and language change

Emily Clem
University of California, Berkeley

1 Introduction

One important tool of sociolinguistic analysis that emerged in second-wave sociolinguistics was the study of the role of social networks in the propagation of linguistic change. As studies investigated the role that network properties played in the diffusion and adoption of linguistic variants, they began to cast doubt on the view that prestige (either overt or covert) was the main factor influencing the dynamics of language change. Researchers such as James and Lesley Milroy (1985, 1992) suggested that diffusion patterns that could not be adequately described via traditional notions of prestige could perhaps be better understood in terms of the network properties of the communities in which they were distributed. However, one of the major issues faced by such studies is that actual social networks are difficult to study in detail due to their size and complexity.

Because of the difficulties associated with studying real social networks in language change situations, researchers have more recently attempted to further evaluate the findings of these earlier studies and to explore the role of network properties in language change through the use of artificial networks. The advantage of a computational model of social network structure is that it allows the researcher to manipulate and control various aspects of the network structure. However, the inherent disadvantage of such models is that they require the researcher to make and implement various assumptions about how language change proceeds and how speaker interactions are conducted and evaluated. Therefore, while some studies have been able to produce results that mirror the predictions of sociolinguistic studies (Fagyal et al. 2010), others have shown that the priors of the model, rather than any population dynamics within the model, can be the biggest factor that influences the outcome (Kirby and Sonderegger 2013).

In this paper, I investigate whether the structure of a social network necessarily predicts certain diffusion dynamics within that network. Building on the methodology used by Fagyal et al. (2010), I explore to what degree the assumptions of the model affect the propagation of changes and compare this to the degree of influence the structure of the network has. By implementing more sophisticated models of linguistic variables and speaker accommodation, I show that network structure affects the rate of change more than the qualitative outcome. Since the role of network structure compared to other social and linguistic factors involved in language change is not well understood, the findings of this paper cast doubt on the assumption that certain network structures are the key component, or at least a necessary component, in instances of linguistic change. This serves to call into question overly deterministic views of language change based on network properties, echoing the criticisms of scholars such as Holmes and Kerswill (2008) and Coupland (2008).

In the following sections I explore the proposed role of network structure on change and demonstrate the effects of different modifications on a model of social networks and language change. In §2, I give a brief overview of one proposal of how network structure facilitates change. In §3, I discuss the basic properties of the model I use along with the structure of the three types of networks that I explore. The results for an incremental model of change are given in §4 with three slightly different versions of the model. In §5 I explore a discrete but probabilistic model of change involving speakers with command of multiple variants. I discuss the general outcomes and implications of these trials in §6 and offer concluding remarks in §7.

2 Background

In the study of social networks and language change, one property of networks that has played a central role is the importance of weak ties. This idea was pioneered in the field of sociology by Granovetter (1973), and was later borrowed into linguistics. Granovetter (1973) proposed that weak interpersonal ties were necessary for the spread of changes within a network, because they form "bridges" between points in a network that otherwise would be connected only by longer paths of transmission. He claimed that individuals with many weak ties are in the best position to facilitate diffusion, and that they can serve to link members of different groups, allowing changes to be spread. This research also suggested that networks with dense, strong ties result in local cohesion within small groups, but lead to fragmentation in the larger structure.

Ideas about the role of social status and position within a network had been appealed to in linguistics in studies such as Labov's (1973) early work on Black Vernacular English. In his terminology, a "lame" would be an individual with weak ties to a network. Labov also remarked on the role of contacts both within and outside of a neighborhood in his Philadelphia neighborhood study (1980). However, the view of network structure pioneered by Granovetter (1973) was not adopted explicitly within linguistics to examine the spread of linguistic variants in a community until the work of Milroy and Milroy (1985). Through their work in Belfast, they argued that the same view of the importance of weak ties could be extended to linguistic change. They claimed that networks that contain individuals with many weak ties facilitate rapid change, while communities with mostly dense networks and strong ties lead to norm enforcement and permit only slow change (1985, 1992). Within linguistics, the role of network structure was also investigated by Lippi-Green (1989), who examined how network integration correlated with linguistic conservativity in a small German-speaking community in western Austria. She demonstrated that males who were most integrated maintained more conservative linguistic variants, supporting the hypothesis that a network composed of dense, strong ties facilitates norm enforcement rather than change. While these are perhaps some of the most explicit cases of network structure being implicated in patterns of change, the role of population dynamics in linguistic change has continued to be examined in research such as Trudgill et al.'s (2000) study of New Zealand English.

One challenge faced by authors attempting to find a correlation between network properties and the spread of linguistic innovations lies in the inherent complexity of social networks and interactions. Because mapping each interaction within a population of speakers is not a realistic goal, researchers have resorted to other methods of measuring network density and the strength of interpersonal ties. For example, Milroy (1980) and Lippi-Green (1989) used scores based on information collected about kinship, workplace, and voluntary associations as a metric for network strength. However, questionnaires about employment and involvement in social clubs serve only as an indirect measurement of the density and strength of ties within a network.

One reaction to this shortcoming in methodology and the inherent difficulty in studying actual networks has been to attempt to model the properties of social networks and to examine the patterns of diffusion that result from different network structures. Fagyal et al. (2010) represent one attempt to mirror the findings of Milroy and Milroy (1985) with a network model. They claim that "loners", members with only a few ties most of which are weak, and "hubs", members with many strong ties, are both necessary for the spread of linguistic innovations. They argue that a network without loners will not allow for innovation and that a network without hubs will not allow for a norm to be established. However, one issue with this model is that the implementation of the transmission of linguistic variables is not very sophisticated. When a speaker encounters a new variable, they immediately adopt this variable, completely losing the variable that they had before. Other work in modeling linguistic change, such as Kirby and Sonderegger (2013) has suggested that the priors of the model affect the qualitative outcome more than the population dynamics within the model. Therefore, while the findings of Fagyal et al. (2010) may be attractive, the question remains as to how much the outcomes of this model are affected by assumptions about language use and linguistic variables that do not very closely approximate reality. Ultimately, I will demonstrate that such a deterministic view of language change that relies so heavily on network structure is likely an oversimplification of the many complex properties that interact in language change.

3 The model

3.1 Generating the networks

To model social network properties I draw significantly on the model used by Fagyal et al. (2010). This decision was made in order to maintain as much similarity between the starting points of the two studies as possible. By doing this, it is clearer what the effects of the type of variable transmission assumed in the current simulations are compared to the effects of the network structure.

A closed influence network was constructed as a Recursive Matrix (R-MAT) as outlined by Chakrabarti et al. (2004). To generate an R-MAT graph, edges are placed in a network consisting of a given number of nodes. The location in the matrix where an edge is placed is probabilistically determined. The matrix is divided into four quadrants, each with a set probability of being chosen. Once a quadrant is chosen, it too is divided into quadrants, with one of its quadrants being probabilistically selected. This process continues recursively until one point in the matrix is chosen for the placement of an edge. In this model, the R-MAT graph was generated in Python 2.7 using the Snap.py tool (Leskovec and Sosič 2014). The probabilities of each of the quadrants being chosen were as follows: 0.5, 0.1, 0.1, 0.3. [1]

Within this R-MAT graph, different nodes have different densities of ties to other nodes. That is, some nodes have many ties to other nodes, while some nodes are relatively isolated, having few ties to other nodes. In this model, the ties between nodes are asymmetric. For any given node (x) with a tie to another node (y), an in-degree tie represents a flow of influence from x to y, while

[1] For the sparse network, the probabilities were equal for all quadrants, as discussed below.

an out-degree tie represents a flow of influence from y to x. Therefore, for the purpose of this study, the in-degree (count of in-degree ties) of a node represents how many other nodes can be influenced by that node. The out-degree (count of out-degree ties) of a node, on the other hand, represents the number of other nodes that can influence a node. The out-degree of a node, then, will be equivalent to the number of "neighbors" that a node has, for the purpose of determining possible sources of influence. The in-degree of a node, however, will be a measure of a node's social influence in the network.

