Centers and peripheries: Network roles in language change

Zsuzsanna Fagyal a,*, Samarth Swarup b, Anna María Escobar a, Les Gasser c, Kiran Lakkaraju d

a School of Literatures, Cultures and Linguistics, University of Illinois at Urbana-Champaign, 707 S. Mathews Avenue, Urbana, IL 61801, United States
b Virginia Bioinformatics Institute, Virginia Polytechnic Institute and State University, Research Building XV, 1880 Pratt Drive, Blacksburg, VA 24061, United States
c The Graduate School of Library and Information Science, University of Illinois at Urbana-Champaign, 501 East Daniel Street, Champaign, IL 61820, United States
d Sandia National Laboratories, 1515 Eubank SE, Albuquerque, NM 87123, United States

1. Introduction

1.1. The social dynamics of language change

Language change is a historical process rooted in synchronic social dynamics. In countless communicative interactions between individuals, novel linguistic variants can emerge, diffuse widely, become integrated into the grammar of the speech community and be transmitted to future generations. When the time course of the cumulative adoptions of novel variants is plotted as a graph (Fig. 1), the resulting distribution takes the form of a well-known, idealized S-shaped curve, long assumed to represent the time course of language change (Weinreich et al., 1968:114; Pintzuk, 2003:512; Chambers, 2004:361). The process begins with a new lexical, phonological or morpho-syntactic variant (form or function) emerging in the speech community and setting in motion an unstable competition between incoming and existing variants (Labov, 1994:65–72; Joseph and Janda, 2003; Mufwene, 2008:115–132). At first, the new variant is rare and present only in the idiolect of a few members of the community. Thus, the pressure on individual speakers to adopt it is weak and its use progresses only at a relatively slow rate or, in some cases, might not extend at all. This stage is referred to as the innovation
phase (Croft, 2000:117–165). However, if communicative interactions continue and the new variant acquires some shared social meaning attracting more members of the community, the variant can start to spread at an increased rate. As more and more speakers find themselves surrounded by individuals who use the new variant, the social pressure for adopting the variant increases. The rate of change is the highest when the most individuals are switching to the new variant, i.e. near the midpoint of the idealized logistic regression, or S-curve. This period of accelerated progression is called the selection and propagation, or diffusion, phase. The process slows down and the S-curve ends in an asymptote when the new variant has been adopted by most members of the speech community. We will refer to this final stage of the change as the establishment or fixation phase. In speech communities, this phase coincides with a change in the social evaluation of linguistic variants and/or the reallocation of social and linguistic functions attached to the competing variants in the broader speech community (Trudgill, 1986; Britain and Trudgill, 1999; Eckert, 2008).

Many long- and short-term grammatical (Labov, 1972a) and lexical (Wang and Cheng, 1970) changes have been suggested to follow such an S-curve-like diffusion trajectory (Kroch, 1989; Chambers and Trudgill, 1992; Denison, 2002). This sigmoid function has been found appropriate in so many contexts that it became established as a sort of “template for language change” (Chambers, 2004:361). However, the S-curve is an idealized depiction of a complex process. Lexical diffusion (Labov, 1994:421–439) and syntactic change, for example word order changes from Old to Middle English (Kroch, 1989; Pintzuk, 2003:509–519), have been found variable across speakers and/or contexts. Language-internal and external factors can propel the diffusion of certain variants, but impede the selection and propagation of others (Thomason and Kaufman, 1988; Hock, 1991; Jones and Esch, 2002; Trudgill, 2004; Auer et al., 2005; Jones and Singh, 2005; Mufwene, 2008:128–131). For instance, structural similarities between languages and the prestige of foreign cultures can help advance lexical borrowings, while group ideologies, such as linguistic nationalism, tend to impede on the large-scale diffusion of new foreign words (Blom and Gumperz, 1972; Hock, 1991:412–425; Hinskens et al., 2000). New variants might spread only locally or sporadically in some communities, groups and stylistic contexts, or merely survive in written records as curiosities of their era. Some segments of a population speaking the same dialect can also “go through a period of fluctuation between the use of the old and new variant” (Wolfram and Schilling-Estes, 2003:717), depending on context, register, gender, or socio-economic status. Finally, there is also much uncertainty about when a novel variant becomes established as widely-shared convention of use, or norm, since the spread of innovations can be conditioned by a unique combination of language-internal and social factors (Chambers, 2004:361; Mufwene, 2008:128–131). As Denison (2002:56) concludes: “The whole thing can last hundreds of years altogether [and], indeed, may never be wholly completed”. And yet, nothing in the smooth, prototypical path associated with the S-curve accounts for these probabilistic aspects of change. It is unclear when and how the plethora of language-internal and social factors identified in previous studies exert their probabilistic influence on the selection, adoption and durable establishment of novel forms.

Previous studies have devoted much attention to the innovation phase of language change. Various types of innovation (e.g. Andersen, 1973, 1988, 1989; Hock, 1991; Croft, 2000; Trudgill et al., 2000; Kerswill et al., 2008) and a variety of alterations in form (e.g. reduction, loss, insertion or extension of elements or features) and meaning (e.g. lexicalization, grammaticalization, exaptation) have been identified (Hopper and Traugott, 2003; Brinton and Traugott, 2005; Heine and Kuteva, 2007). However, while novel variants can emerge and be initially selected as a result of endogenous, language-internal, processes, the mechanisms behind their spread and large-scale adoption, or the lack of thereof, are inherently social

Fig. 1. Cumulative adoption of novel variants over time in three stages: emergence (innovation), selection and propagation (diffusion) and community-wide adoption (fixation).

1 Social motivations underlying the selection process have been interpreted in terms of prestige (overt vs. covert prestige, Labov, 1972a) and group solidarity (Le Page and Tabouret-Keller, 1985; Milroy, 1992).

2 Many have questioned whether logistic regression is the appropriate mathematical model (Niyogi and Berwick, 1997; Denison, 2002; Kretzschmar and Tamasi, 2003), while others have used it with different types of empirical data (Kroch, 1989; Labov, 1994:65; Chambers, 2004 for a review).

3 Ayres-Bennett (2004) documents innovations by wealthy 17th century women, depicted in Molière’s comedy: Les Précieuses Ridicules. She points out, however, that these linguistic innovations did not spread: “Despite the prestige attributed to women’s language, many of their usages remained confined to the particular social groupings in which they originated.” (228).
1.2. Simulating the diffusion process

Although skeptical of the feasibility of such analyses, Bloomfield (1933:394) raised the idea of an experiment with the potential to reveal the dynamics of large-scale social interactions behind the spread of novel linguistic variants. He suggested that “fluctuations in the frequency of forms could be accurately observed if we had a record of every utterance that was made in a speech-community during whatever period we wanted to study”. The idea consisted in “keeping a tally sheet” of all occurrences of competing variants, such as “he ran away; he fell down in contrast with away he ran; down he fell”, in the speech of many individuals, which would yield valuable empirical data from a very large, heterogeneous social network of speaker/hearers communicating over a long period of time. “Every time any speaker uttered a sentence, an arrow were drawn into the chart pointing from his dot to the dot representing each one of his hearers. At the end of a given period of time, say seventy years, this chart would show us the density of communication within the community” (Bloomfield, 1933:46).

Links between individuals in tight-knit local sub-groups that communicate primarily with each other would, thus, reveal “lines of weakness” between subgroups conditioning the spread of new, and the maintenance of old, linguistic variants in this large communication network. Formulated as the principle of communication density5 and illustrated with dialectal data from Germanic and Romances languages, the density of “lines of communication” and “the relative prestige of social groups” were identified as the two main conditioning factors of “the spread of linguistic features” (Bloomfield, 1933:345).

Several decades later, Labov (1994, 2001) found Bloomfield’s speaker-oriented view of the diffusion process “a mechanical approach attractive in many ways”, as it is “conceptually economical and […] based upon the act of speaking itself” (Labov, 2001:326). He related Bloomfield’s idea to studies of information flow in influence networks in other social sciences, but also pointed out that most network studies in sociolinguistics, “devoted to one or two isolated groups of a dozen speakers or so”, are still a far cry from the massive databases envisioned by Bloomfield. As the next section will reveal, however, network studies in sociolinguistics can provide a starting point for theoretical models of social interactions underlying the spread of novel linguistic variants.

