

RESEARCH

Bipartite network structures and individual differences in sound change

Robin Dodsworth

North Carolina State University, US
robin_dodsworth@ncsu.edu

This paper assesses the influence of social network structure, and the role of the individual, in shaping the loss of the regional vowel system in the Southern U.S. city of Raleigh, North Carolina. The entire front vowel system, including monophthongal /aɪ/ as in *ride*, is shifting toward the national standard. Previous network studies in sociolinguistics have focused on individual-level network characteristics, such as integration in dense local networks or contact with speakers from different neighborhoods or ethnic groups. By contrast, the Raleigh study focuses on individuals' positions in the community network structure as represented by a bipartite network of people and the schools they attended. Bipartite networks indicate social proximity between people via their shared participation in an event or organization.

With a 189-speaker sample of Raleigh natives, the network measure of *structural equivalence* offers a view of Raleigh's community network structure and of the individual's role in advancing the shift away from the Southern vowel system. Structural equivalence is the extent to which nodes inhabit similar positions within a social network. In this case, it describes the extent to which pairs of speakers attended the same schools. A distance matrix containing each pair's network proximity is used to predict speakers' linguistic similarity. The role of the individual and of the social indexicality of Southern variants is considered in the context of aggregate patterns of variation.

Keywords: social networks; indexicality; Southern Vowel Shift; social class; vowel change

1 Introduction

Most sociolinguistic network studies aim to uncover the aspects of daily interactional life and characteristics of speakers' ego networks that influence their production of local linguistic variables. These studies often focus on the local social and stylistic meanings of linguistic variables, as individuals use them, rather than larger-scale sociolinguistic patterns as in the tradition of Labov (1966). In our recent work, we have taken a different approach to sociolinguistic network analysis: rather than investigating speakers' daily interactional practices or ego networks, we borrow quantitative techniques from sociology and adjacent fields in order to explore the relationship between *aggregate network structure* and linguistic variables in the dialect contact setting of Raleigh, North Carolina. In the current analysis, however, we ask whether and how some of the aggregate patterns that we have found in recent work can contribute to an understanding of the social indexicality of linguistic variables in their local context and ultimately the role of the individual in linguistic change.

2 Sociolinguistic network analysis

Most previous sociolinguistic network studies test one of two hypotheses, broadly conceived. The first is that speakers who are well integrated into a local community are more likely to fully acquire or maintain a local vernacular than speakers peripheral to the community. For example, Labov (1972) investigates the relationship between adolescent peer group membership and use of core features of African American English in the neighborhood of Harlem in New York City. One conclusion is that peer group membership affects not only *rates* of vernacular variants, but even the internal constraints: for the members of two local social groups, phonological context is the stronger constraint governing (t/d) deletion; but for the boys who are not members of the social groups, as for many White speakers in New York City and elsewhere (Guy 1980), morphological status is the stronger constraint. A further network-based look at the linguistic consequences of peer group membership in Harlem involves the Jets, one of the local social groups. 36 boys, all members or would-be members of the Jets, are classified as core, secondary, peripheral, and “lame” (non-) members on the basis of observation as well as friendship naming by the boys. Core members delete the copula with the highest rates, followed by secondary members, then peripheral and lame members, in that order (Labov 1972: 279, Table 7.8). Other linguistic variables, however, do not clearly pattern with network integration.

Cheshire (1982) similarly asks adolescent boys who hang out on a playground in Reading, England, to name the others on the playground who they considered friends. This yielded three categories of boys: core members, secondary members, and non-members. Seven out of nine vernacular linguistic features showed higher frequencies among core members than among secondary members.

In Milroy’s (1987) study of three lower working class neighborhoods in Belfast, Northern Ireland, the central question is whether speakers’ integration into the local community network, as determined not by core vs. peripheral membership but by measures of density and multiplexity of their ego networks, correlates positively with their use of local vernacular features. Milroy draws upon anthropological research to form the hypothesis that dense, multiplex networks in working class Belfast communities enforce local linguistic norms, even in the face of stigma within the broader urban community; therefore, speakers who are well integrated into those communities should show the highest rates of vernacular variants. In line with Labov’s (1972) study of vernacular use and hyper-local peer group membership in Harlem, Milroy proposes that what matters most for a speaker’s use of local linguistic norms is not the overall density of his/her ego network, but rather the density of certain areas of the ego network, such as family, work, or neighborhood ties. Accordingly, Milroy indexes community integration via a “Network Strength Scale” built on measures of a speaker’s interaction in particular domains of everyday social life: family, work, and leisure activities (see also Lippi-Green 1989 for a similar focus on domains of everyday life). Significant network effects emerge for /a/ backing among older women, /ð/ deletion for both older and younger women, and, for older men, the frequency of a low, rather than mid, vowel for /ɛ/ in monosyllabic words. Labov (2001) reanalyzes Milroy’s Belfast data using multivariate regression and finds that four of the linguistic variables show significant effects of network: three for females (/ð/ deletion, /a/ backing, and /o/) and one for males (/ɛ/ lowering).

The predominance of network effects for women, rather than men, in the Belfast data underlies Labov’s decision to carry out the analysis of network factors for men and women separately in Philadelphia. Labov (2001) uses two network indices. The first index encodes how many neighbors on the local block a speaker reports interacting with in daily life. The second index represents the percentage of the speaker’s named friends who live off

the block. For 6 out of 8 linguistic variables, one of the two network indices or their combination has a significant, additive effect (i.e., not erasing the effects of other social factors). The female leaders of linguistic change in Philadelphia have the highest proportion of friends off the block and/or the highest density of interaction on the block; that is, the female speakers who have the most intense social interaction are the most likely to adopt linguistic innovations.

The second hypothesis often tested in sociolinguistic network analysis is that speakers with significant contact with another social group, usually another ethnic group, are more likely to use linguistic features associated with that group, relative to speakers with little inter-group contact. Ash & Myhill (1986), for example, assess the linguistic effect of ethnic diversity in speakers' lives and personal networks. The network data come from speakers' answers to questions about four aspects of their social interaction:

[...] the racial composition of the speaker's present neighborhood, the racial composition of the speaker's high school, the number of friends the speaker presently has from the "opposite" ethnic group, and the number of spouses and/or lovers of the "opposite" ethnic group that the speaker has had. (Ash & Myhill 1986: 33–34)

Inter-ethnic contact scores are assigned to Black, White, and Puerto Rican speakers in Philadelphia, allowing the researchers to ask whether inter-ethnic contact correlates with rates of 10 linguistic variables associated with African American Vernacular English. With respect to most of the phonological variables, Black speakers who have little contact with White speakers show the highest rates of the AAE variants, followed by Blacks who have regular contact with Whites, followed by White speakers with more and less inter-ethnic contact, respectively. For the grammatical variables, however, Blacks with little inter-ethnic contact show high rates of the AAE variants, and all other groups show very low rates. The different patterns that emerge for the phonological vs. the grammatical variables suggest that the relationship between language and network is not merely an exposure effect wherein greater familiarity with White variants correlates with their greater use among Black speakers. Instead, there is apparently strategic use of Black and White linguistic variants, hinging on the variants' prestige or salience.