All of the networks discussed here began with 900 nodes, but the number of edges or links between them varied by trial. For each of the four conditions, three trials were run, with each trial representing a different network type. The first trial was run with 7,500 unique edges. In this network, there were both loners and hubs. Loners were at the peripheries of the network, and were not influenced by other nodes in the network. Hubs, on the other hand, were nodes that had many ties and influenced many other nodes. The maximum in-degree for a hub in this network was 47, and the maximum out-degree was 60. Because influencers were probabilistically chosen based on in-degree, as discussed below, these hubs had a great potential to influence many other nodes. Fagyal et al. (2010) observe that a network with similar composition can be used to demonstrate the effects of scale-free small-world properties (Watts and Strogatz 1998) on the diffusion of linguistic variants. The relevant properties are small diameter, high clustering, and scale-free degree distribution. Small diameter refers to the fact that there is a low degree of separation between nodes. High clustering refers to the fact that if two nodes are linked to a third node, they are likely to be linked to each other as well, as discussed by Granovetter (1973). Finally, scale-free degree distribution refers to the fact that in this type of network there are relatively many nodes with low in-degrees and relatively few nodes with high in-degrees.

A second trial was run for each condition with 22,500 unique edges. This created a dense, multiplex network where there were few loners. The maximum in-degree for a node in this network was 98, and the maximum out-degree was 116. This means that many of the nodes had ties to a large proportion (over 10%) of the other nodes in the network. Fagyal et al. (2010) show that for a network that lacks loners, less frequent variants disappear until only one variant remains. This rapid convergence on one stable variant demonstrates the role of dense, multiplex networks in maintenance and norm-enforcement and reinforces claims about how such networks can often be resistant to change (Milroy and Milroy 1992). This is in contrast to a network with loners, where these nodes with zero out-degree cannot be influenced by other nodes and therefore always retain their variant. This allows for infrequent variants to survive and eventually re-enter the rest of the network. This pattern is in line with claims that have been made about the role of individuals with weak ties to a network in the process of diffusion of innovations (Granovetter 1973).

Finally, for each condition a third, slightly different, trial was run. Beginning again with 900 nodes, a network was created with fewer hubs. To do this, edges were placed randomly, with equal probabilities for each quadrant in the matrix. Additionally, only 3,000 unique edges were generated. The maximum in-degree for this network was only 15, and the maximum out-degree was only 19. This network then lacked any individuals that exerted influence over a large portion of the network. In the simulations of Fagyal et al. (2010) a random, sparse network without hubs did not allow for a norm to emerge. Though looser networks have been claimed to facilitate change (Milroy and Milroy 1985), without hubs no one variant was able to dominate a majority of the network.

All three types of networks examined here, then, were shown to result in radically different outcomes in the model of Fagyal et al. (2010). I will demonstrate, however, that by manipulating the way that variants are represented and transmitted, such extreme effects of network structure do not hold.

3.2 Modeling linguistic variables

The specific properties of variant transmission and adoption varied by trial. These specifics will be discussed further in §4 and §5. However, some basic properties remained constant across trials.

To model the presence and diffusion of linguistic variants in the network, each node was randomly assigned a value, representing a variant, at the start of a trial. The number of possible variants in the network varied by condition.

The value update process also varied across conditions, but some properties were held constant. At each time step, a single node (x) in the network was uniformly randomly chosen. The neighbors of this node were defined as those to which it had a direct link through out-degree links. From this set of neighbors, a single node (y) was chosen for x to potentially be influenced by. Following the choice of Fagyal et al. (2010), y was chosen using the in-degree biased voter model. This model is intended to approximate the effects of the social influence of speakers on others in the community. For each neighbor of x, the in-degree of that neighbor was calculated. This in-degree was then divided by the sum of all in-degrees of all of the neighbors of x to arrive at a probability that that neighbor would be chosen. From the set of neighbors, y was then probabilistically chosen as a potential influencer of x. After this influencer was chosen, the actual effects of its value on the value of x varied by condition. The conditions are discussed individually in §4 and §5.

4 Condition 1 - incremental change

Linguistic variants were first represented as incremental changes along a scale of values. In this implementation of the model, each node was randomly assigned one of eight variants (0-7). The values 0 and 7 represented two ends of a scale, and the values in between represented intermediate points on that scale. This scalar relationship was used to capture non-arbitrary relatedness of linguistic variants. This scale of values could linguistically represent a phonetic continuum of realizations, such as a more or less fronted pronunciation of a vowel or a shorter or longer VOT. In the variant update process, rather than a node changing its value to that of its neighbor, it can move in the direction of its neighbor along this continuum. This scale, therefore, allows for the modeling of gradient changes, rather than immediate categorical changes.

Three slightly different conditions were run with this incremental model of the variant update process. Across conditions, for each set of trials utilizing the same network structure (normal, dense, or sparse), each node in the network began with the same value as it had in each of the other conditions at the start of every trial. This was kept uniform so as to control for possible effects of value distribution on the results.

4.1 Condition 1A - basic accommodation

For the most simple trial, the variant update process proceeded in a uniform and scalar fashion. Once an influencer (y) was chosen for node x at time t, the value of node y was compared to the value of x. If the two nodes shared a value, nothing happened. However, if the value of y was greater or less than the value of x, x's value moved in the direction of the value of y by one point on the scale. Thus if y had a value of 5 and x had a value of 2, rather than x changing completely to 5, the value of x was increased to 3.

The analogue of this in actual linguistic interactions would be partial accommodation. An example of this would be if speaker x increased their VOT in an interaction with speaker y who had a longer VOT, but speaker x did not increase their VOT enough to match speaker y entirely. As Giles et al. (1991) claim, convergence can often be partial. That is, a speaker can accommodate to an intermediate point between themself and their interlocutor rather than converging fully to their interlocutor. Likewise, they claim that the extent of convergence can vary based on certain sociodemographic variables. Such views of accommodation as not representing actual and total convergence of participants are also consistent with an identity-projection model of accommodation (Auer and Hinskens 2005). Rather than converging to their interlocutor, it has been proposed that speakers accommodate instead to a stereotyped interlocutor. Because of this, their speech may not become equivalent to that of their interlocutor.