1.3. Individual speaker roles in social networks

1.3.1. Poles of density

Social networks bring together people who are the most likely to communicate with, and thus influence each other in daily life. Sociolinguistic studies have shown that these micro-social structures are interactional sites where linguistic variation can acquire locally-relevant social-indexical meaning (Eckert, 1989, 2000).

Networks are comprised of individuals (agents) and their complex web of relationships (ties) to friends, kin and acquaintances. Agents are connected by ties of different types and strengths.6 Connection densities and tie strength7 determine agents’ network positions that can be quantified, for instance, relative to these agents’ centrality, i.e. the number of ties by which

---

5 See real and apparent time constructs (Bailey, 2004; Sankoff, 2006; Sankoff and Blondeau, 2007) and geographical diffusion (Britain, 2004; Taeldelman, 2005).

6 “The spread of linguistic features depends upon social conditions. The factors in this respect are doubtless the density of communication and the relative prestige of social groups” (Bloomfield, 1933:345).

7 Density is the proportion of ties in a network relative to the total number of possible ties. Tie strength denotes a more complex phenomenon, which has been defined variably as frequency of interaction, level of intimacy and/or closeness.
they relate directly or indirectly to others in the network. The more direct (or first-order) and indirect (or second-order) connections an agent has, the more centrally-connected s/he is in the network. Conversely, the sparser the amount of ties linking an individual to others, the more isolated or peripheral that individual’s position is in the network. In sociolinguistics, individuals representing various poles of density in their personal networks have motivated multiple ethnographic and quantitative studies. These studies pointed to the crucial, albeit somewhat contradictory, role of central and peripheral members in the selection and propagation of competing linguistic variants. Three relevant studies are reviewed in the following sections.

1.3.2. Leaders and lames

Labov’s (1972b) study of Black English Vernacular (BEV) in three adolescent street gangs of about a hundred people in Harlem demonstrated for the first time that individual social influence (or popularity8) among peers can have linguistic consequences in small personal communities. Network relationships in Labov’s study were established based on responses to open-ended questions, among them about the popularity of group members. Each individual in the network was asked to name all members they routinely hang out with. The largest number of reciprocal namings defined centrally-connected members who were the most popular among their peers, participating extensively in the local adolescent street culture, and also showing the most salient use of BEV features, such as r dropping, t/d deletion, and copula contraction or deletion. These individuals exerted the most influence on their closely connected peers, forming core (tight-knit) and secondary (less tight-knit) subgroups in the network. More sporadic reciprocal namings designated peripheral individuals. These individuals did not share in the local adolescent group life and most of them used the vernacular quantitatively differently from their better-connected peers. Among the peripheral members, lames were the most isolated, having received the fewest mutual namings for popularity among their peers. Lames were largely disconnected from the daily practices of their peer groups, which made them more accessible to outside influence. In their use of the vernacular, they came across as less prototypical speakers. Having “brought their rule system into alignment with that of the dominant white society” (Labov, 1972b:269), their use of BEV was tainted with standard features from mainstream American English.9

Thus, Labov’s (1972b) Harlem study demonstrated that greater network centrality correlates with greater social and linguistic influence. Members who were central to their peer groups, i.e. perceived as the most popular, exerted a strong norm-enforcing influence over their own and others’ social life and, thus, the use of the vernacular. This form of social pressure diminished with group members’ distance from core practices, and was mirrored by a less salient use of the vernacular. Peer influence spread through the network in a wave-like fashion: the most influential central members were its locus; the least influential peripheral members were its weakest manifestation.10

1.3.3. Insiders and outsiders

Social network studies in Belfast’s protestant enclaves (Milroy and Milroy, 1978, 1985) have further elaborated the social aspects of the above findings. In some neighborhoods in Belfast, phonetic variables, such as (th) and (a), two of the most overtly recognized stereotypes of local working class speech, for instance, were primarily used by older males. A five point network strength scale (Milroy and Gordon, 2003:121) tracking ties to kin, work, friendship, and joint leisurely activities revealed that older males’ loyalty to the local vernacular correlated strongly with the density and diversity of their ties to others in the neighborhood.11 These individuals also had few ties to acquaintances outside the group. Thus, just like in Harlem, leaders and their cohorts were shown to be agents of stability of local speech patterns. On the other hand, the Belfast studies have also revealed a well-known pattern: people with ‘weak ties’ act as conduits for innovative variants to flow towards more central members of the network: “Weak ties provide people with access to information and resources beyond those available in their own social circle” (Granovetter, 1983:209). Women and younger males working outside the neighborhood with no local kin and only a few ties to local friends in certain neighborhoods in Belfast were indeed less integrated and used fewer phonetic variants typical to the local dialect. Due to their greater contact with outsiders, their speech showed evidence of innovative supra-local forms in the broader speech community. These individuals, the Milroys argued, were agents of change, referred to as innovators, channeling novel variants towards the local vernacular.

In light of these studies a simple picture of individual network roles emerged: individuals central to a group’s practices exert a norm-enforcing influence that weakens as a function of distance from core practices. Innovations enter via individuals with loose or ‘weak’ ties to these practices.

1.3.4. Icons, loners and in-betweens

Findings from many subsequent studies, however, turned this picture somewhat upside down. Two findings are particularly relevant for the purposes of this paper: (1) contrary to the Milroys’ weak-tie model, charismatic leaders with strong ties to the local community have also been identified as innovators, i.e. primary agents of linguistic change in some social settings, and (2) dichotomous network positions (center vs. periphery) were redefined in terms of a complex social

---

8 Popularity was defined based on ratings of perceived leadership among peers.
9 For instance in subject-verb agreement, the vernacular 3rd person singular forms do, went, and have were dispreferred for the more standard variants does, went, and has.
10 The spatial equivalent of this diffusion pattern in dialect geography is the wave model (Schmidt, 1872 cited in Wolfram and Schilling-Estes, 2003:721).
11 Other studies also quantified network relationships. Cheshire’s (1982) ‘vernacular culture index’ in Reading (UK) was based on six criteria, Lippi-Green’s (1989) dialectal study in Austria used sixteen indicators, and Edwards’ (1992) ‘vernacular culture index’ drew on ten different variables.
continuum of group members, clear or transitional cases of social position and practice. Eckert’s (1989, 2000) ethnographic account of adolescent peer groups’ language use in Belten High, for instance, revealed multiple small networks with at least one leader in each micro-community engaged in activities chiefly characterizing that group. In the center of some of these communities of practice (Lave and Wenger, 1991; Davies, 2005), interactional sites of situated social learning, stood the flamboyant stylistic icons displaying a salient use of pronunciation variants indexing on-going sound change. Well-known for their sociability, these Burnout girls’ extensive network ties were means of gaining visibility and influence among their peers (Eckert, 2000:199–204). Personal styles and tastes in clothing, speech, posture, as well as patterns of consumption rendered these fashion leaders salient prototypes of adolescent cultural practices. They were powerful models for interpersonal accommodation,12 imitated by their closest peers. The near-equivalents of such central figures in other studies (Labov, 2001; Mendoza-Denton, 2008) led to the proposal that leaders of language change are centrally-connected, highly visible individuals whose influence can extend beyond their own personal networks.

Leaders and their cohorts were surrounded by individuals referred to as in-betweeners, “transitional in practice as well as in social space” (Eckert, 2000:189). In-betweeners formed smaller clusters and were conscious about “being on the outs” of more influential groups. The third category of individuals was “the least engaged in the social scene” (Eckert, 2000:192). They referred to themselves as loners. Similar to lames in Harlem, these isolated individuals varied considerably in their speech patterns and personal histories. They were neither the most advanced, nor the most conservative in their use of the most stereotypical phonetic variants, but rather those who appeared to be true outliers in their “noticeably anomalous” use of the local dialect features (Eckert, 2000:207).