Edwards (1992) investigates both community integration and inter-ethnic contact in a study of Detroit AAE speakers, asking each speaker five questions having to do with whether speakers have friends, family, interactions, and workplaces in the immediate neighborhood, and five additional questions having to do with attitudes toward the local neighborhood and its culture. These 10 questions constitute a Vernacular Culture Index (VCI), which is used to measure speakers' integration in the ethnically homogeneous local community. Having jobs outside the neighborhood tends to correspond to having contact with White speakers, while having jobs only in the neighborhood tends to mean insulation from contact with Whites. (See also Hoffman & Walker 2010, who employ an Ethnic Orientation questionnaire to assess various aspects of speakers' linguistic practices, attitudes toward ethnic culture, and experiences related to ethnic culture.) Three linguistic variables associated with AAE, (ay) ungliding, coda /r/ deletion, and copula absence, correlate with the elements of the VCI. The individual elements of the VCI showing the strongest correlations with the linguistic variables are one of the "physical integration" elements ("Most of the jobs I have held have been in this neighborhood") and another element having to do with social integration ("I do not have White friends with whom I interact frequently"). Edwards explains the correlations between the linguistic variables and the VCI as arising first from regular social interaction, which produces shared group identity and ultimately shared linguistic norms:

[Social network] theory proposes that the network multiplexity that results from frequent daily interactions fosters strong local affiliations that typically lead to linguistic focusing in the sense that Le Page and Tabouret-Keller (1985) used the term. This means that certain linguistic items become normative in the community. Thus, multiplex social networks are norm-enforcing social mechanisms. This line of argument leads to the conclusion that the [r]-deletion, [ay] monophthongization and zero copula rules are social norms conformed to by neighborhood residents who are strongly integrated into or positively oriented toward their neighborhood. These linguistic behaviors can be seen as manifestations of a sense of group identity or group solidarity. (Edwards 1992: 108)

Cheshire et al. (2008) investigate the role of the ethnic composition of friendship networks in the diffusion of Multicultural London English features. In Hackney, an inner London area, young Anglo speakers are showing evidence of the progression of a regional chain shift that involves the TRAP, STRUT, FOOT AND GOOSE vowels. Non-Anglo young speakers in Hackney have a higher STRUT vowel and a fronter GOOSE vowel than the Anglo speakers overall. Both the Anglo and non-Anglo young speakers were asked about the ethnicity of each of their close friends, and were then assigned a score indicating the percentage of their friends with a different ethnicity. The vowels that differ significantly by ethnicity of network are FACE (which is becoming more monophthongal) and GOOSE; in both cases, young non-Anglo speakers lead the change, followed by young Anglo speakers with multi-ethnic networks, followed by young Anglo speakers with predominantly Anglo networks. Khan (2006) similarly finds, in Birmingham, England, that White speakers with high percentages of inter-ethnic ties are more likely to produce linguistic variants associated with non-White ethnic groups.

Many early sociolinguistic network studies were motivated by the belief that local interactional factors were more important than economic differences in the context of daily life. For that reason, the Labovian speech community model was seen as irrelevant. For example, Lippi-Green (1989) reports that in the dairy farming village of Grossdorf, Austria:

[I]t is not occupation that determines one's place in the hierarchy, but rather the degree of integration into the established structures. A successful farmer from a well-established clan may not make as much money or build as nice a home as someone well-placed in the provincial government, but it is the farmer who is more likely to be elected to the Council. It is not so much a matter of class or status, but who you know, and who knows you. (Lippi-Green 1989: 216)

Milroy (1987: 12–17) similarly offers that people in many working class communities are first and foremost *locally* oriented, having low levels of geographic and social mobility and close relationships with neighbors. In the course of ethnographic observation in Belfast, Milroy finds that community members prioritize local community cohesion above economic status. Therefore, linguistic variation can have more to do with local neighborhood residence, and neighborhood status, than with economic characteristics or with mere exposure to linguistic variables. People in the same dense, multiplex network are exposed to the same linguistic variants, but they have agency in using those variants to index membership in a localized network.

Recent simulation-based research offers evidence that the structure of social interaction, and thus exposure to linguistic variants, influences linguistic change, social identity notwithstanding. For example, Stevens, Harrington & Schiel (this volume) use an agent-based model to show that interaction among speakers can promote sound change in the

context of asymmetric phonetic overlap between two sounds. Other agent-based modeling work and related simulation-based work (e.g., Fagyal et al. 2010; Garrett & Johnson 2013; Stanford & Kenny 2013; Harrington & Schiel 2017) similarly supports the perspective that interaction between speakers propagates linguistic change. Taken together, the simulation-based studies and sociolinguistic network studies involving real communities suggest that both the social *meaning* of linguistic variables and speakers' *exposure* to them (resulting from patterns of social interaction) influence linguistic variation and change. While sociolinguists have discussed the ways in which a linguistic variable's social meaning and its distribution in a social network are mutually influential (Edwards 1992; Milroy & Milroy 1992) this remains an area for future research (cf. Eckert this volume).

3 Reconsidering sociolinguistic network data and methods

In order to motivate the kind of network data that we use in the present study, we first make a few observations about the usual data and methods in previous sociolinguistic studies.

In previous studies, the network data are principally speakers' self-reports about who their friends are or who they interact with (or what types of people they interact with, e.g., people in the neighborhood vs. elsewhere), in some cases supplemented by the investigators' observations. In addition to being self-reported and thus prone to inaccuracy, the data derived from this approach are *ego-centric*, having to do with the characteristics of individual speakers' first-order networks rather than with the community's overall network structure (Sharma 2017). An exception, of sorts, is Bortoni-Ricardo's (1985) analysis of migrants to Brazlândia, Brazil. Each speaker was asked to name the three people they most often talk to outside of their homes, and then a community network structure was built out of the speakers plus the people they named. But even this study's data are entirely self-reported, first-order network data.

On a theoretical level, the near-exclusive focus on first-order network ties rather than community network structure reflects the belief that some characteristics of ego networks, such as the presence of inter-ethnic ties, promote or inhibit speakers' adoption of new or non-local linguistic variants. But the more pressing and more practical reasons are that 1) the researcher typically cannot directly observe more than a small subset of interactions among community members, and so self-reported data are essential; and 2) sociolinguistic samples are (for good reasons) often too small or not representative enough to allow a good view of the overall community network structure even if it could be observed directly.