Taking these factors into consideration, it seems that a model such as the one proposed here which relies on incremental change towards an "interlocutor" is perhaps a closer approximation of accommodation in interactions than the categorical change implemented by Fagyal et al. (2010). Figure 1 shows the effect of such a model of value-updating in a network with both hubs and loners, that is, the network with 7,500 links.[2]

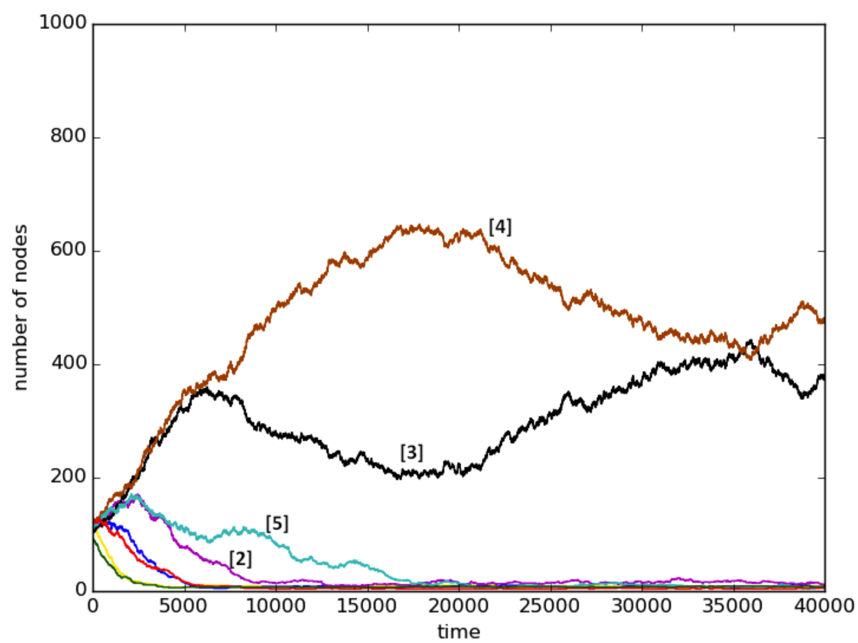

Figure 1: Basic Accommodation

[2]In all of the figures in §4.1 through §4.3, the following colors correspond to the following values: yellow = 0, blue = 1, magenta = 2, black = 3, brown = 4, cyan = 5, red = 6, green = 7.

As can be seen in the figure, most of the variants quickly decrease in frequency until most of the network converges on two adjacent intermediate values, 3 (black) and 4 (brown). Though these two adjacent values oscillate, there are no more increases in any more extreme variants after this point of converge is reached.

This behavior is similar to that outlined in Kerswill and Williams (2005). They show that for dialect leveling situations, networks that begin with diverse inputs will eventually converge on a new norm. In many cases, this norm will be a form of "compromise" between the competing variants. In this model, as in a koineization situation, the simulation is initiated with a variety of values. Eventually, the network converges on an intermediate (compromise) variant.

We must now examine how network structure affects this type of convergence. In Figure 2, results for a dense network are shown. In this trial, we see very similar results to those shown in Figure 1. The network quickly converges on two adjacent intermediate values, with all other values disappearing nearly entirely in this network that lacks loners. The one main difference in this network is that one of the two intermediate values gradually wins out over time. This enforcement of a single norm is to be expected based on the usual behavior of dense, multiplex networks (Milroy and Milroy 1985).

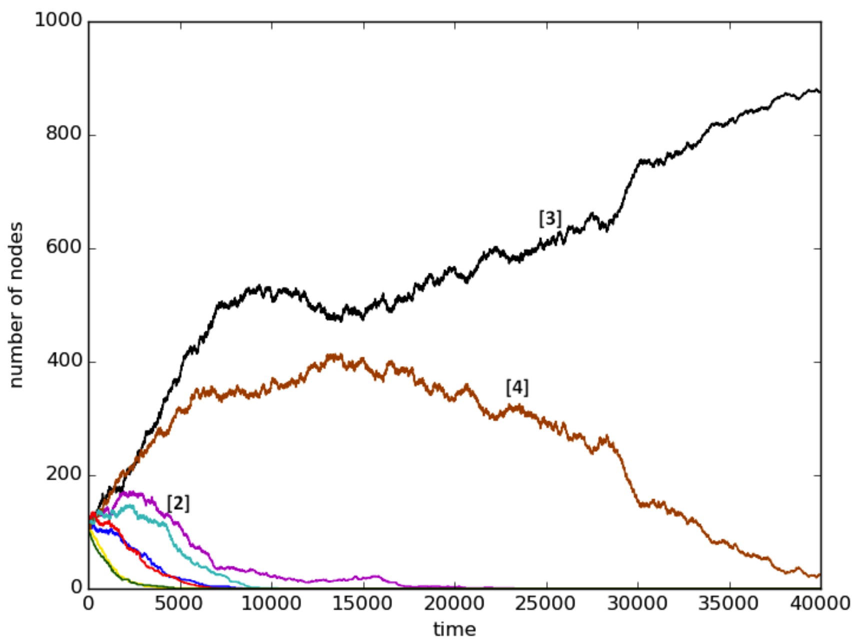

Figure 2: Basic Accommodation - dense network

Finally, we turn to the third trial in which we see the effects of a random and sparse network that lacks hubs. These results are shown in Figure 3. In this network, we see a similar picture as before with the two intermediate values quickly becoming the most frequent and remaining so over time. However, a difference is that the next two least extreme values, 2 (magenta) and 5 (cyan), do not disappear entirely. This seems to represent a convergence on an intermediate form, but with a broader range. It seems that without hubs, the effects of convergence, though similar, are less

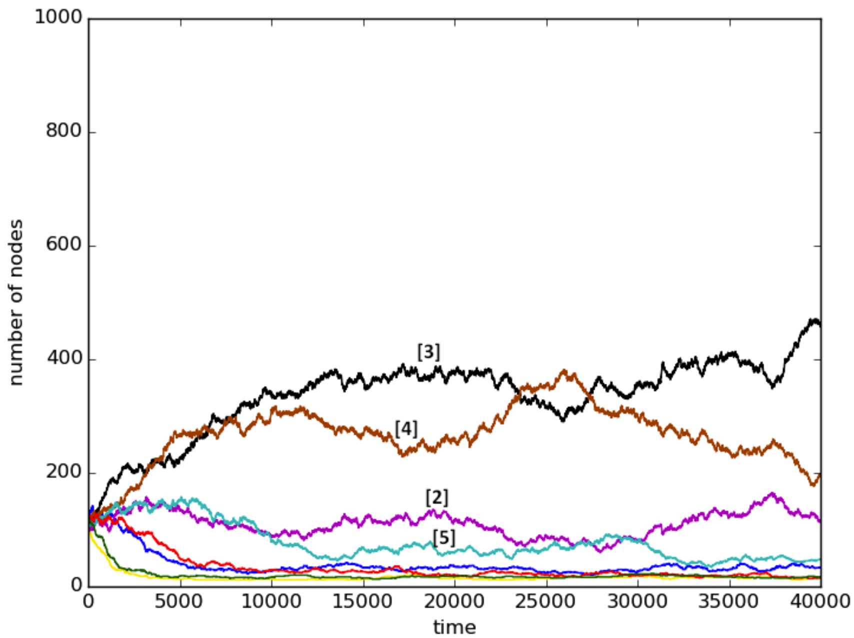

Figure 3: Basic Accommodation - sparse network

extreme. That is, the overall trend is towards the same intermediate values, but the network does not converge on a single value or pair of values to the exclusion of all others.

Taking the results of all three trials together, the convergence on and maintenance of an intermediate value or small set of values, seems to fall out directly from the more incremental, scalar nature of change or "accommodation." There are some effects of network on the degree of "focusing" in the sense of Trudgill et al. (2000) with the density of the network correlating with the degree of focusing. Likewise, the rate of convergence seems to be affected by network type, with the density of the network also correlating positively with the rate of convergence. However the same qualitative result of the intermediate values winning out is achieved regardless of the properties of the network.