Isolated individuals represent a particularly diverse cultural category, variably referred to as outsiders, lames, loners, marginals and oddballs in the sociolinguistic literature (Chambers, 2009:96–114 for a review). Loners are a subgroup of individuals who are not entirely disconnected from local practices, as some might even be tied to the community by strong family ties, but they tend to keep it to themselves and not interact extensively with any specific sub-group in the community. In this study, we will test whether this type of behavior can have a decisive impact on the diffusion of innovations and the community-wide establishment of norms.

1.3.5. Leader and weak-tie models

Studies outlined above support two competing models of speaker roles in language change. The main difference between the two models lies in the source of novel variants. Findings from Labov’s Philadelphia neighborhood (Labov, 2001) and Eckert’s Detroit high-school studies, along with the wave-like diffusion patterns of vernacular features found in Harlem (1.3.4), support the so-called two step flow of influence model (Lazarsfeld et al. 1949 cited in Labov, 2001:356–365). In this model, centrally-connected leaders are the source of novel variants and, primarily influenced by other leaders,13 their linguistic influence percolates through their local personal networks. This model is consistent with other observations of individual leadership in language variation and change: by virtue of their salient positions, leaders have both broader connections to a pool of innovative variants in the linguistic market (Kerswill and Williams, 2000; Trudgill, 2004) and also have the legitimacy to introduce such variants in their communities. On the other hand, the Belfast study showed evidence for the weak-tie model of influence, in which individuals with looser ties to the locals bring about change by virtue of their distance from the regulatory influence of local centers and increased contact with outsiders (Granovetter, 1973, 1983).

Based on the three studies presented above, centrally-connected members and their cohorts are considered conservative agents of stability in some contexts and innovative agents of change in others. Conversely, peripheral individuals are sometimes seen as the source of novel variants and other times appear as outer edges of networks barely touched by innovations. How can both roles be legitimate alternatives? What are the systematic explanations behind these, apparently conflicting, speaker roles? Reformulated as a research question: what roles do more or less centrally-connected members of social networks play in the diffusion process?

In this paper, we use computer simulations to provide some answers to the above questions. In keeping with the view that the spread of linguistic innovations is framed by the broader question of social change, we examine the dynamic role of centers and peripheries in a large influence network. We model diffusion as the probabilistic uptake of one of several competing variants by agents of unequal social standing. We simulate Bloomfield’s thought experiment and systematically vary two factors identified as the most crucial in diffusion: density of communication, modeled as presence or absence of centers and peripheries, and relative prestige, modeled as social influence or popularity among agents. We also examine the social conditioning of the final stage of the process that has received less attention in the past: the establishment (fixation) of novel variants as norms.

2. The computational model

2.1. Wiring the network

We simulate the diffusion of linguistic innovations in a closed influence network. We analyze the diffusion model to the spread of phonetic, lexical and morpho-syntactic forms possessing some indexical value (Silverstein, 2003; Eckert, 2008), i.e.

---

12 For the notion of accommodation and its critics, see Giles and Smith (1979) and Meyerhoff (1998).
13 The equivalent diffusion pattern in dialectology is referred to as the gravity or hierarchical model (section 4.2.1).
conveying group-membership, speaker stance and style, or any group-based ideology that can motivate speaker/hearers to adopt one competing variant over another. We carry out the simulations in an artificial but socially realistic influence network, whose local characteristics bear close resemblance to properties of small local communities studied by sociolinguists (1.3). Within this speaker-oriented theoretical framework (Croft, 2000:53–59), we rely on computational research on networks in the social sciences (Valente, 1995; Macy and Willer, 2002; Mason et al., 2007) to bridge the gap between empirical findings on small-scale vs. large communities. We analogize the resulting network to a large, highly polarized social space, analogous to urban working-class neighborhoods or clusters of adolescent peer groups in school settings. We use the structural (topological) characteristics of this network to simulate social distance among individuals taken to mirror fixed class- or caste-based social inequalities, as agents only communicate directly with their closest neighbors (2.3).

Consistent with Bloomfield’s hypothesized system of “dots and arrows” connecting a large number of speaker/hearers in multiple tight-knit communities (1.2), the network is modeled as an influence graph. Links (or ties) correspond to symbolic information-carrying connections between nodes representing individual agents. Fig. 2 provides a schematic representation of the type of influence between agents. Links from node A to nodes B, D, E and F indicate that A listens to B, D, E, and F, and thus can be influenced by all four of them. The link between agents F and E is reciprocal, which means that these two agents can mutually copy each other’s variant. The number of arrows pointing to an agent corresponds to that agent’s in-degree, which is indicated by the first number in parentheses in Fig. 2. The number of arrows pointing away from an agent represents that agent’s out-degree, which is expressed by the second number in parentheses. The higher an agent’s in-degree, the more influential that agent is considered. Conversely, the lower an agent’s in-degree, the fewer is the number of agents willing to adopt the variable used by that agent. Agents B and D who have links pointing to them but not from them are loners who listen to no one else in this closed network. Agents E and F, with the highest in-degree, are the most influential.

Fig. 2 depicts possible paths of interpersonal influence. As opposed to Bloomfield’s implied directionality of arrows pointing from speakers to hearers (1.2), the orientation of arrows in our network represents the point of view of the hearer: arrows point from hearers susceptible to pick up linguistic influence to speakers possibly spreading it. These paths can or cannot be activated, depending on the type of biasing selected (2.3). The idea of an influence network remains consistent with Bloomfield’s conception of direct conversational interactions: “[in the network] some speakers would turn out to be in close communication’, while others ‘widely separated […] had never heard each other speak and were connected only by long chains of arrows through many intermediate speakers” (Bloomfield, 1933:46–47). The model can also capture patterns of reciprocal and non-reciprocal links seen between peers in Harlem (1.3.2), by showing that individuals belonging to the same network are not equally likely to influence each other’s linguistic behavior. Our influence network has scale-free small-world properties (Watts and Strogatz, 1998; Barabási and Albert, 1999) that can be characterized by three features:

1. **Small diameter**: Short average distance between pairs of nodes. Most pairs of agents are connected to each other by short chains of acquaintances.
2. **High clustering**: Nodes linked to the same node are also likely to be linked to each other. Most pairs of agents tend to belong to small groups or cliques.
It is not known why many large-scale social networks have these particular structural properties. It is recognized, however, that diameter, clustering, and network topology have important consequences for the flow of information (here social and linguistic influence) within and between social groups. For instance, the smaller the diameter (fewer intermediates between agents), the faster rumors and news of emergency travel down the network. The higher the clustering, or cliquishness in a group, the more likely it is that members would share similar attitudes and unite for instance in various subgroups of e-communities (Latané and Bourgeois, 1996). Also, agents with a scale-free interaction topology can achieve consensus of opinion very quickly (Delgado, 2002; Pujol et al., 2005).

These structural properties are also relevant for speech communities. The large majority of sociolinguistic studies of social networks have targeted “one or two isolated groups of a dozen speakers or so” (Labov, 2001:326), grouped in “high-density networks in which everyone knows all the others” (Chambers, 2009:80). Thus, personal networks studied by sociolinguists can be described as small in diameter and exhibiting high clustering. Lack of sociolinguistic studies of very large groups currently prevents us from knowing much about the structure of such linguistically-relevant large influence networks. However, judgment samples of over a hundred people in the Philadelphia Neighborhood Studies (Labov, 2001) and of about two hundred high-school girls’ social networks in Belten High (Eckert, 2000:173) point to highly heterogeneous networks. The social continuum of friendship clusters in Belten High (1.3.4), encompassing anyone from the most flamboyant leader to the most isolated loner and a number of in-betweens, can also be viewed as analogous to a scale-free small world social structure, in which most agents have average influence over their peers, a number of them are very peripheral, but it is not surprising to see a few disproportionately influential leaders.