Another salient characteristic of previous studies is the common focus on small, dense neighborhood networks, to the exclusion of speakers outside these networks. This practice of focusing on hyper-local communities carries the advantage of allowing a deep, perhaps even comprehensive look at the community, and it also avoids the difficulty in collecting and representing network data for speakers with relatively loose-knit personal networks (cf. Milroy 1987: 197–198). It also means, however, that we end up with little information concerning how the linguistic practice of members of the dense local network fit into the linguistic distribution of the broader community, and we also do not make progress in developing strategies for studying network effects in larger, lower-density community networks.

A further methodological observation is that across studies in which speakers self-report information about their social contacts, the ways in which the data are used can lead to entirely different views of speakers' network characteristics. In the case of Harlem (Labov 1972), observation and friendship naming result in a proposed core/periphery structure, and individuals are located categorically within the structure. In the cases of Belfast and Philadelphia, individuals lie along a continuous axis and no particular group network

structure is proposed. Distinct from both is Gal's (1992) approach in Oberwart, Austria. Speakers were asked to name their recent contacts, but rather than building a model of the community network out of these data, Gal constructs a "peasantness of network" index because this is compatible with the research question (i.e., whether speakers use Hungarian or German). The diversity in researchers' handling of network data has been an asset insofar as different representations of network data are compatible with different linguistic questions, community types, and sample sizes. But this diversity has also made it difficult to directly compare conclusions across studies and to form generalizations about the relationship between language and social networks.

Finally, in most previous studies, the network criteria are selected by the researcher specifically for that community. Although there is usually an effort toward using network characteristics found to be important in other communities (either in linguistic or sociological or anthropological studies; cf. Milroy 1987), the criteria are usually tailored to the community in a way that maximizes their relevance for that study but obstructs the generalizability of the study's conclusions. For example, Labov's (2001) use of criteria having to do with friendships on the block and off the block are useful when the linguistic question is along the lines of 'what is the individual's level of interaction in the immediately local social space'. But we would not expect this question to be equally important in our assessment of the Southern Vowel Shift in Raleigh or in most present-day urban environments because day-to-day social interaction has wider and more varied geographic reach, even if some neighborhoods remain locally cohesive.

A final methodological observation is that, with the partial exception of Bortoni-Ricardo (1985), sociolinguistic network studies have tended not to engage contemporary quantitative social network procedures that have gained currency in sociology and related fields. This is not so much a shortcoming as a reasonable consequence of differences between sociolinguistics and sociology: in sociolinguistics, our sample sizes are often necessarily smaller and our questions are different. Nevertheless, as we will argue here, some contemporary social network methods can serve sociolinguistic questions.

The Raleigh study, an ongoing investigation of the retreat from the Southern Vowel Shift, is in part an effort to take sociolinguistic network analysis in some new methodological and empirical directions. Our primary goals are, first, to represent speakers' positions in the community's overall network structure rather than looking at self-reported ego network data; second, to use data and methods that can be replicated across communities, thus facilitating generalization; and third, to engage with contemporary social network methods in other disciplines.

The Raleigh study, in contrast with many previous sociolinguistic network studies, is not an effort to understand the linguistic behavior of individuals in a small community; instead, it deals with a large urban speech community as in Labov (2001). Second, the use of network data in the Raleigh study is not motivated by any difficulty in representing social class (cf. Milroy 1987; Lippi-Green 1989). Raleigh's economic structure lends itself to an occupational model that resembles those in other urban sociolinguistic studies (though see Forrest & Dodsworth 2016 for a more complex occupational paradigm). A central goal, in fact, is to assess simultaneously social network and social class influences on linguistic variation. Finally, our focus here on the community's social network structure, rather than characteristics of ego networks, rests on the assumption that broad patterns of interaction within a community can influence the spread of linguistic variables. This assumption is supported by the simulation-based work discussed in section 2, such as the agent-based model of sound change in Stevens et al. (this volume).

In the current analysis, however, we ask whether and how the aggregate, relatively abstract trends that we have found in recent work (Dodsworth & Benton 2017; Dodsworth & Benton forthcoming) can reveal individual agency in adopting or resisting linguistic changes in the context of the local community. The strength of many previous, small-scale sociolinguistic network studies lies in the view they construct of the individual's place in the local society and the individual's symbolic use of linguistic variables. At the same time, individual-level cognitive differences such as working memory or production planning, as discussed by MacKenzie (this volume), could open the door to community-level change, as could articulatory differences among individuals (Baker et al. 2011). Aggregate network analysis of the kind we are after will be most useful if it can not only uncover large-scale patterns, but also elucidate the role of the individual and identify social and cognitive characteristics of individual leaders of sound change, in the context of those patterns. Labov (2001) finds that the leaders of new sound changes in Philadelphia are women central in the socioeconomic range who interact socially with a lot of other people, either on the neighborhood block or further away. In the context of similar generalizations about the Southern Vowel Shift in Raleigh, we aim to identify exceptional individual speakers and explore their social characteristics, including their network positions.

4 The Southern Vowel Shift in Raleigh, NC

Our data are from a subset of a conversational corpus of speakers native to, and living in, Raleigh, North Carolina. Data collection began in 2008 and is ongoing. Raleigh lies in the Southern dialect region of the U.S. (Labov 1991; Bailey 1997; Thomas 1997; 2001; Tillery & Bailey 2004; Labov et al. 2006), and most Raleigh natives born before about the third quarter of the 20th century have vowel systems characterized by the Southern Vowel Shift (Dodsworth & Kohn 2012; Forrest & Dodsworth 2016; Dodsworth & Benton 2017; Dodsworth 2018). The triggering element of the Southern Vowel Shift (SVS), according to Labov et al. (2006), was the monophthongization of the diphthong /ai/ as in *high*. In some Southern regions, monophthongal /ai/ may occur in any environment, but Raleigh belongs to a region in which it occurs only before voiced consonants or in open syllables (*high* and *tide* but not *tight*). The second element of the SVS is the lowering of the /e/ nucleus, so that the vowel in *state* or *take* is a diphthong with a central nucleus and a high front offglide. The nucleus of the other mid-front vowel, /ɛ/ as in *head* or *step*, is raised and fronted, resulting in a diphthong with a peripheral nucleus and central offglide. The high front pair, /i/ and /ɪ/, underwent an analogous but weaker shift in Raleigh: the /i/ nucleus is slightly lower and fronter than in non-Southern regions, and the /ɪ/ nucleus is higher and fronter. (The high front pair shifted more dramatically in other regions; Labov et al. 2006.) Finally, the /æ/ nucleus is raised and fronted, resulting in a diphthong and sometimes a triphthong.