4.2 Condition 1B - frequency effects

As mentioned in the previous section, the results of this type of model of accommodation approximate the outcome of some cases of new dialect formation. Namely, the network converges on a compromise form. In their discussion of new dialect formation, Trudgill et al. (2000) claim that variants that occur with higher frequency will win out in the context of dialect formation. In this model, the network starts with a random distribution of values on a scale, much like a community may start with a heterogeneous mixture of variants in a new dialect formation setting. A question of interest, then, is whether the results seen in the previous section will change if the overall distribution of variants in the network affects the accommodation process. According to the predictions of Trudgill et al.'s (2000) third-stage leveling, we would expect that the majority form would win out in such circumstances.

For this condition, the variant update process was adjusted so that it was not obligatory for one node to update its value at each time step. Instead, an influencer (y) and influencee (x) were chosen at each step, as before. However, whether x actually accommodated towards y by one step was probabilistically determined. This probability was calculated based on the mean value for the variant of all nodes in the network, with more extreme means leading to more extreme probabilities. The probability was then adjusted based on whether the direction of accommodation of x to y would be moving x's value closer to or further from the mean value in the network. Node x then could accommodate to y by moving one step closer on the scale, or x could retain its value with the choice based on this probability.

The results for this set of adjustments in a network with both hubs and loners can be seen in Figure 4. As we can see, the network converged on an intermediate value even faster than before, leveling off after only 5,000 time steps. Nearly identical results are seen in Figure 5 with a dense network and Figure 6 with a sparse network. The only difference between these networks is how many nodes end up with the intermediate value. This is based on how many loners there are in the network that can retain their own variant.

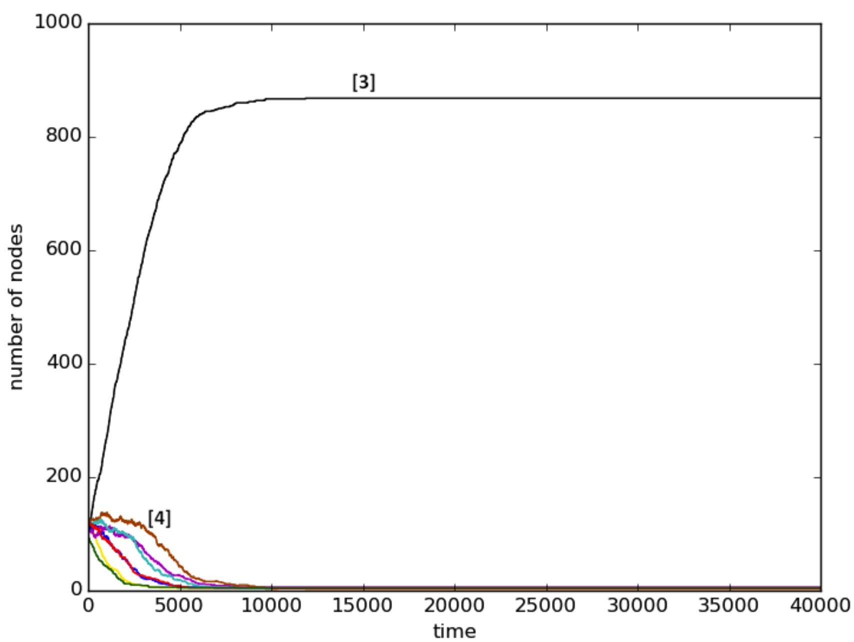

Figure 4: Frequency Effects

These results show focusing that is relatively independent of network type. They demonstrate that when the overall tendency of the network affects the variant update process, not only do variants still converge on an intermediate value, but they also do so more rapidly and more completely than in the first condition. Therefore, we see that if a variant's status as the most frequent variant is allowed to affect the probability of a node adopting the variant, this variant will win out very quickly.

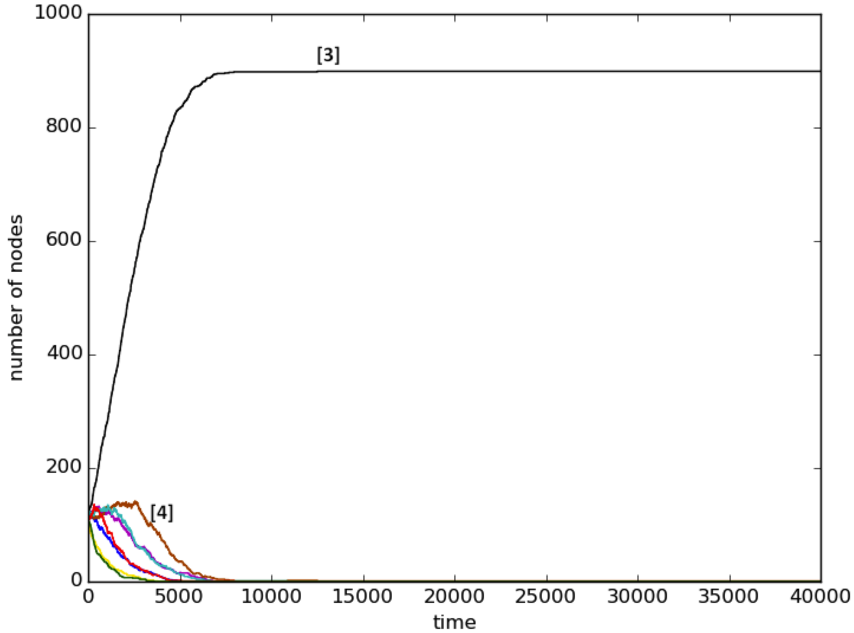

Figure 5: Frequency Effects - dense network

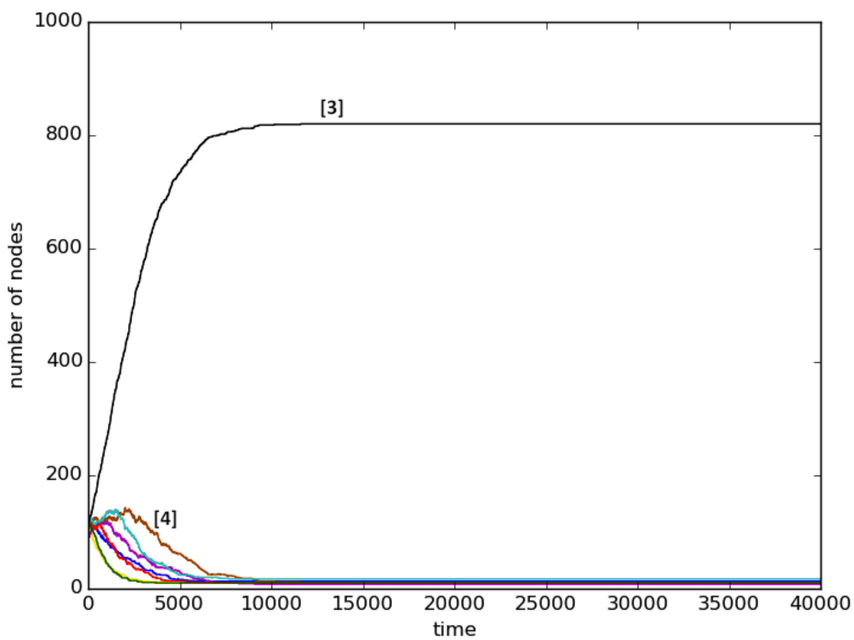

Figure 6: Frequency Effects - sparse network

4.3 Condition 1C - variant markedness

Up until this point, the nature of the model has assumed that nodes are equally as likely to accommodate to each value on the scale from a language-internal standpoint. However, in real speech communities linguistic factors may make certain variables more "attractive" than others. This is independent of the network structure of the individuals who use the variants.