To generate such an artificial social network, we used the R-MAT algorithm of Chakrabarti et al. (2004). R-MAT or Recursive MATrix is an algorithm that operates on the adjacency matrix of the network to create a set of nested communities in the network (Fig. 3). It works as follows: if agent \( x \) is influenced by agent \( y \) (i.e. there is a link from \( x \) to \( y \)), then we place a 1 at row \( x \) and column \( y \) of the adjacency matrix, otherwise we place a 0 at that location. The R-MAT algorithm uses four parameters, \( a; b; c; d \), which correspond to four quarters of the adjacency matrix (Fig. 3). We start with an adjacency matrix filled with zeros. We then choose a quarter of the matrix with probability corresponding to its parameter. We chose the parameters \( a = 0.5, b = 0.1, c = 0.1, \) and \( d = 0.3 \). These parameters mean, for example, that half the time we choose the upper left quarter of the matrix. We then treat the chosen quarter as a new matrix, divide it into quarters, and again choose one quarter with the same set of probability parameters. This process is repeated recursively until we end up with a single cell, whereupon we set the value at that cell to 1. The above process is executed from the beginning for a pre-determined number of times to create the network. Any particular cell in the adjacency matrix might end up being chosen more than once, in which case the final number of links in the network will be slightly fewer than the number of times the process is carried out.

In order to make sure this network has the properties identified above, i.e. small diameter (members in close reach), high clustering (tight-knit community) and scale-free degree distribution (a few very influential and many isolated members), we created a relatively densely-wired network with a large number of agents: 900 nodes with links added to the adjacency matrix 9000 times, resulting in 7561 unique, i.e. single new, links. Since the number of possible links is 810 000, this should be regarded as an average density of connections, and it will be varied during the simulation (3.2).

The number of agents is a hypothetical number chosen for two reasons. The first is to generate a relatively large population of social agents. Although most sociolinguistic studies of social networks have not worked with more than a few

---

Footnote:

14 Models based on spatial distribution (Wong et al., 2005) and competitions of people for limited resources (Anghel et al., 2004) have been proposed as possible explanations.

---

Fig. 3. Adjacency matrix recursively divided into quarters, with each quarter having a probability \( a; b; c; d \) associated.
dozen individuals, and thus could not yield an estimate of how large such populations should be, there were a few noticeable exceptions. The largest samples of working class neighborhoods (Milroy and Milroy, 1985), adolescent peer groups (Eckert, 2000) and bilingual (im)migrant communities (Bortoni-Ricardo, 1985) extended up to about 150 individuals, which is the presumed cognitive limit of the number of individuals with whom any one person can maintain regular and stable face-to-face relationships in tight-knit social groups (Dunbar, 1993). Research in the social sciences also indicates that the emergence of positional, i.e. class-based, inequalities is linked to group size (Reuben, 2004). For this reason, our simulations of the spread of linguistic influence in a large heterogeneous social structure had to assume many more than 150 agents. The second reason for a very large population of agents was to guarantee a network with a large enough population of individuals to be grouped into many nested communities of different size showing structure at all scales in the network. The maximum in-degree of 53 in our model is well within Dunbar's number15 and represents only a small fraction of network size. This means that the network is truly populated by many small hubs rather than just a few large 'super-nodes' that could influence everyone else.

2.2. Simulating structural social influence

We model the spread of linguistic variants as follows. We assume \( k (k > 1) \) possible variants for a given phoneme, lexeme, morpheme or syntactic structure. For example, we can think of the selection of flapped or fully released /t/ in the word 'mitten', or the choice between a yes/no question formulated with or without periphrastic do in American English as \( k = 2 \). The variants can stand for stylistic and dialectal lexical variants of voiture 'car' as 'véhicule', 'bagnole', 'char' and 'taco' (\( k = 4 \)) in French, or variants of the Spanish diminutive as –ico, –ito or –illo (\( k = 3 \)). We use eight variants in the simulations. Eight is a somewhat arbitrary number, chosen to reflect a collection of coexisting, competing variants that present significant choice complexity across the population.

To initialize the model, a uniformly randomly chosen variant is assigned to each agent in the network at time \( t = 0 \). At each following time step, i.e. single instance of interaction between agents, one of the agents chosen uniformly randomly updates (replaces) its variant by one of its neighbors’ variant. The term neighbor stands for an agent pointed to by direct links from another agent. If the neighbor happens to have the same variant as the agent, the latter does not update its variant. Agents maintain only one variant at a time.

Fig. 4(a) shows the in-degree distribution, i.e. number of incoming links, for groups of nodes in the network. The maximum in-degree and out-degree in the network are 53 and 50 respectively, both for agent 0. Thus, no single agent has direct influence over a large fraction of the population, which is consistent with the cognitive limit of individual social maximum in-degree and out-degree in the network are 53 and 50 respectively, both for agent 0. Thus, no single agent has maintain only one variant at a time.

Fig. 4(a) shows the in-degree distribution, i.e. number of incoming links, for groups of nodes in the network. The maximum in-degree and out-degree in the network are 53 and 50 respectively, both for agent 0. Thus, no single agent has direct influence over a large fraction of the population, which is consistent with the cognitive limit of individual social influence in large groups (2.1). There are several agents with relatively high in-degree. These agents are called hubs. They are major centers of influence and stand for structural equivalents of some of the charismatic leaders identified in linguistic studies of social networks (1.3). In keeping with the idea of a continuum of network positions, these hyper-influential hubs are surrounded by other central, but somewhat less influential nodes. Fig. 4(b) depicts the dense core and the sparse edges of the network. Agents on the edges of the network are not influenced by the rest of the population. They are loners, i.e. peripheral agents who are very little engaged and therefore not influenced by other agents (zero out-degree), though they may exert some small influence (low in-degree) on them. They are analogous to lames in Harlem (1.3.2) and other isolated figures well-known in the sociolinguistic literature.

2.3. Simulating individual social influence

Agents’ influence is modeled by a rule that specifies which neighbor they should imitate. We take this rule to be a rudimentary model of inter-personal accommodation. In the first simulation, agents use a modified opinion/imitation dynamics rule known as the voter model (Sood and Redner, 2005; Tavares et al., 2007).16 The simplest version of the voter model assumes that: (a) agents are ready to imitate any of their neighbors holding a variant different from theirs, and (b) all neighbors can be treated equally. We consider both assumptions unrealistic in real-life social interactions. Agents’ social distance and social impact on other agents have been shown, empirically and computationally, to play a crucial role in the spread of innovations (Nowak et al., 1990; Nettle, 1999; Mason et al., 2007). For this reason, agents are situated in a heterogeneous social network structure that allows them to communicate only with their closest neighbors (2.1). To simulate unequal social impact between individual agents, we also modify the voter model. We call the modified version in-degree-biased voter model consisting of agents copying neighbors with probability proportional to those neighbors’ in-degree. The equation for an in-degree-biased voter model is written as follows:

\[
P(i) = \frac{D_i}{\sum_k D_k}, \quad \forall i, k \in N,
\]

where \( P(i) \) is the probability that the agent picks neighbor \( i \), \( D_i \) is the in-degree of neighbor \( i \), and \( N \) is the set of all neighbors of the agent.

15 To obtain a higher maximum in-degree, a network greater than 900 nodes has to be built. The solution adopted in this paper is preferred because it increases agents’ in-degree without altering the fraction of loners present in the network.

16 The voter model has a long history in statistical physics, and has been used to simulate linguistic interactions in other contexts (Castelló et al., 2006; Castellano et al., 2007).
In-degree biasing does not model language-internal bias for different variants (Nettle, 1999), but it introduces a biasing factor in the selection/propagation process in terms of individual social distance: it assumes that agents know the extent of individual influence over other agents, and thus are more likely to copy a highly influential neighbor. This “socially-informed choice” by agents is not deterministic; it is just probable. Agents are biased but not determined by their neighbors’ in-degree, which means that they can, at times, ignore their neighbors’ known influence on, or popularity with, others, and choose to adopt a linguistic variant from a less popular neighbor instead. This makes the dynamics of the diffusion process a more realistic model of inter-personal accommodation that was shown to be sensitive to group members’ perceived popularity with others (1.3). In sociolinguistic terms, agents in our simulations are part of the same speech community by virtue of the fact that they are modeled to have a shared system of social evaluations. All of them are aware of the more or less prestigious language use of peers that they know the best, i.e. their most immediately connected neighbors in the network. On the other hand, agents are also free to follow newly emerging trends and fashions in their specific choice of a linguistic variant.