Since the middle of the 20th century, Raleigh as a community has been in retreat from the Southern Vowel Shift, toward a regionally unmarked vowel system. One of the reasons is that a technology industry research center, Research Triangle Park, was established next to Raleigh in 1959, and it became successful especially after IBM opened a headquarters there in the 1960s. Raleigh's population and geographic area grew quickly as thousands of well-educated technology sector workers moved to Raleigh from New York and other non-Southern regions (Dodsworth & Benton 2017). We expect that the shift away from the SVS was driven by the mixing of Raleigh natives and non-natives in schools, leading to the formation of stable new community linguistic norms (Kerswill & Williams 2000).

From a social network perspective, the trajectory of vowel shift across time in Raleigh is potentially interesting for two reasons. First, the areas of Raleigh that first grew with the incoming migration were to the north and west of downtown (closer to Research Triangle Park), leaving the eastern and southern neighborhoods relatively insulated from dialect contact. Second, the incoming migrants were predominantly affluent (Rohe 2011), and so we can expect that they had disproportionate contact with affluent Raleigh natives than with working class Raleigh natives, in schools and elsewhere in the community.¹

The data in the present analysis come from a 189-speaker subset of the Raleigh corpus. All of the speakers in the present subset are White. The conversational interviews were transcribed and force-aligned, and the vowels were measured automatically and hand-corrected. Token counts for each of the six elements of the SVS appear in Table 1.

Figure 1 shows apparent-time change for each of the six variables. Each circle represents one speaker’s mean Z2–Z1 (Lobanov-normalized F2–F1) measured at 25% of the vowel’s duration (the nucleus), except that /ai/ was measured at 75% of the vowel’s duration in

Table 1: Post-correction token counts.

Vowel	N
/i/	12,010
/ɪ/	12,447
/e/	15,827
/ɛ/	13,285
/æ/	17,729
/ai/	16,637

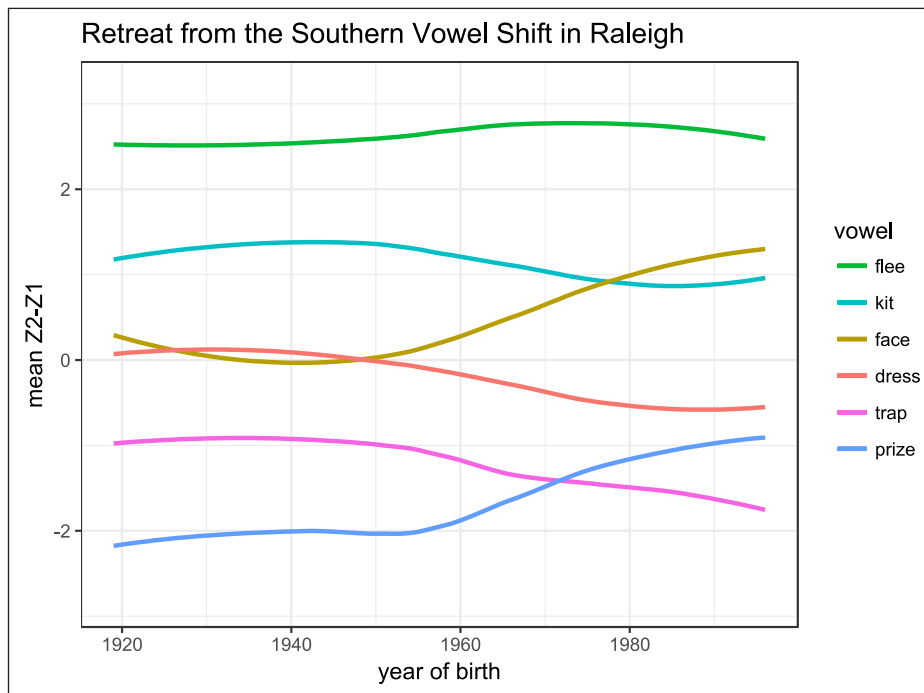


Figure 1: Apparent-time change in Raleigh’s front vowel system. Each circle represents one speaker’s mean.

¹ For a great deal of further discussion on these points, see Dodsworth & Benton (2017) and especially Dodsworth & Benton (forthcoming).

order to capture the glide. Higher values of Z2–Z1 correspond to higher, fronter positions in the front vowel space. The front vowel system was stable among speakers born during the first half of the 20th century. As Research Triangle Park, the technology industry, and Raleigh’s population grew during the second half of the century, all of the front vowels gradually retreated from their SVS positions. The high front tense vowel, /i/, was never dramatically Southern-shifted in Raleigh, and so its change across the 20th century was slight compared to that of the other vowels. In particular, the /e/ nucleus shifted from being indistinct from the /ɛ/ nucleus to occupying a more peripheral position than the /i/ nucleus. The /ai/ glide, in voiced and syllable-final contexts, followed a strikingly similar apparent-time trajectory, ultimately becoming roughly as peripheral as the /ɛ/ nucleus.

As illustrated for /e/ in Figure 2, white collar speakers generally led the retreat from the SVS and maintained the lead throughout the second half of the 20th century. Nevertheless, blue collar and unskilled white collar (e.g., secretary or office manager) speakers shifted in the same direction as white collar speakers, and at similar rates.

We investigate the social factors associated with the retreat from the SVS within each of three generations (Figure 3); the generations are defined in relation to the establishment of Research Triangle Park and in relation to a change in the public school system.

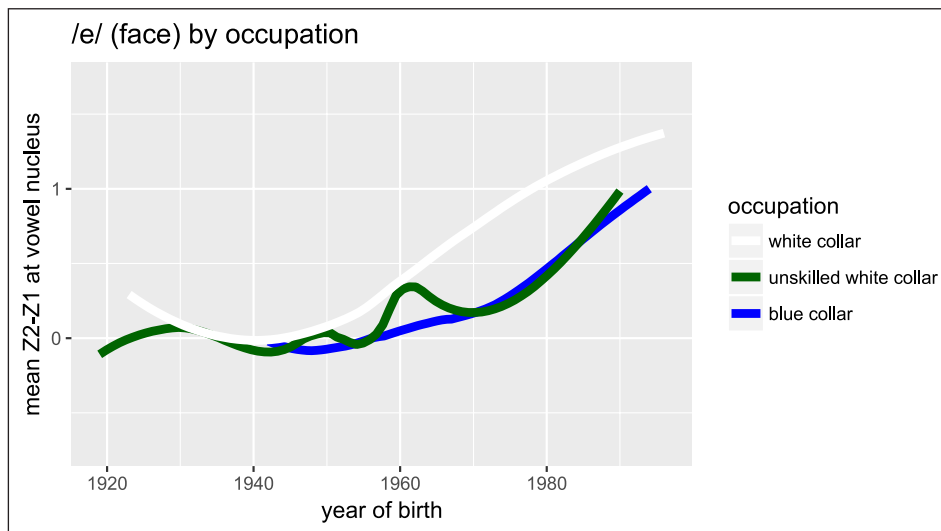


Figure 2: Apparent-time change for /e/ by occupation. Each circle represents one speaker’s mean.