In order to incorporate this concept into the model, accommodation was weighted more heavily towards one end of the value scale. More precisely, once a node selected which neighbor to copy, if the direction of resulting accommodation would be towards the higher end of the scale, the node increased its value by 2. However, if the direction of accommodation would be towards the lower end of the scale, the node only decreased its value by 1.

This weighting could be conceptualized of as a variety of different factors in actual language change. One possibility is that it could reflect a language-internal bias towards a particular variant due to markedness or ease of production or perception. In cross-linguistic tendencies of sound change, we often see the effects of markedness or phonetic "naturalness." This results in asymmetries in the types of patterns and processes we find crosslinguistically. For example, Garrett and Johnson (2013) reference many instances of phonetically natural sound change, such as velar palatalization before front vowels or vowel rounding harmony, but there are no attested examples of similarly imaginable but unnatural changes, such as velars backing before front vowels or vowel length harmony. Similarly, they note asymmetries in the direction of sound changes, with changes such as intervocalic stop voicing being very common, but intervocalic stop devoicing being unattested. Thus, in language change we expect that linguistic factors like phonetic naturalness influence the direction of change to reduce articulatory effort or to increase perceptual cues.

Another possibility is that the weighting in this model could represent some bias due to the relationship of the language systems that are in contact. As Anderson (1988) points out, the result of dialect mixing when the competing variants are not ascribed particular social values by the community, will be change towards the simpler of the two systems. This is like the situation seen with Kashubian and Polish vowel insertion rules, where speakers are more successful at acquiring the simpler Polish rule and shift has happened in this direction (Anderson 1988).

With this understanding of the motivations for this adjustment to the model, we turn now to the results. Figure 7 shows the results for the network with hubs and loners. In this trial, the network quickly converges on the two values at the higher end of the scale. This end represents the less marked variant, or simpler system. Once this end of the scale becomes dominant, the nodes oscillate between the two highest values, due to the fact that the accommodation process continues.

For the dense network we observe a similar pattern, as seen in Figure 8. The network quickly becomes dominated by the two highest values. However in this network, the highest value begins to clearly win out starting around time step 20,000. This could be due to the fact that there are no loners to preserve any lower values in this network.

Finally, in the sparse network, we see a slightly different pattern. The network once again converges on the less marked end of the scale. However, in this network, it is the three highest values that remain common in the network. This demonstrates that with a sparser network the results of focusing are less extreme.