Analogous to Bloomfield’s dynamically updated ‘tally sheets’ of competing linguistic variants (1.2), we keep a running count of variants that are selected by the agents at each time step during the simulation. Time is quantified as number of interactions between agents. Since the approach taken here is theoretical, no correspondences with specific time units in social history (years, decades, centuries, … etc.) were established.17

2.4. Hypotheses

Our first hypothesis is that the scale-free small-world network architecture, together with the in-degree-biased voter model of social influence between interacting agents should lead to complete diffusion and establishment of a variant as norm. We define complete diffusion as the adoption of at least one competing linguistic variant by a large majority of agents over time. Thus, we expect patterns of density of communication and relative social prestige, i.e. structural and individual

---

17 In one simulation of a specific historical event, the spread of New Zealand English, retention in memory of a particular linguistic variant was used as time unit, bridging the gap between historical time and time of simulation (Baxter et al., 2009). Future research might focus on the best correspondence between these two scales.
aspects of relative social influence in the network, to be optimal for the successful diffusion and establishment of one linguistic innovation as norm. However, since loners in the network are not sensitive to their closest peers’ influence, and other agents’ selective attention to popularity is only probabilistic, our second hypothesis states that there should be more or less long periods when none of the competing variants clearly wins over the others. We analogize these periods to transitional stages showing possible local reversals and advances in the use of old vs. new variants in certain groups of speakers (1.1). Our third and last hypothesis is that loners and hubs are essential elements of a successful diffusion process. Since their exact, dynamic contribution is difficult to predict, we propose to systematically eliminate them and observe the behavior of the system in each condition.

3. Results

3.1. Diffusion dynamics

Fig. 5(a) shows the diffusion curves of competing variants on a condensed time scale. Although only five out of eight variables are shown, all variables exhibit the same path of cumulative adoptions over time: all agents in the network (y axis) adopt each of the competing variants after some period of time (x axis). Thus, hypothesis one is confirmed: simulating the spread of innovative variants as individuals’ probabilistic preference for relative prestige in a large, socially heterogeneous population leads to the establishment of one variant as a generalized convention, or norm, over time. By using in-degree-biased modeling in a scale-free small-world social influence network, we successfully replicated the three stages taken to be the generic path of language change: initial stasis (innovation), steady propagation (diffusion) and final fixation (establishment) of new variants.

At first glance, the particularly steep slopes of the diffusion curves in Fig. 5(a) appear to be jumps rather than gradual transitions. This is due to fact that the x axis represents only every 100th time step, and the total time depicted on the x-axis is very long with respect to the rise time of any norm. This compresses each curve and makes them appear steeper than they actually are. Fig. 5(b) depicts the diffusion curves of two variants in greater detail, though still at every 100th time step. It reveals, as hypothesized, that the trajectory leading to complete diffusion is, at times, quite rugged. Although both variants

Fig. 5. (a) Number of agents adopting each of the competing variants in the network over time; (b) periods of partial diffusion preceding establishment as norm for the last two variants; (c) middle portion of an S-like diffusion path of one variant on an extended time scale.
shown in Fig. 5(b) end up stabilizing as norms over time, each has demonstratively undergone a longer phase of partial diffusion during which none of them has been propagated beyond a little over half of the agents in the network. At these instances, the variables are shown to compete with one or more old variants that they end up forcing out over time. This means that, for a substantial length of time, more than one variant was simultaneously used by different segments of the population. This confirms hypothesis two: the simulation is able to generate a probabilistic diffusion path.

Fig. 5(c) shows the middle section of the cumulative adoption curve of a single variant on an even more extended time scale than Fig. 5(b). The curve is rugged and shows gradual, rather than jump-like, transitions. The slow initial rise from time steps 20 800 to 21 250 is followed by a more rapid rise between time steps 21 250 and 22 000, leading to a final deceleration from time step 22 000 until the end of the curve. Although the curve is not prototypically S-like, its direction and recognizable phases indicate one possible path of diffusion arising as a result of a stochastic process. In some cases, the diffusion process can be temporarily halted and then boosted again until eventually all agents but loners switched to the most widely preferred variant. In other cases, such as the one shown in Fig. 5(c), diffusion is slow but progressing steadily: once agents start adopting one variable in greater numbers, the process eventually culminates in a nearly complete adoption phase, with one variant more easily winning over the others. Note that all these scenarios are empirically valid (Chambers, 2004:360–366 for different types of curves). Change in monolingual situations is typically slower than in contact situations, and gradual morphological change can co-exist with lexical changes that are accomplished within the lifespan of a single generation (Joseph and Janda, 2003). Although, in our simulations, each new variant ends up forcing out the preceding form, which is not always the case in real-life scenarios of language change, the diversity of diffusion scenarios and the variability in the rate of diffusion indicate that the model is able to account for the variable use of competing forms in the population. As change is typically underlain by variation, we take this to demonstrate structured heterogeneity, such as dialectal and/or sociolectal variation in different segments of the broader speech community.

We also ran a number of other simulations to identify the factors that lead to the emergence and change of these diffusion dynamics in the network. Three factors turned out to be decisive: (1) existence of loners, (2) existence of hubs, and (3) in-degree-biased selection of a variant. We show next that the removal of any of these three factors drastically alters the behavior of the system: norms either fail to emerge, or emerge once and become fixed forever.

3.2. Manipulating network structure

3.2.1. Lack of loners

In keeping with Bloomfield’s prediction that the density of lines of communication is a decisive factor in the diffusion process, we first altered the current network structure by making it so dense as to lack isolated individuals. We ran R-MAT with the same parameters as before, only this time we carried out the link addition process 27 000 times, resulting in a network with about three times as many links as before. This resulted in the multiple linking of all 900 individuals, and lead to a very dense network where all extremely peripheral individuals, or loners, were removed. We ran the simulation on this new network again with $k = 8$ competing variants. The result is shown in Fig. 6. We see that the population converges very quickly to a norm, which then stays fixed: all 900 agents adopt one variant, eliminating all the others completely. In the previous simulation, there were a few loners whose chosen variants never changed and, therefore, only matched the rest of

18 Two of our reviewers suggested that the lack of a long and flat initial trail and the ruggedness of the rise are not characteristic of an S-curve.

19 We could have removed loners by simply adding a link from each loner to some other node in the network. This would have resulted in the same dynamics as in Fig. 6. However, increasing the density of links to remove loners has the secondary effect of showing that more dense networks are less likely to have loners and therefore are less likely to show the behavior demonstrated in Fig. 5(a).
the population by chance (Fig. 5). These loners served as either sources of new variants, or repositories of past norms, and thereby multiple variants could be simultaneously maintained in the population. This allowed the norm to change over time as, probabilistically, a loner’s variant could be picked up by some other agents and propagated through the rest of the population. In absence of such very peripheral agents, this clearly cannot happen, as very high density leads to hyper-rapid convergence at the expense of all future network-internal innovations. Thus, as hypothesized, the contribution of peripheral individuals is crucial for the social dynamics underlying the diffusion process.

3.2.2. Lack of leaders
To test whether the existence of very influential leaders and their local communities are also a crucial factor in the establishment of norms, we changed the topology of interactions by removing hubs from the network. Instead of a scale-free network, we created a random network with approximately the same number of loners as in the network shown in Figs. 4 and 5. We created a random network by starting with an empty network of 900 agents. We then chose a from agent and a to agent independently and uniformly randomly and created a link between them, 3000 times. The choice of the number of links was dictated by the resulting number of loners. Since, at this step, we investigated the effect of hubs on the dynamics, it was important to keep the number of loners comparable in, both, the scale-free and the random network conditions. The number of loners in the former was about 2% (20 out of 900 agents). Adding 3000 links resulted in 4% of loners (35 out of 900 agents) in the random network. This range of variation was considered minimal and, thus, the number of loners in the network comparable.