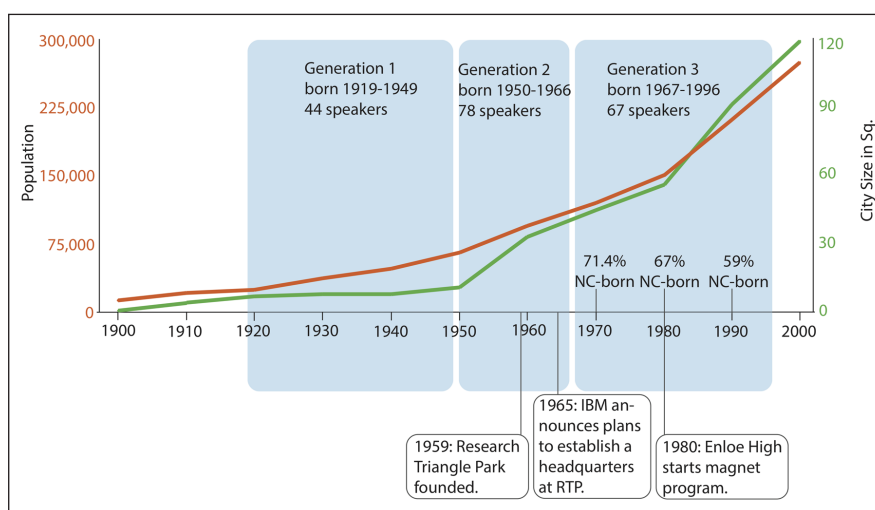


Figure 3: The 3 generations defined in relation to Research Triangle Park and the school magnet program.

In Generation 1, Raleigh wasn't yet a contact setting in the current sense because the speakers grew up prior to the large-scale migration from outside the South. Speakers in Generation 2 grew up in the midst of incoming Northerners and the growing technology industry. Speakers in Generation 3 also grew up in the midst of growth, with the added dimension that Raleigh's school system was affected by the beginning of a magnet program at one high school, followed by others. Magnet programs in the U.S. school system allow qualified students living outside a given school's district to attend the school in order to take part in a program focused on an area of study such as music or foreign language. As a consequence, high school students in Generation 3 no longer necessarily stayed in their home neighborhoods for school.

The social factors of age, sex, and occupation vary by generation in their effects on the front vowel system in Raleigh (Forrest & Dodsworth 2016; Dodsworth & Benton 2017). We model occupation with a three-way paradigm: white collar/college degree required, white collar/no degree required (e.g., secretary), blue collar. In the present sample, we have found (Dodsworth & Benton forthcoming; full models in Appendix 2) that there are no significant social effects in Generation 1. This is consistent with the fact that the community, in the aggregate, had not yet started to retreat from the SVS relative to earlier points in apparent time, even if some individual speakers were variably adopting non-Southern vowels. In Generation 2, however, white collar speakers are significantly less Southern than others for three of the six vowels (/i/ and /e/), and older speakers are significantly more Southern for /e/ and /æ/, even though this generation spans only 16 years. In Generation 3, white collar speakers continue to lead for in the case of /e/, and year of birth remains a significant factor for /e/, /ɛ/, and /æ/. In Generation 3, females lead the strengthening of the /ai/ glide, and Generation 3 females are also less Southern than males in the case of /æ/. These effects are not unexpected for a dialect contact setting in which children with different native dialects are coming together at school and gradually establishing a new variety from diverse input (Kerswill & Williams 2000). In fact, Raleigh, as an urban dialect contact setting, can be considered an ideal community for studying network effects specifically because it embodies a setting that sociolinguists understand well, in the sense of knowing that linguistic change occurred because people, especially children, from diverse dialect backgrounds came together (Trudgill 1986; Britain 1997; Thomas 1997; Kerswill & Williams 2000; Al-Wer 2007; Al-Rojaie 2013). In this context, we now consider Raleigh's school-based network structure.

5 Bipartite networks and the Raleigh school attendance network

Whereas many Generation 2 speakers attended school with the children of migrants from the Northern U.S., the majority of speakers in the sample from Raleigh's historically working class East Raleigh neighborhoods report not knowing any Northerners while growing up; those kids, they say, went to school in North Raleigh. In view of the preponderance of new settlement by affluent northerners in the northwestern areas of Raleigh (Rohe 2011), we hypothesized that the reversal of the SVS was more advanced among people who grew up in the North Raleigh areas because they had more exposure to non-Southern varieties and less exposure to the local Southern variety.

In most cases, it is impossible or unreasonable to directly observe interactions between people in a community, especially in a large urban community such as Raleigh. An alternative approach is to represent a relational social system by describing how people are indirectly tied through their collocation in physical settings, events, or organizations. This approach often makes sense if social relationships are mediated through institutions or events. For instance, consider an example where a researcher is interested in the social relationships among the CEOs of large corporations in the United States. Given this

population's size and status, it is probably infeasible to directly survey all of the CEOs of the Fortune 500 to ask about their relationships with one another. However, CEOs commonly belong to multiple corporate boards, non-profit boards, and civic clubs, and these affiliations provide information about their social relationships (Galaskiewicz & Wasserman 1981). Researchers can construct a network in which CEOs are one type of node, boards and clubs are another type of node, and a tie between a CEO node and a board or club node indicates affiliation. When two CEOs belong to the same social club or board of directors it can be said that they are more socially proximate than a pair who do not belong to the same organization. Consequently, the relational structure among CEOs can be inferred from their co-membership in organizational and institutional settings.

Networks in which the relations are mediated through shared settings, such as corporate boards or social clubs, are often referred to as “bipartite” or “two-mode” networks because the nodes fall into two distinct classes. Bipartite networks are distinct from conventional one-mode networks because their structure allows ties between nodes in separate partitions or classes (a person belonging to an organization) but not between nodes in the same partition because the researcher did not observe direct relations between individuals. A structure of relations among people can be inferred by observing the patterns of overlapping co-memberships (which people belong to the same organizations; which organizations have the same members).

Bipartite networks are particularly useful in contexts where individuals' colocation in a common event or organization provides meaningful information about their social context, their propensity to interact, or their differential exposure to common or distinct social settings or norms. Even when it cannot be assumed that two individuals interact when they are co-located in an organization, it is nevertheless generally safe to assume that the pair know many of the same individuals and may share important assumptions and reference points.