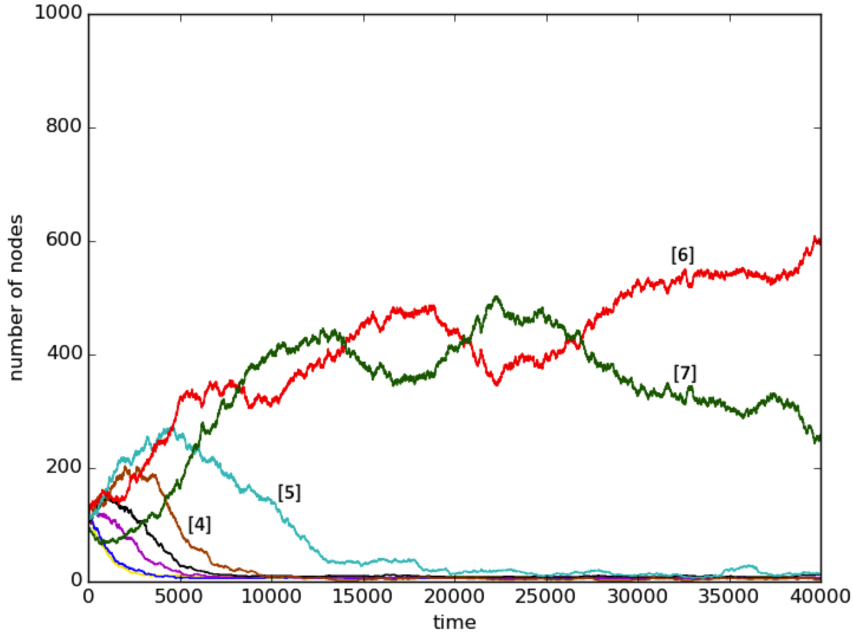

Figure 7: Variant Markedness

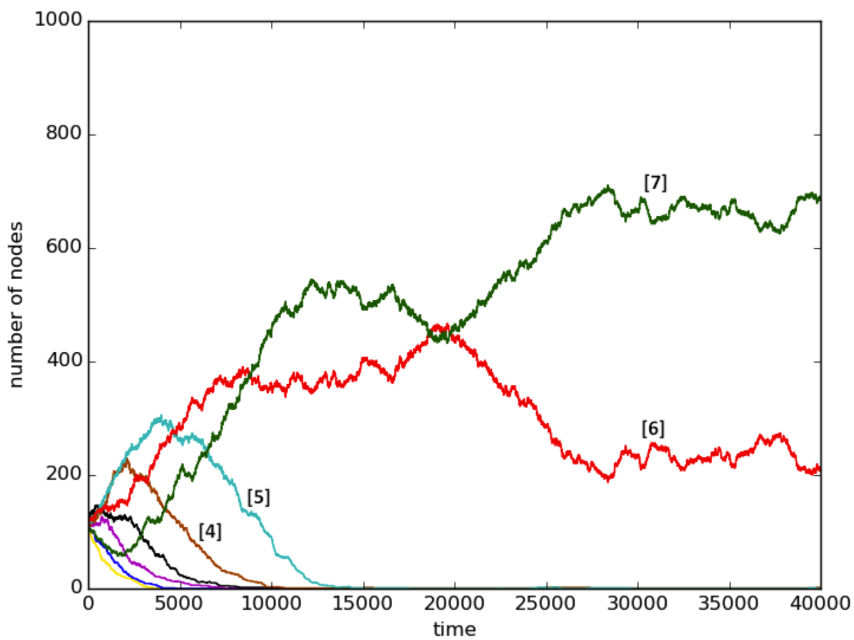

Figure 8: Variant Markedness - dense network

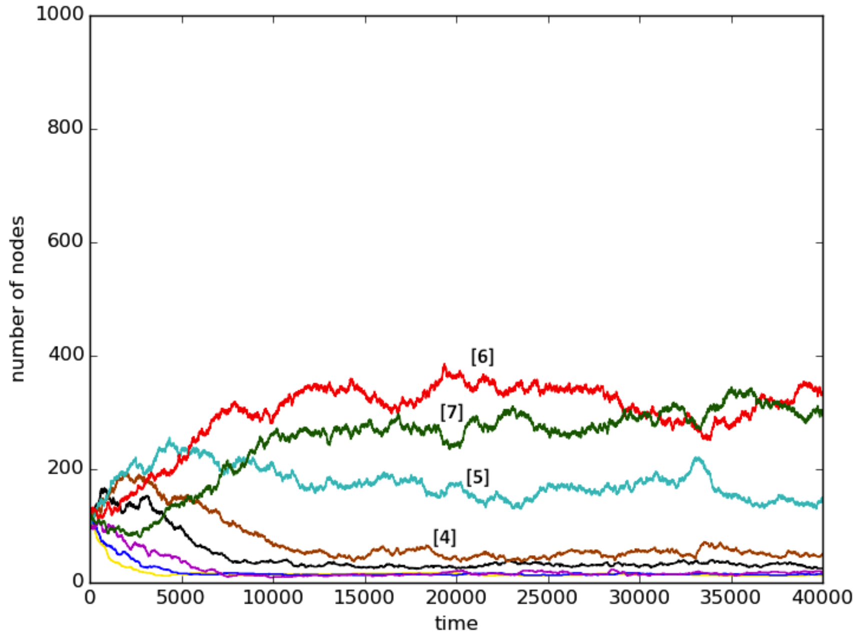

Figure 9: Variant Markedness - sparse network

The results of this condition are in line with Trudgill et al.'s (2000) claims about third-stage leveling patterns in dialect mixing situations. Just as third-stage leveling may result in the survival of majority forms, as seen in §4.2, third-stage leveling can also favor unmarked forms. In this condition, we see similar results for all three network types with the unmarked forms qualitatively winning out in all three trials. The difference between network types, is, once again, not the forms that win out, but the degree of focusing and the rate of convergence, with network density positively correlated with both of these.

5 Condition 2 - saltatory change

The three previous iterations of the model have relied on an incremental view of linguistic variants, with values related along a scale, just as linguistic variants may be related along a continuum of realizations. However, not all language change is along a continuum of gradient realizations. Most changes are instead saltatory in nature, with no intermediate variants.[3] However, in such situations of change, speakers do not immediately shift categorically from one variant to another. Instead, we expect speakers to maintain both variants in their repertoire for some time, with one becoming gradually favored over the other.

Therefore, in this version of the model, each node was randomly assigned one of only two variants (0,1). Additionally, each node was associated with a certain probability of using each of

[3]This is true of some sound changes, but is especially the case for morphological and syntactic changes which are typically non-gradient.

the two variants in the network. At the start of each trial, all nodes began with a probability of 1.0 of using the variant they began with. In the update process, when an influencer (y) was chosen for node x, the variant that x "heard" from y in the interaction was determined probabilistically based on the probability associated with each variant in y's "grammar." Therefore, if y had a probability of 0.8 associated with variant 0 and a probability of 0.2 associated with variant 1, x had an 80% chance of hearing variant 0 in the interaction. The variant that x actually heard from y then shifted the probabilities in x's grammar by 0.2 in favor of the variant it had just heard. For example, if x already had a probability of 0.6 associated with variant 1 and heard variant 1 from y, the probability associated with variant 1 in x's grammar increased to 0.8. At each time step in the model, the mean probability of hearing each variant in the entire network was recorded as a measurement of the prevalence of each of the two variants in the network as a whole.

This type of saltatory change is similar to the type of linguistic change seen in the shift from apical [r] to dorsal [R] in Montréal French (Sankoff and Blondeau 2007), a sound change which also took place across Europe (Trautmann 1880; Chambers and Trudgill 1998). In this model, as in the Montréal study, speakers have one of two distinct variants, but vary in the percentage of use of each variant. Likewise, over time, the percentage of use of each variant changes. This type of saltatory, rather than gradient, change could also reflect the pattern of use of two distinct lexical items or grammatical constructions, where the choice is not along a continuum but is rather the selection of one of two (or more) competing discrete forms, both of which are present in a speaker's lexicon or grammar.

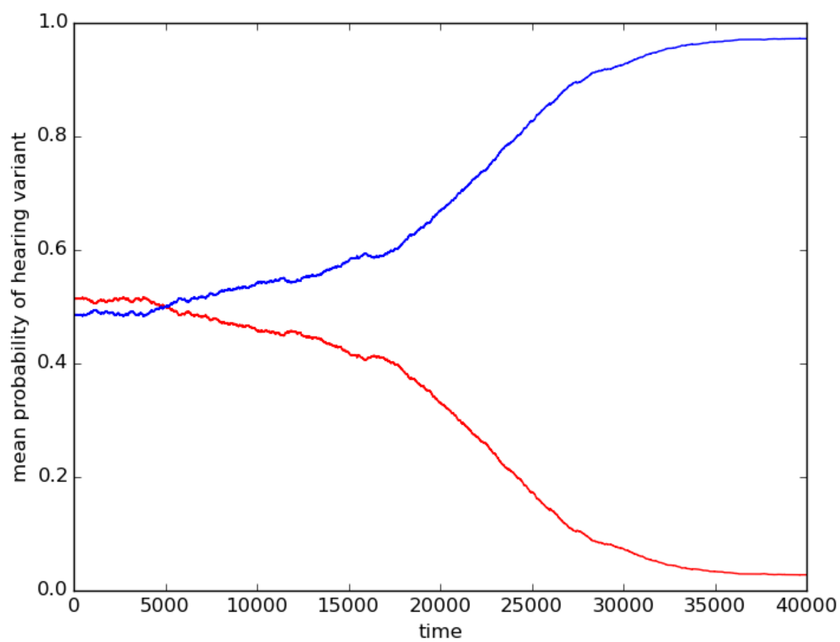

Figure 10: Saltatory change

The results of this version of the model for a network with both loners and hubs are shown in Figure 10. We can see that both variants were present in roughly equal proportions in the network with one of the variants increasing in use around time 5,000 and drastically increasing in use around time 18,000. By the end of the simulation, speakers had reached near categorical use of one variant.

The results for the dense network are shown in Figure 11. In this network, both variants stayed in a more stable distribution for longer with one variant rapidly increasing in use around time 15,000 and reaching categorical use by time 31,000. The lack of loners in this network allowed for one variant to reach this categorical use, suggesting that the lack of categorical use in the network with loners was due not to individuals never reaching categorical use of the variant but rather to some loners in the network always maintaining categorical use of the less popular variant. This is because the grammars of these loners could not be influenced by their neighbors.

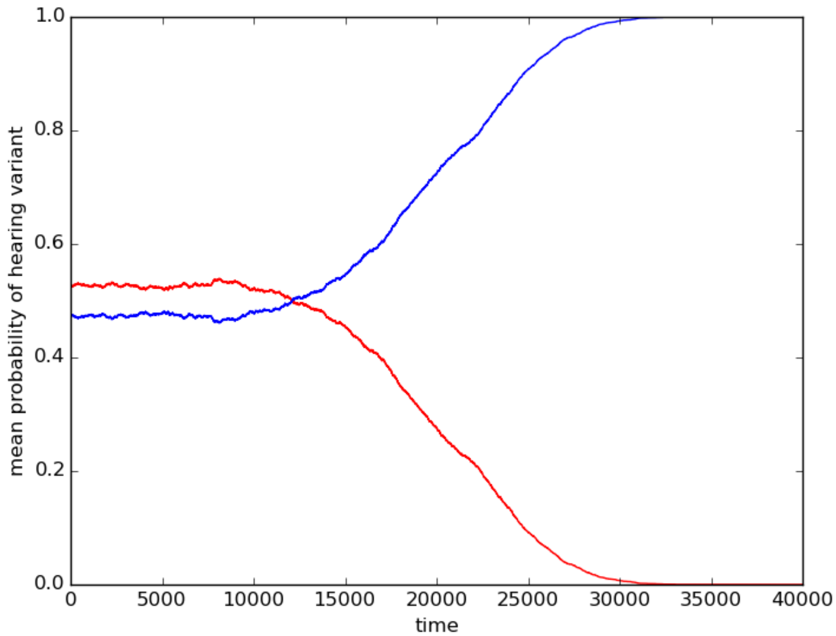

Figure 11: Saltatory change - dense network

Finally, the pattern for the sparse network is seen in Figure 12. In this network, we see only a gradual increase in one of the variants that continues over the entire course of the simulation. This might suggest that a norm would not emerge in this network. However, one possibility is that if a sparse network is allowed to continue in this process of change, it would eventually result in one variant becoming much more frequent than the other. The results for such a simulation are shown in Figure 13, where a dense network is allowed to continue for 80,000 time steps, rather than 40,000. In this trial, we see that, yes, a sparse network can result in the emergence of a clear norm. However, the most drastic increase in the use of one variant does not occur until around time 50,000.