The resulting random network showed the diffusion dynamics illustrated in Fig. 7. We can see that in absence of hubs norms fail to appear. There is some propagation of variants over time, i.e. there is intense competition between several variants, one of which emerges as the most widely adopted for some time, but none of the variants is spread to the majority of agents. Thus, the cumulative adoption of novel variants by centrally-connected, hyper-influential individuals is crucial for a successful diffusion process. This confirms hypothesis three. Hubs diffuse an existing change by propelling locally salient variants quickly through the majority of the population. This causes norms to appear and be maintained over some period of time.

3.3. Manipulating individual prestige
As a final step, we return to the scale-free network structure used to generate the diffusion dynamics shown in Fig. 4. This time, we change the so-called update rule. We do so with the intent of changing agents’ probabilistic bias for adopting more “popular” individuals' variants. We simulate two scenarios. In the first case, we revert to a simple version of the voter model (2.3). Instead of biasing the choice of a neighbor to imitate by that neighbor’s in-degree (number of times chosen as popular), agents treat all their neighbors equally. This has the effect of diminishing the number of times hubs are chosen to be copied, because agents no longer select the most influential agents around them. This automatically results in raising the number of times loners are chosen to be copied, because individuals with virtually no positional (network structural) influence over others in the population may also be chosen by their peers. In this scenario, loners’ variants are expected to enter the population more often.

The outcome of this scenario is shown in Fig. 8(a). The population of agents show no collective sensitivity to individual agents’ social influence, which leads to failure of appearance of norms. The competing variants diffuse in a chaotic way: some are sporadically propagated for a short amount of time by some influential hubs, but since too many equally attractive variants are present in the population, a single variant cannot emerge and become established as a norm for an extended period.
In the second scenario, we simulated individual social influence by giving agents random sensitivity to the relative prestige of their peers. We assigned a random number between 0 and 1 to each agent at the beginning of the simulation. Instead of being biased by their in-degree, the agents used the numbers assigned to their neighbors to decide whom to imitate.

The outcome of this case, shown in Fig. 8(b) falls in between the unbiased (Fig. 8(a)) and in-degree-biased cases (Figs. 5–7). Although different variants became dominant at different times in the population, there was considerable fluctuation over the number of agents adopting one variant rather than another as a stable norm over time. The reason for this is that hubs were hampered in their role of spreading popular variants and enforcing established norms. Since individual agents were not specifically biased towards more popular members, there were no widely shared assumptions as to what agents' variants others should “value” as prestigious in the broader community. This biasing was not sensitive to other agents’ actual social standing, since the degree of social distance that agents took with respect to their neighbors was not based on information from the network. Interestingly enough, however, random biasing did go some way of simulating individual social distance. As opposed to purely mechanical selection, agents here did not assume that they can treat all neighbors equally. However, agents also did not share a common view, or joint convention, as to whose language use everybody is supposed to value and, thus, replicate in face-to-face interaction.

4. Discussion

4.1. Centers

The dynamic approach to diffusion adopted in this paper yields the following answers to our research questions (1.3.5). Centers (hubs or leaders) and peripheries (loners and other outsiders) represent two facets of the same diffusion dynamics. These poles of density can be agents of stability or agents of change depending on how their roles in the diffusion process are interpreted in specific social and historical contexts.

Centers are best thought of as actuator switches of the spread of social influence. The analogy borrowed from mechanics is not fortuitous. Similar to mechanical components, whose basic function is to convert energy to movement, centers’ most essential role is to serve as conduits for innovations to propagate to others in the network. Although new norms cannot be established without preference towards variants held by the most popular central agents (4.4), perceptions of popularity do not need to be perfectly correlated with these agents’ in-degree for variants to be propagated through their extended social networks. Evidence for this comes from our simulations with random biasing (Fig. 8(b)) where some variants occasionally became dominant even though their selection was uncorrelated with agents’ actual in-degree. Agents’ action, however, is neutral to the specific outcome of locally-bound social dynamics. Centrally-connected charismatic leaders can advance a new and vigorous vowel shift, such as the Northern Cities Shift in their extended adolescent peer groups in Belten High, Michigan, and thus be perceived as agents of on-going change. They can also propagate local forms in their extended working-class communities in Belfast, resisting the intrusion of mainstream Irish influence, and thus coming across as safekeepers of local norms. The stability of established norms, however, is tied to centers’ perceived social influence: in closed, tight-knit social networks their influence depends on how many agents remain sensitive to their relative prestige over time.

Our simulations have also demonstrated that centers are propagators and enforcers of norms, but they are not necessarily innovators. They can, but do not have to, be “introducers” of novel variants (Croft, 2000:179; Mufwene, 2008:67). The distinction between these two roles is crucial. Recall that variants in our simulations were assigned uniformly randomly to agents rather than systematically associated with the most densely-connected centers (or hubs) in the network at the start. For this reason, the direction of the changes we simulated cannot be argued to have come either from the less influential

---

20 This type of selection is also referred to as “neutral interactor selection” (Baxter et al., 2009:271).
peripheries or from densely-connected centers that can influence many agents in the network. Regardless of what agent started out with what variant during the initialization process, whenever hubs were present in the network, they successfully propagated one of the competing variants. Highly-influential centers can, of course, also introduce novel variants, copy innovations from other leaders and their cohorts. Calling them innovators, however, obscures their essential function of spreading, rather than inventing or introducing, the most socially prestigious variants. Thus, our results lend support to the view that the dynamics underlying the propagation phase should be examined separately from constraints motivating the innovation process (Croft, 2000).

4.2. Loners

4.2.1. From safe-keepers to innovators

Similar to centers, certain types of peripherals (referred to as loners) are neither innovative nor conservative. Their action of holding on to linguistic variants has been interpreted in historical dialectology as the peripheries’ tendency to conserve vestigial forms: “The more isolated area usually preserves the earlier stage”, states Bartoli’s well-known first law (Mańczak, 1988:349). This, however, is only one part of the picture. Our results show that when loners are part of a population structure that allows their influence to reach centrally-connected hubs, they can have a decisive impact on the linguistic system over time.

Empirical studies provide numerous illustrations of the spread of peripheral variants to a majority of speakers in a speech community. Peripheral individuals have been shown to channel new phonetic and lexical variants from outside (Milroy and Milroy, 1985) or to “recycle” old forms that might be perceived as novel after some time. Such “recycling” of old linguistic material has been attested, for instance, in teenage talk (Trudgill, 1999). The use of well as an adjective modifier (e.g. well bored, well hard) in the speech of London teenagers is analyzed as a feature going “all the way back to Beowulf and the eight century [but] in dormant existence until the late 20th century when it was taken up again and revived in the London teenage talk” (Stenström et al., 2002: 158 – 159). Similar examples are attested in working-class youth talk in French (Gadet, 2003). If, as Wolfram and Schilling-Estes (2003:729) suggest, “the social network model may be seen […] as a potential explanation for the diffusion patterns that dialectologists and sociolinguists have already observed”, then the influence of loners in our simulations can also be taken to illustrate contra-hierarchical spatial diffusion in dialect contact. This type of diffusion from the peripheries towards the center, rather than the other, most frequently attested “hierarchical” way, has been shown for old dialectal forms, such as the low-back vowel merger in Oklahoma and the traditionally low-back realization of /ay/ in Ocracoke, North Carolina (Bailey et al., 1993). Our findings strongly suggest, however, that group ideologies do not necessarily have to be evoked to account for contra-hierarchical diffusion. Recall that in our simulations, all agents shared the same probabilistic bias towards competing variants (2.3 for biasing by in-degree). While the choice of a given variant was probabilistic, i.e. agents could or could not copy the form held by their most popular neighbor, agents did not have more or less individual propensity to adopt innovations. Consequently, the spread and community-wide adoption of the loners’ variants in the population cannot be attributed to agent- or group-specific social-psychological biases inducing greater convergence. This means that the “backward flow” of vestigial variants might be a natural outcome of the stochastic processes underlying the spread of innovations in closed and tight-knit social networks. In such networks, individuals are at close reach from each other and can converge relatively rapidly on the type of language use to be valued the most. Thus, in some cases, analyzing contra-hierarchical diffusion in terms of “revitalization of traditional norms” (Bailey et al., 1993:386) might just be a post hoc reinterpretation of the diffusion dynamics characterizing relatively closed social networks with scale-free small-world properties.