As described in Dodsworth & Benton (2017) and in Dodsworth & Benton (forthcoming), we model Raleigh's network structure at each of the three generations using bipartite networks in which one class of nodes represents speakers and one class of nodes represents elementary, middle, and high schools. A tie between a speaker and a school indicates that the speaker attended the school for at least a year. We are interested in a school-mediated network structure because we assume that the speakers mainly acquired their dialects – including their vowel systems – during childhood and adolescence, and that changes to their vowel systems after adolescence have been modest relative to the difference between speakers born a generation apart (Chambers 1992; Sankoff & Blondeau 2007; Siegel 2010; Rickford & Price 2013; Johnson & Nycz 2015). We do not assume that two speakers who attended the same school at the same time directly interacted with one another or had the same groups of friends, but we take for granted that they interacted with many of the same people and encountered many of the same linguistic and cultural norms.

Figure 4 depicts the bipartite network corresponding to Generation 1. The larger, unfilled circles represent elementary, middle, and high schools, and the smaller blue circles represent speakers. A tie between a speaker and a school means that the speaker attended the school for at least a year. Raleigh was relatively small at this point and schools were racially segregated, so most White students attended Broughton High School near the center of town after finishing at one of a small set of elementary schools. A few speakers in the sample went to Enloe High School in the working class Eastern part of Raleigh, or to Garner High School to the South.

In the Generation 2 network (Figure 5), Broughton High School remains a high-degree node, but now it is less central to the network defined by the sample because multiple North Raleigh high schools, middle schools, and elementary schools have emerged. Enloe

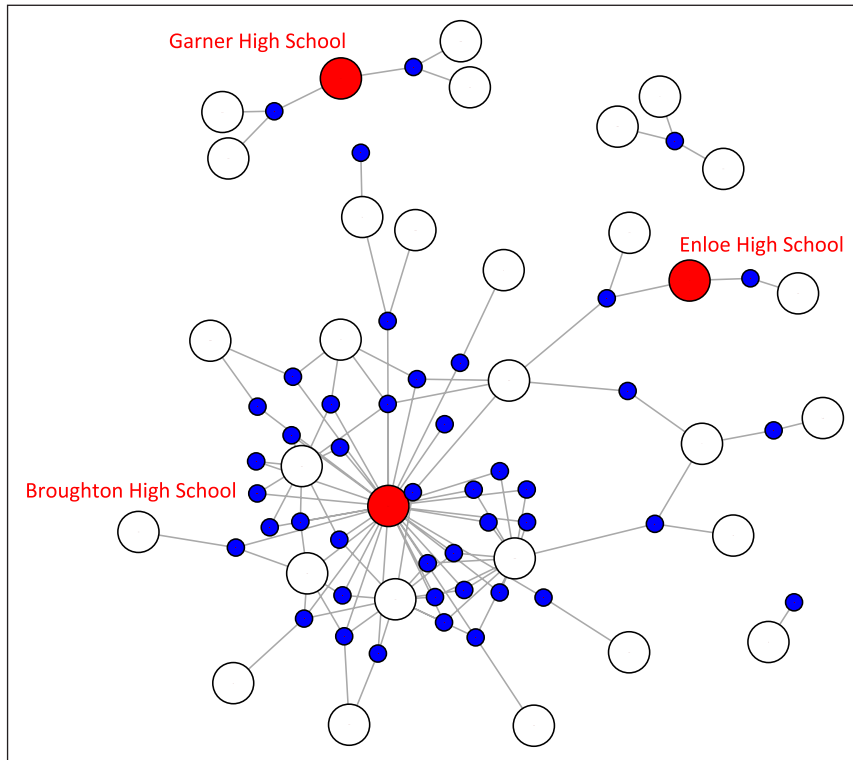


Figure 4: The bipartite network of schools and speakers, Generation 1. Small blue circles represent individual speakers and large white and red circles represent schools. A tie represents a speaker’s attendance at a school for at least a year.

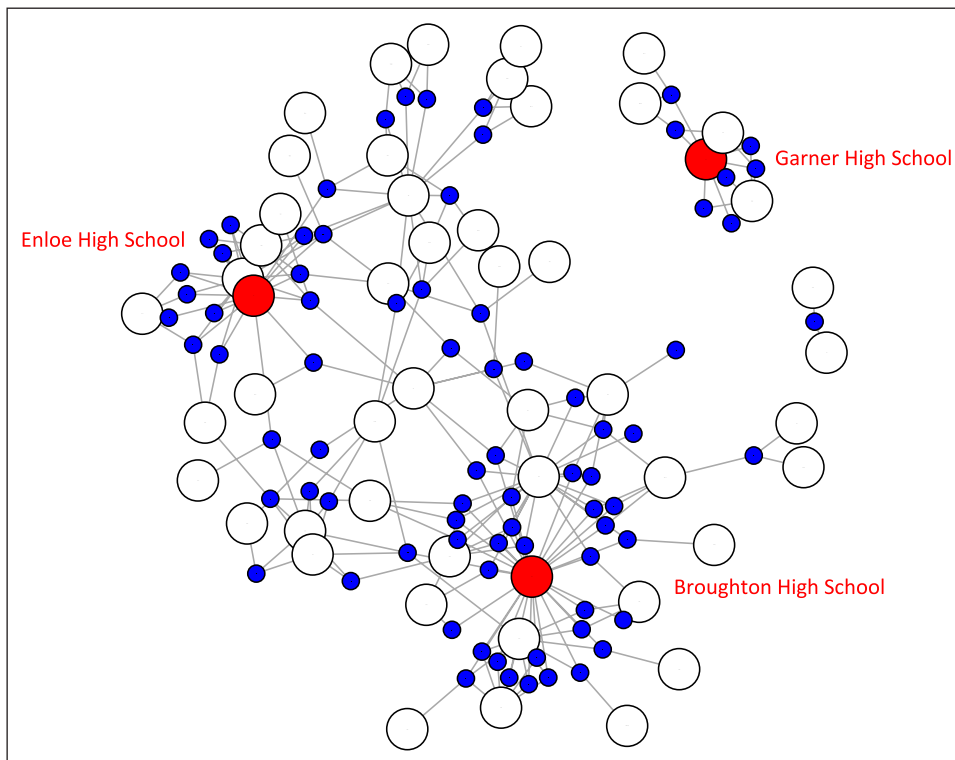


Figure 5: The bipartite network of schools and speakers, Generation 2. Small blue circles represent individual speakers and large white and red circles represent schools. A tie represents a speaker’s attendance at a school for at least a year.

High School was still the local school for East Raleigh, and it was the first high school to desegregate. Garner schools remain disconnected from the rest of the school network, indicating that no speakers in the sample who attended Garner schools during elementary or middle school then transferred to other Raleigh schools later. Broughton High School is close in the network to one of the first North Raleigh high schools, Sanderson in particular; this is partly because some students first attended Broughton and then Sanderson, and partly because some students lived in central Raleigh during elementary and middle school before moving to North Raleigh in time for high school. Enloe High School, on Raleigh's East side, is close in the network to a different North Raleigh high school, Millbrook, for the same reasons. Some Raleigh speakers report that their families moved from the Enloe district to the Millbrook district in order to avoid Enloe as it became integrated, or to avoid East Raleigh's racially mixed neighborhoods.