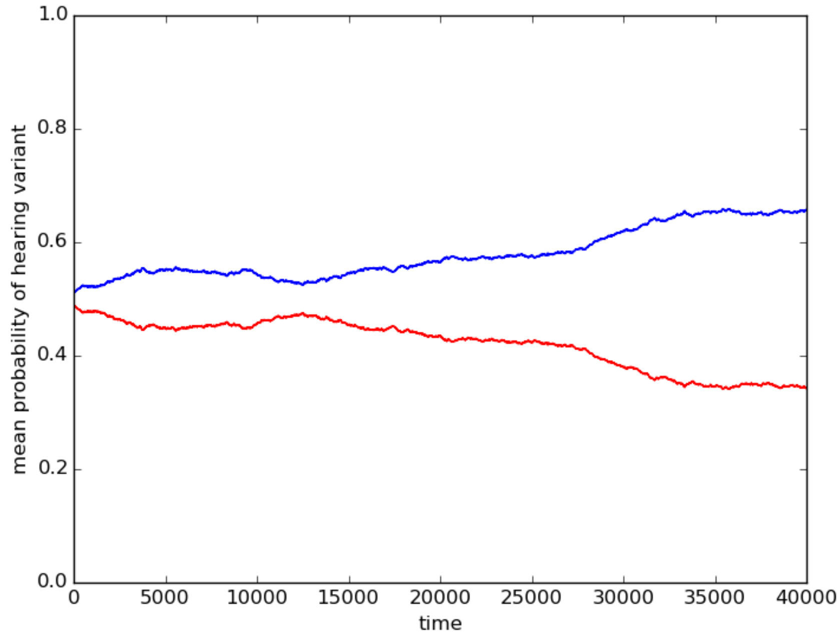

Figure 12: Saltatory change - sparse network

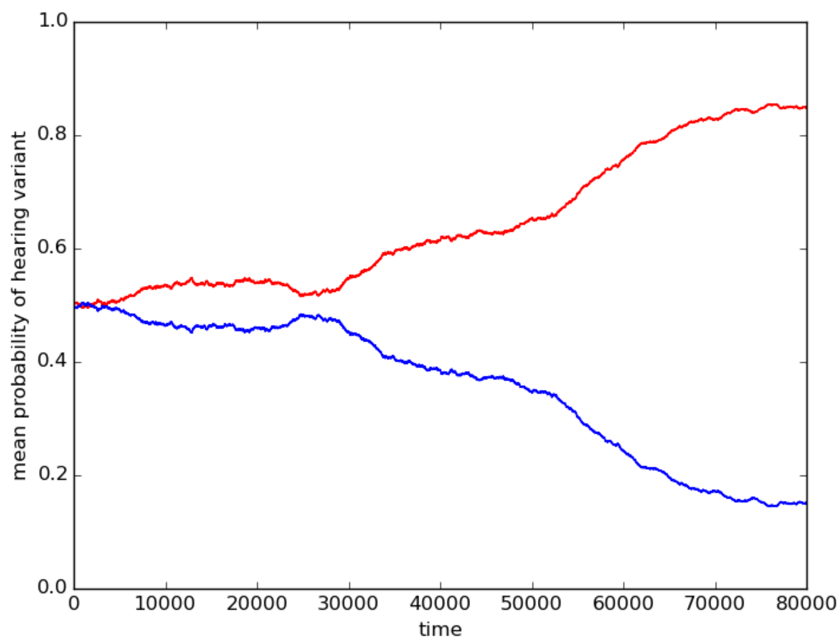

Figure 13: Saltatory change - sparse network, extended time

For this model of change, then, we have seen that the biggest effect of network type is once again on the rate of diffusion. The results of all trials are in line with the predictions of Sankoff and Blondeau (2007) who claim that stable variation over an extended time span is uncommon. Instead, once entering a stage of variability, speakers tend to converge rapidly on near categorical use of a variant. In these trials we see that result. Once one variant begins to become more common, meaning that some speakers are beginning to shift into variable use of that variant, the change to near categorical use in the network proceeds relatively quickly.

A final point about comparing network types has to do with the illusion of stable variation, especially in the sparse network. In both the network with hubs and loners and in the sparse network, the network never converges on categorical use of one variant as seen in the dense network. This, however, is due to the construction of these networks themselves. Both of these network types of have loners whose values cannot be affected by other nodes, resulting in some of the nodes always retaining categorical use of the less common variant. In the sparse network, there are many of these loners, contributing to the illusion that there is stable intra-speaker variation. However, the fact that the sparse network only ever reaches roughly 0.85 mean probability of use of the more common variant actually reflects the presence of many loners with categorical use of the less common variant. It is interesting to note, then, that even with speakers who maintain categorical use of the competing variant, once the use of one variant becomes near categorical for most speakers, there is no oscillation back towards the other variant.

6 Discussion

In the preceding sections we have seen that the properties of the social network have not affected the qualitative outcomes of the trials as much as they have affected the rate of change. Instead, the assumptions of the model about the nature of linguistic variables and the mechanisms of diffusion have a much stronger affect on the qualitative outcome than does network type. This finding seems to contradict the strong effect of network type found by Fagyal et al. (2010) and instead is more in line with the findings of Kirby and Sonderegger (2013) who demonstrated that priors of the model rather than differing population dynamics affect the qualitative outcome observed in modeling language change. With assumptions about linguistic variables and change that are arguably more sophisticated and more closely approximate reality, the effects of network type decrease fairly drastically.

The question then becomes, what does this tell us about the patterns of change attributed to network structure in observed communities, such as those studied by Milroy and Milroy (1985) or Lippi-Green (1989)? If network structure can truly matter for the facilitation of language change, why don't we see those effects here? One possibility is that the patterns seen in these studies were wrongly attributed to the structure of the social network. Since social networks in these studies were quantified only indirectly by scores based on factors like occupation and kinship, perhaps these findings exaggerated the effect of network or overestimated the ability of these scores to correctly mirror network structure.

However, another possibility that seems more likely is that the effects of network on diffusion patterns pertain more to the entrance of linguistic variables into a network rather than the adoption of variables once they are already present in a network. In Milroy and Milroy (1985), for example, the Clonard girls acquired innovative variants through weak tie encounters. As early adopters, they were then able to bring these innovations back to their own communities. Weak ties, therefore, served as a medium for transferring variants between otherwise disjoint networks. This is in line with the sociological work of Granovetter (1973) that claims that the importance of weak ties lies in their role as bridges, that is, as the only connections between groups that would otherwise be isolated. The importance of the role of network structure in admitting new variants is also echoed in Anderson's (1988) distinction between "open communities" that have interdialectal communication, and "closed communities" which lack interdialectal communication. It is in open communities that he claims we most often find the type of simplification or compromise that was shown in §4.1.