4.2.2. The spread of rare variants

The sociolinguistic literature provides rich empirical evidence that a small number of speakers can exert lasting influence over the language use of a large community. With respect to the spread of the labio-dental approximant over the apical trill in Norwich, for instance, Trudgill (1999:319) argues that the seemingly “idosyncratic speech impediment” of about five percent of the speakers using the labio-velar approximant in the late sixties has become “the majority form among younger speakers” in a single generation. However, loners are not just a group of speakers whose linguistic influence would automatically spread over time. Loners represent a particular type of extreme peripherals best portrayed as the recluse “go it alone” who often behave more like spectators than participants in local daily life. While these individuals are connected to a few of their peers in the network, they conduct their daily business largely outside of others’ influence. The social significance of these loner figures transcends cultural and political boundaries. Their strange personalities, peculiar language use, and often unusual social histories have been documented in a variety of settings. Take, for instance, Ignaz from Grossdorf, Austria (Lippi-Green, 1989). Ignaz, a venerable farmer with kinship ties to some people in the community, was a loner by every measure. Unlike residents who worked outside the village but spent most of their free time with the locals, Ignaz tended to avoid such activities, showing “disinterest or even apathy about the social life of the village” (Lippi-Green, 1989:230). In Ignaz’s small, close-knit, and secluded alpine village, the use of one of two low back vowels was correlated with residents’ allegiance towards local village life. Like other locals, Ignaz used the older form of the vowel, but he used it quantitatively differently and often with the non-local variant in the same context. For any long-time resident, these idiosyncrasies would be surprising; for a loner, they are emblematic. According to Chambers (2009:100), “if [Ignaz] were young and black in Harlem, he would be a lame”. 
Another well-known loner figure in the sociolinguistic literature is Reg, “a level crossing gatekeeper in a very isolated rural area between Peterborough and March” in the region of the Fens, England (Britain, 2003:195). He was also compared to other peripherals in his local community, and also identified as a peculiar type that, unlike nomads, i.e. peripherals who move around (e.g. an individual named Zak in the same study), Reg was a secluded man and an ultra-conservative speaker. He was a recluse in one of the most isolated communities of the region where “we can […] expect to find changes either being much slower than in more central areas, or the preservation of (often complex) earlier linguistic structure” (Britain, 2003:205).

Loners in our simulations also represent such isolated figures. They do not listen to others, but there are a few members in the community who are in touch with them and can be influenced by them. Note that both characteristics are important for the diffusion dynamics observed in the model. If the loners did not have zero out-degree, i.e. if they would occasionally copy the variable of any other agent, then they would sooner or later adopt the dominant form in the population, and the vestigial forms would be lost over time. In the same way, if the loners did not have low in-degree, i.e. they would have more impact on others, the variant they hold would enter the population too often for norms to be established. Both scenarios would eventually be accomplished despite the probabilistic nature of the in-degree-biased voter model. Also, while it can easily be imagined that some loners could be connected to other networks (not modeled here), which would make them in many ways similar to Eckert’s in-betweens (1.3.4), this type of connection would not change the dynamics in our simulations. No matter how loners acquire variants that they hold, i.e. by safe-keeping old or adopting new variants, the most important point is that they would never let go of them. Loners’ safe-keeping action stabilizes old forms in the network, which ultimately makes the dynamics evolve in ways observed in our simulations.21

Loners modeled in our simulations are not the only type of peripherals documented in the sociolinguistic literature (Chambers, 2009:100–108 for a review). Aspirers, outsiders longing to become members of a group, or interlopers, outsiders who move from one area to another might have a different effect on the diffusion dynamics modeled in this paper. In our simulations, we focused on a single social network with a complex community structure where one type of peripherals, referred to as loners, had the greatest influence. Our results depend entirely on the fact that these individuals do not change their behavior with respect to the eight variables in competition in the network. If they did, over time the previous norms would be eroded and one variant would remain fixed forever. Loners’ very few incoming links were sufficient to ensure that their influence could travel through the network to all other agents, which ultimately makes the loners’ network position more important than what it might seem. Their indirect influence, or “reach”, extending over the rest of the network is due to the structure of the heterogeneous social network, in particular the three properties described earlier as: small diameter (members in close reach), high clustering (tight-knit community) and scale-free degree distribution (very few influential and many isolated members). This explains the remarkable fact that only 2–4% of loners were sufficient to create dynamics that can propagate a rare form to all other agents over time. The conditions of the spread of rare, or functionally disadvantageous, linguistic innovations have also been investigated in agent-based computer simulations. One of them is particularly relevant for our study.

4.2.3. The threshold problem of language change

Nettle’s (1999) study investigated the so-called threshold problem of language change, i.e. variants only becoming “fixed in a language if they can pass a threshold of frequency which in the early stages they never have” (Nettle, 1999:98). In his model, agents arrayed in a grid-like social network had to choose between two competing linguistic alternatives, p and q. They passed through five “life-stages”, analogous to age, in which they could variably influence, and be influenced by, others. An agent decided which variant to adopt based on the number and social status of individual agents and the social distance22 between agents already using the variant. In addition, agents also had some functional bias, i.e. perceptual or cognitive preference, towards a particular linguistic variant. Nettle showed that minority variants were adopted only when some speakers were socially much more influential than others, and referred to these as “hyper-influential” agents. Although functional bias played a large role in the selection of a variant, without heterogeneous social influence the population still could not overcome the threshold problem, i.e. passing the point at which a given variant begins to propagate and end up being used by the majority.

The stochastic nature of Nettle’s model is important to stress in light of the present study. Nettle simulated, for the first time, a noisy and fluctuating diffusion process – one variant selected, propagated or failing to do so – that is typical of realistic scenarios reported by speaker-oriented approaches to language change. Our simulations have also been successful in modeling such a non-deterministic process. One major limitation of Nettle’s work was that the social network structure it assumed was unrealistic. As we argued above, real-life social networks are clearly not regular grids where individuals only listen to their most immediate neighbors.23 Nettle’s study was the first to demonstrate, however, that agents endowed with a statistical language model could select and propagate even rare variants, or variants with a low functional bias, to all agents. Functional biases determine the direction of a change, but “changes are adopted because some speakers are much more

21 Books can also be thought of as such immutable repositories of vestigial forms that might be selected and diffused again in the community (p.c. Richard Sproat).
22 These social factors were motivated on social–psychological grounds, such as Social Impact Theory (Nowak et al., 1990).
23 Prior to Nettle’s work, “small-world” properties of some social networks were recognized (Milgram, 1967), but structural properties of large networks were still unknown (Watts and Strogatz, 1998; Barabási and Albert, 1999).
influential than others as social models” (Nettle, 1999:112). How this was achieved remained an open question at Nettle’s time. Our simulations can now provide an answer that brings support to Nettle’s intuition: the key to diffusion is differential social selection. In a socially heterogeneous population, even vestigial or rare forms can be adopted as norms, provided that loners are allowed to preserve it, hubs are influential enough to propagate it, and agents can unanimously agree selecting it.

4.3. The nature of linguistic conventions

Our findings are consistent with the view of linguistic conventions advocated by Croft (2000:175) who states: “There would of course be a few stubborn diehards who insist on using an alternative form even when almost everyone has abandoned it; but we must allow for the existence of such types in a realistic theory of convention in language use.” Our simulations show that those who do not adhere to specific conventions of language use might, indeed, not be a nuisance in an otherwise perfect scenario of diffusion, but rather could be an essential component of it. In a socially realistic heterogeneous influence network, future changes do not occur without the preservation of marginal variants by outliers. This dynamic view of individual network roles, therefore, runs against an all or nothing view of language change, according to which “one cannot assume that a change entirely ends, if there is a residue of relic forms or vestigial variants” (Trudgill quoted in Croft, 2000:185). Of course whatever happens to relic forms in languages over time does also depend on many other sources of influence in the speech community. There could be numerous amplifiers and barriers to diffusion, some of which are related to the variant, while others might have to do with agents’ propensity to accommodate trends and fashion in language use (Rogers, 1983; Wejnert, 2002 for parallels in sociology). In principle, loners’ actions of preserving vestigial forms can also be counteracted by innovation processes in the community. While such processes have not been modeled in this paper, it has to be emphasized that more fine-grained studies of the conditions of establishment of linguistic conventions are needed to take into account both the rate of innovations and the preservation of old forms in the same community.