During Generation 3 (Figure 6), Enloe High School began a magnet program in addition to its regular academic program, meaning that students from across the public school system could apply to Enloe in order to take advantage of special academic opportunities. For that reason, a wider range of elementary and middle schools are close in the network to Enloe in Generation 3 than in Generation 2, and some students transferred from other high schools to Enloe. The magnet program began in 1980, and the oldest Generation 3 speakers were born in 1967; therefore, all of the Generation 3 speakers who attended Enloe were there after its magnet program had begun. During Generation 3, more schools at all levels were built as North Raleigh and other areas continued to expand.

While we will refer readers to Dodsworth & Benton (forthcoming) for detailed quantitative comparisons of the three bipartite network structures, we will mention here that the average geodesic distance – the average shortest path length – between speakers increases from one generation to the next. That is, as Raleigh grew and became more geographically

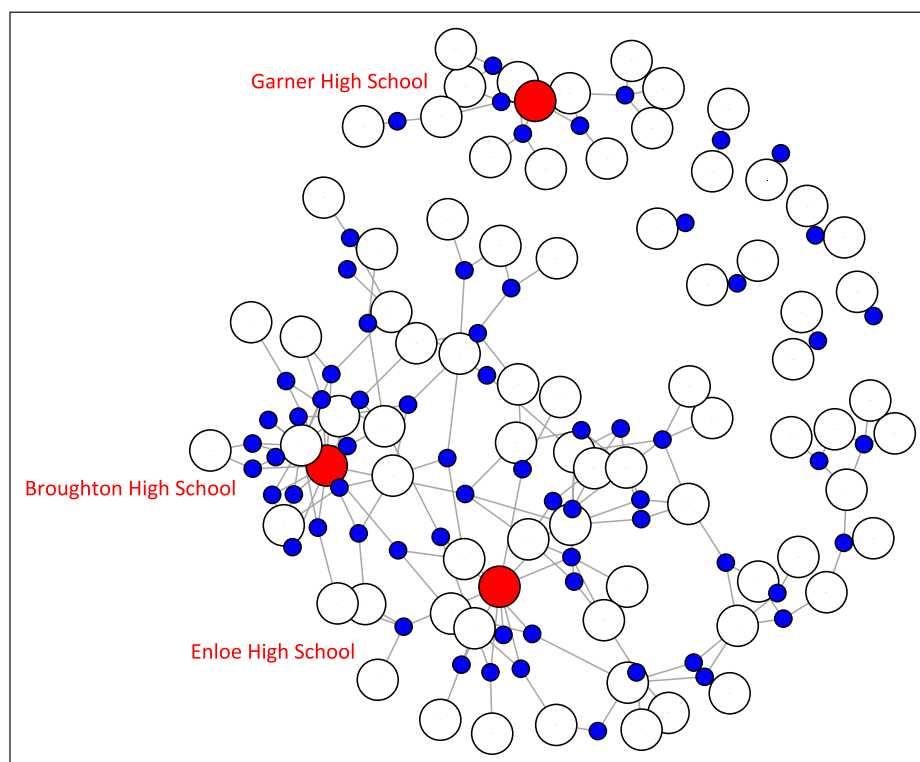


Figure 6: The bipartite network of schools and speakers, Generation 3. Small blue circles represent individual speakers and large white and red circles represent schools. A tie represents a speaker's attendance at a school for at least a year.

dispersed, social distance between pairs of speakers increased on average. In the early cohort, for any two individuals who did not attend school together, the odds were good they had a common connection through a third individual. However, for the most recent cohort, this distance has increased and any two individuals are more likely to have to rely on longer chains of two or three individuals to find a connection. A common theme in earlier sociolinguistic network analysis is that dense local networks can enforce linguistic norms, promoting the maintenance of a local vernacular even in the context of pressure from an external dialect. If this is true, then we can expect that the Raleigh community's capacity for maintaining the SVS was weakened as the social proximity among speakers decreased.²

In a previous analysis (Dodsworth & Benton 2017), we found that the interactions between network integration and age difference indicate that a central, dense area of the overall network corresponds to linguistic homogeneity during dialect contact and dialect shift: when two speakers are both embedded in deep areas of the network (which is to say they are near the dense urban core, rather than in peripheral North Raleigh or elsewhere), they are more alike linguistically. A large difference in age between two speakers generally corresponds to linguistic difference between them, but the linguistic difference is less when the two speakers are both well-embedded in the network. This is a useful conclusion insofar as we've reached a standard sociolinguistic finding using a replicable method that uses a contemporary network analysis routine. However, in keeping with the goal of developing new hypotheses about the relationship between social network structure and linguistic variation, we now strive for a more general measure of network similarity that could be useful in a broader range of sociolinguistic settings and for a broader range of questions. We specifically want a measure that allows us to consider the interactions among network position, occupation, sex, and age. For this we've used a measure of *structural equivalence* as the network variable.

6 Structural equivalence

Two nodes are structurally equivalent when they have ties to the same alters (Marsden & Friedkin 1993). This network metric follows from the intuition that two individuals face nearly identical social worlds (norms, interaction opportunities, and preferences) if they're connected to identical network contacts. We measure the degree of structural equivalence between every two speakers (nodes) as their Jaccard similarity, which is defined as the number of elements the nodes have in common divided by the sum of the elements in common and the elements not in common (the intersection divided by the union). In this case, the Jaccard similarity is calculated as in Figure 7: the number of schools they both attended divided by the sum of the number of schools they both attended and that only

$$1 - \frac{\text{number of schools both speakers attended}}{\text{number of schools both speakers attended} + \text{number } i \text{ but not } j \text{ attended} + \text{number } j \text{ but not } i \text{ attended}}$$

Figure 7: Jaccard distance formula.

² It is possible that dense, multiplex areas of the community network retained the SVS even as the community network as a whole became less dense. We find evidence for this in Dodsworth & Benton (2017).

one of them attended. This ratio is then subtracted from 1. The result is a number between 0 and 1, with numbers closer to 1 indicating greater distance. The higher the distance between two nodes, the further away the two nodes are from being structurally equivalent, and the lesser their network similarity.