It is important, then, to consider what types of situation the models presented here represent. In all of these models, the linguistic variants began with a random even distribution in the network. This starting distribution, then, is not like a situation where two distinct communities have two distinct variants or systems of contrast. Instead, it represents an integrated community where both variants are present throughout the speaker community. This type of starting point is similar to a dialect mixing situation where speakers of two or more dialects are all part of the same community.

One such situation that we could compare this to, then, is the development of New Zealand English, as studied by Trudgill et al. (2000). Trudgill et al. classify New Zealand English as a "mixed dialect", and they observe three stages of change, which they attribute to accommodation. The first two stages observed involve leveling, and the third involves focusing, of the type seen in §4.1. In this mixed dialect situation, they note that it is often the case that majority forms and unmarked forms survive, as seen in §4.2 and §4.3, respectively. In the models discussed here, we have seen that the network structure can affect the rate at which changes occur and can contribute to the degree of focusing observed. However, these other factors that Trudgill et al. (2000) observe as important factors in the development of New Zealand English, such as frequency of the variant, and markedness, contribute more to the qualitative outcome. This finding that network structure is not the main factor determining the outcome of a change under conditions similar to those found in new dialect formation lends support to the position of scholars such as Kerswill and Williams (2005), who for the formation of a new dialect in Milton Keynes suggest that network structure alone cannot explain all of the observed patterns of change.

The results of this study suggest that while certain properties of network structure, and specifically weak ties, might be important for allowing variants to enter a network, as has been claimed by studies such as Milroy and Milroy (1985), once variants are present in a network, such as in a mixed dialect or koineization situation, network structure is not as crucial for determining the propagation of a variant. This distinction perhaps reflects the difference between the actuation problem and the transition problem (Weinreich et al. 1968). Network structure may be important for allowing an outside variant to enter into a network and lead to an innovation in that community. However, network structure may be less crucial for accounting for the transition from variable use of a variant to its categorical adoption throughout the community.

7 Conclusion

In the models presented here, we have seen the effects of four different assumptions about linguistic variables and transfer and three different network structures on the propagation of linguistic variants in a speech community. Whether an incremental view of change with variants in a scalar relationship is assumed or a saltatory view of change with two distinct variants available in each speaker's grammar is assumed, the results are similar. The effect of network structure on the qualitative outcome of the model is less robust than its effect on the rate of change. That is to say, regardless of network structure, roughly the same qualitative outcome is eventually reached.

This finding seems to be contrary to the findings in some of the sociolinguistic literature on networks (Fagyal et al. 2010; Lippi-Green 1989; Milroy 1992, 1980; Milroy and Milroy 1985, 1992). These works have attributed many of the patterns of change seen to what they see as crucial properties of social networks, such as the presence of weak ties. However, in this paper, I have demonstrated that factors such as frequency and markedness affect the dynamics of change more than network structure, and that both sparse and dense networks allow for change, just as networks with both hubs and loners do. These finding suggest that while network structure and weak ties may be important in allowing novel variants to enter a speech community, once variants are present throughout a community, a network structure with both hubs and loners is not crucial in allowing a change to be propagated in the community.

References

Anderson, Henning. 1988. Center and periphery: Adoption, diffusion, and spread. In *Historical dialectology: Regional and social*, ed. Jacek Fisiak, 39–83. Mouton de Gruyter.

Auer, Peter, and Frans Hinskens. 2005. The role of interpersonal accommodation in a theory of language change. In *Dialect change: Convergence and divergence in european languages*, ed. Peter Auer, Frans Hinskens, and Paul Kerswill, chapter 13, 335–357. Cambridge: Cambridge University Press.

Chakrabarti, Deepayan, Yiping Zhan, and Christos Faloutsos. 2004. R-MAT: A recursive model for graph mining. In *SIAM International Conference on Data Mining*.

Chambers, J. K., and Peter Trudgill. 1998. *Dialectology*. Cambridge: Cambridge University Press, second edition.

Coupland, Nikolas. 2008. The delicate constitution of identity in face-to-face accommodation: A response to Trudgill. *Language in Society* 37:267–270.

Fagyal, Zsuzsanna, Aamarth Swarup, Anna María Escobar, Les Gasser, and Kiran Lakkaraju. 2010. Centers and peripheries: Network roles in language change. *Lingua* 120:2061–2079.

Garrett, Andrew, and Keith Johnson. 2013. Phonetic bias in sound change. In *Origins of sound change: Approaches to phonologization*, ed. Alan C. L. Yu, 51–97. Oxford: Oxford University Press.

Giles, Howard, Nikolas Coupland, and Justine Coupland. 1991. Accommodation theory: Communication, context, and consequence. In *Contexts of accommodation*, ed. Howard Giles, Nikolas Coupland, and Justine Coupland, chapter 1, 1–68. Cambridge: Cambridge University Press.

Granovetter, Mark S. 1973. The strength of weak ties. *American Journal of Sociology* 78:1360–1380.

Holmes, Janet, and Paul Kerswill. 2008. Contact is not enough: A response to Trudgill. *Language in Society* 37:273–277.

Kerswill, Paul, and Ann Williams. 2005. New towns and koineization: Linguistic and social correlates. *Linguistics* 43:1023–1048.

Kirby, James, and Morgan Sonderegger. 2013. A model of population dynamics applied to phonetic change. In *Proceedings of the 35th Annual Conference of the Cognitive Science Society*, 776–781. Cognitive Science Society.

Labov, William. 1973. The linguistic consequences of being a lame. *Language in Society* 2:81–115.

Labov, William. 1980. *Locating language in time and space*. Academic Press.

Leskovec, Jure, and Rok Sosič. 2014. Snap.py: SNAP for Python, a general purpose network analysis and graph mining tool in Python. http://snap.stanford.edu/snappy.

Lippi-Green, Rosina L. 1989. Social network integration and language change in progress in a rural alpine village. *Language in Society* 18:213–234.

Milroy, James. 1992. Social network and prestige arguments in sociolinguistics. In *Sociolinguistics today: International perspectives*, ed. Kingsley Bolton and Helen Kwok, 146–162. Routledge.

Milroy, James, and Lesley Milroy. 1985. Linguistic change, social network, and speaker innovation. *Journal of Linguistics* 21:339–384.

Milroy, Lesley. 1980. *Language and social networks*. University Park Press.

Milroy, Lesley, and James Milroy. 1992. Social network and social class: Toward and integrated sociolinguistic model. *Language in Society* 21:1–26.

Sankoff, Gillian, and Hélène Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83:560–588.

Trautmann, Moritz. 1880. Besprechungen einiger schulbücher nebst bemerkungen über die r-laute. *Anglia* 3:204–22.

Trudgill, Peter, Elizabeth Gordon, Gillian Lewis, and Margaret Maclagan. 2000. Determinism in new-dialect formation and the genesis of New Zealand English. *Journal of Linguistics* 36:299–318.

Watts, Duncan J., and Steven H. Strogatz. 1998. Collective dynamics of 'small-world' networks. *Nature* 393:440–442.

Weinreich, Uriel, William Labov, and Marvin I. Herzog. 1968. Empirical foundations for a theory of language change. In *Directions for historical linguistics: A symposium*, ed. W. P. Lehmann and Yakov Malkiel, 95–188. University of Texas Press.