4.4. Influence, prestige, and popularity

Finally, we have also investigated the impact of individual social status (relative prestige or popularity) on the diffusion process. When we modeled it as agents’ popularity with others, using in-degree biased voting (2.3), this rule specified that agents had to be probabilistically biased for adopting those individuals’ variants who were also popular with other agents. We then compared this, so-called, update rule to no bias (imitate anyone) and to random bias (imitate a randomly designated individual regardless of others’ opinion). The simulations showed that the community-wide adoption of an innovative variant as a stable, novel norm only occurs when: there are influential hubs and isolated loners in the network, and all agents (except loners) preferentially select a linguistic variant held by better connected neighbors, i.e. those who enjoy the highest relative prestige in the local community. Lack of sensitivity to individual social influence (no bias) led to lack of convergence at the population level. Agents’ random sensitivity to others’ popularity/prestige (imitate a randomly designated individual) led to some convergence, but with considerable fluctuation. We interpreted this as a lack of widely shared assumptions of what agents’ variants others should “value” as prestigious in the community. Thus, our findings indicate that only when widely popular local agents are also centrally-connected in the social hierarchy, i.e. they are part of an influential social class or group, can they become powerful propagators of linguistic innovations and stabilizers of emerging norms. In other words, structural influence and relative prestige jointly lead to the emergence and maintenance of norms. Of course, these two components of social valuation are not entirely separate: in our model agents are ‘pre-wired’ to possess more in-degree than others (structural influence) and their variants are copied based on the weighting of their in-degree (relative prestige). Thus, the latter is indirectly based on the former. Note, however, that this is the result of modeling a closed social network, in which agents do not migrate, i.e. their network position is not treated as an independent variable in the model. The network structure modeled in the paper mirrors real-life social settings where the likelihood of setting trends (relative prestige) is closely tied to agents’ class- or caste-based privileges (structural influence). We manipulated these two components of social valuation to find out how individual preferences (random or no bias) and the absence of major and minor structural influence (hubs and loners, respectively, taken out) affect the diffusion dynamics in this type of network. Future studies will show whether agent mobility, among others, yields different results. With the above limitations in mind, we suggest that Bloomfield’s formulation of relative prestige of social groups should be understood as the composite of both network structural characteristics and unanimously shared preference for what is socially desirable to imitate in others’ language use.

These findings also speak to the role of inter-personal accommodation in language change, and are consistent with the findings of a recent computational study that targeted the role of social structure and social valuation in new dialect formation. Baxter et al. (2009) have addressed Trudgill’s (2004) claim that the emergence of new dialects in isolated settings (e.g. New Zealand English) can be explained purely in terms of frequency of occurrence of particular variants and of frequency of interactions between speaker/hearers in the community (Trudgill, 2004, 2008a,b). Based on a mathematical model of the usage patterns of two variants documented in the ONZE project (Gordon et al., 2004), Baxter et al. (2009) found that the emergence of isolated new dialects is very likely non deterministic and depend on social factors that “operate in
conjunction with the frequency effects of pure accommodation (neutral evolution), social network effects (neutral interactor selection), and social valuation of one’s interlocutors (weighted interactor selection)" (Baxter et al., 2009:291). Although our model did not extend to frequency, our results bring further support to the importance of social selection in the emergence of new community norms.24 Indirectly, our results also underscore views of linguistic accommodation as behavioral coordination (Keller, 1994) informed by sensitivity to “the social marking of linguistic variants” (Holmes and Kerswill, 2008:275).

5. Conclusion

In this paper, we brought support for the essential role of communication density and relative prestige in the spread and establishment of linguistic innovations as norms. As first outlined in Bloomfield’s (1933:345) principle of communication density, the spread of linguistic features does depend upon social conditions. The two main poles of density in influence networks, equivalents of leaders and loners in empirical studies of speech communities, play different roles in the diffusion process. Leaders advance on-going change, while loners are repositories of variants considered old or new depending on the current state of the rest of the population. The lack of highly-connected agents, structural equivalents of leaders in empirical studies, results in failure of appearance of norms. The absence of isolated individuals, or loners, leads to lack of innovation. No norm emerges when members ignore, or do not share a common view of, the popularity of individuals they imitate. Thus, innovations spread following an S-curve and stabilize as norms only if: (1) the network is socially heterogeneous (scale-free) and comprised by tight-knit communities (high clustering) that keep their members in close reach (small diameter), and (2) members imitate the language use of those whom all perceive as the most popular.

We modeled the cumulative adoption of novel variants over time as a probabilistic and noisy path of change. This result aligns well with many documented cases of social and geographical spread of sound changes. Demographic factors, such as population density, and linguistic factors, such as the coexistence of old and new forms, have been previously dismissed as impossible to fully document based on historical data (Wolfram and Schilling-Estes, 2003:715–717 for a review). It appears, however, that the systematic study of individual social dynamics underlying language change, which Bloomfield thought could not be realized, is emerging. In recent years, computational methods have become a powerful new tool that allows experimentation with possible worlds of social agents. If used judiciously, these tools can help experiment with the necessary and sufficient conditions for the emergence of various large-scale social phenomena, among them the dynamic process of language change (Wichmann, 2008). In these systems, analogized to speech communities, small incremental changes at the individual level, such as speaker/hearers interactions, can have large consequences and lead to complex emergent phenomena over time (Dras and Harrison, 2003; Harrison and Rainmy, 2007). Similar to our model, many computational approaches treat the sociolinguistic system as such a complex adaptive system (Baker, 2008; Baxter et al., 2009; Troutman et al., 2008), and numerous recent simulations used realistic social networks to study the dynamics of language change (Ke et al., 2008; Minett and Wang, 2008) and societal bilingualism (Sankoff, 2008).

Many vexing questions remain. Our simulation still lacks realism at many points. It is fair to say, for instance, that speakers of large speech communities do not live and die in the same social network where no population change and transmission to future generations occur. This means that computational models of emergent sociolinguistic phenomena will have to pay attention to agents’ mobility and strong (e.g. family) ties to other agents. Furthermore, since the preservation of vestigial variants co-occurs with on-going innovation in the community, it would be particularly useful to investigate both within the same model. Another question might be raised with respect to the selection of competing variants. Would the results of our simulations change if agents used a variant variably rather than categorically during the simulation? While there is empirical evidence in the variationist sociolinguistic literature for both categorical and variable uses of forms, the choice of either of the two would always depend on the type of linguistic phenomenon to model. Baxter et al.’s (2009) model, for instance, is a promising approach to phonological change, as agents maintain a probability distribution over several competing variants, altered through these agents’ repeated interactions. The model presented in this paper, on the other hand, would be more appropriate for categorical adoptions of new words and expressions. Within the validity of a limited theoretical approach, our model’s goal was to bridge the gap between empirical findings on individual speaker roles in small communities and very large societies. Highly polarized social spaces, such as urban working-class neighborhoods or clusters of adolescent peer groups in school settings, for instance, would be analogous to the type of population simulated in this study. While still many questions remain, these and other linguistically meaningful agent-based simulations could help exploring the empirical foundations of a theory of language change, i.e. “determining the set of possible change and possible conditions for change which can take place in a structure of a given type” (Weinreich et al., 1968:101).

Acknowledgements

We thank Professors Hans H. Hock, José I. Hualde, Richard Sproat, and the audience of the NWAV 37 conference in Houston for their comments on earlier versions and presentations of this paper. We are especially thankful to three anonymous reviewers for their insightful reviews and constructive criticism. All shortcomings are ours.

24 The fact that we see robust fixation of norms in our model while fixation in Baxter et al.’s (2009) model could not be reached within the time frame New Zealand English had emerged would be worth investigating in future work.