We calculate Jaccard distance for every dyad, and this gives a matrix of distances for each generation.³ We then use the distance matrix to ask a “relational” question about linguistic variation: for each linguistic variable, does the mean linguistic difference between pairs of speakers correlate with their Jaccard distance?⁴ If a speaker’s network position influences his/her exposure to and orientation toward the SVS, then we expect that pairs of speakers with lower distance will also have lower mean linguistic difference, net of other social factors. This hypothesis is both very simple and a departure from the common sociolinguistic hypothesis that deeper integration into a dense, multiplex network will correspond to greater use of the vernacular. We further hypothesize that network position interacts with age, sex, and especially occupation. Network similarity was expected to moderate the effect of age difference on linguistic difference; in other words, speakers far apart in age were expected to be linguistically different in the context of linguistic change in progress, but when they have similar network positions, the age effect was expected to be weaker. The strong collinearity between network and occupation has largely prevented detailed statistical treatments of their interaction in previous sociolinguistic network studies. In contrast, several previous sociolinguistic studies illustrate the complex relationship between network structure and sex, often having to do with differences between men and women in the types of jobs or religious roles they typically occupy within particular communities (Gal 1979; Edwards 1986; Milroy 1987; Rosen & Skriver 2015; Sharma 2017). In fact, Labov (2001) finds significant network effects *only* for women.

We test these hypotheses within each generation using Quadratic Assignment Procedure (QAP) models. QAP is analogous to standard linear regression, but all the variables are matrices rather than vectors, and standard errors are calculated with a bootstrapping technique that permutes the rows and columns in the dependent variable matrix in order to deal with the non-independence among values in the same row or column. For that reason, the p-values are more conservative than in standard OLS regression (Krackhardt 1988; Mizuchi 1993). In this case, the dependent variable for each vowel is the matrix of differences in means for every pair of speakers. The fixed effects are a set of matrices corresponding to social differences (Table 2). One is a matrix of differences in age for each pair of speakers; another is the Jaccard distance matrix; and the others have to do with year of birth, occupation, and sex.

While our hypotheses focused on aggregate patterns rather than on the social indexicality of the SVS, the QAP results – especially as they concern the interactions among network, occupation, and sex – suggest a complex sociolinguistic setting. The QAP results are described in full in Dodsworth & Benton (forthcoming). In the present analysis, we use examples from the results to explore the ways in which the social indexicality of the SVS is

³ The choice of Jaccard distance over other common metrics of distance between binary vectors is inconsequential in our data. For example, the correlations between our Jaccard distance matrices and the same matrices built from the Dice and Kulczynski metrics ranges from .95 to .98 across our three generations of speakers.

⁴ A reviewer suggested that the same hypothesis could be approached via the simpler method of assigning each dyad a score in the range 0–3, reflecting the number of schools that *i* and *j* both attended and ignoring schools that they don’t have in common. One advantage in this approach would be to remove some high-leverage values in the distribution of distances, especially in Generation 3 when there are many more schools in Raleigh and thus greater skew toward high distances between speakers. However, this approach would miss the distinction between a speaker who attended schools *a*, *b*, and *c* and a speaker who attended schools *a*, *b*, *c*, and *d*. That is, it would not adequately represent dyads in which one speaker attended more than one elementary, middle, and/or high school. This is the case for 37 (20%) of the speakers in the present sample (7 attended 5 schools; 30 attended 4 schools).

Table 2: Independent variables in QAP regression models.

Matrix name	Description
j's year of birth	each column in the matrix has one value
age difference	the difference in year of birth between i and j
j's occupation	0 = blue collar, 1 = unskilled white collar, 2 = white collar; each column in the matrix has one value
same/different occupation	whether i and j have different occupations in the 3-way scheme (blue, unskilled white, white): 0 = different, 1 = same
j's sex	0 = female, 1 = male
same/different sex	whether i and j are coded differently for binary sex : 0 = different, 1 = same
Jaccard distance	the Jaccard distance between i and j as described above

revealed in our aggregate, school-based network strategy. The regression coefficients and corresponding significance levels from the QAP models for /e/ and /ai/ in the Generation 2 sample and /e/ and /ɛ/ in the Generation 3 sample are reported in Tables 3–6, respectively.⁵ The Generation 2 models omit 4 dyads that were outliers insofar as they had two characteristics: *j* was blue collar and the Jaccard distance between *i* and *j* was below .5. The Generation 3 models omit 10 dyads that were outliers in having Jaccard distance below .25. The omission of these outliers did not result in a different set of significant effects, relative to those reported in Dodsworth & Benton (forthcoming), though it did alter the coefficients slightly.⁶

Unlike in our previous reports of these and other network patterns, we focus here on individual speakers and their locations in the aggregate trends. Our goal is to bring together the different types of explanatory potential available in our relatively abstract, macro-level analysis, on one hand, and the more social-meaning-focused, small-group approaches that have characterized some previous sociolinguistic network studies.

7 QAP results: examples that reveal exceptional speakers

7.1 Generation 2 /e/

In Generation 2, the first generation to grow up amid heavy migration from the North, one of the vowels showing a significant positive effect of Jaccard distance is /e/, meaning that the greater the Jaccard distance between two speakers, the more different their /e/ nucleus on average (Table 3, model 14). This is evidence that a speaker's position in the community network, in this case during childhood and adolescence, influences his/her adult use of Southern linguistic variants. Most of the other social factors show no significant effects in any model in Table 3, which suggests that the effect of Jaccard distance is not an epiphenomenal consequence of the effect of sex, occupation, or age. However, Jaccard distance is significant only when the model also contains the interaction between occupation (which also shows a significant positive main effect, indicating that blue collar speakers are more similar to their alters overall) and Jaccard distance, and the interaction is significant. These effects are illustrated in Figure 8. Each circle represents one dyad (*not* one speaker), and the higher the circle along the y-axis, the greater the mean difference in /e/ between the two speakers in the dyad. The three facets distinguish only *j*'s occupation; so the speakers comprising the dyads in the rightmost facet do not necessarily *both*

⁵ These models' R^2 values are very low, as expected, because we are dealing with dyadic data. Our goal is to determine whether the factors in the model significantly predict linguistic similarity, rather than to account for all of the variance in the linguistic data (cf. Mizruchi 1993; footnote 11).

⁶ We have not corrected for multiple comparisons, as Bonferroni-type correction isn't typically applied to QAP models. We claiming very few significant effects, and they are mostly limited to age and network.

Table 3: Coefficients from QAP models for /e/ in Generation 2. Coefficients are multiplied by 100.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
j's birthyear	0													
age difference		.6	.7	.6	.6	.6	.7	.3	.6	.8	-1.1	.6	.7	.7
same occupation			.8						.9				16.7	
j's occupation				.9						1.2				17.5*
same sex					.5									
j's sex						.8					-5.8	-.2		
Jaccard distance							-5.9	-8.0				-7.4	3.9	40.8*
age difference*same occupation									0					
age difference*j's occupation										-1				
age difference*j's sex											1.2*			
age difference*Jaccard distance								.4						
same occupation*Jaccard distance													-17.6	
j's occupation*Jaccard distance														-18.1*
j's sex*Jaccard distance												1.1		
Adjusted R ²	0	0	0	0	0	0	.01	.01	0	0	.01	0	.01	.01

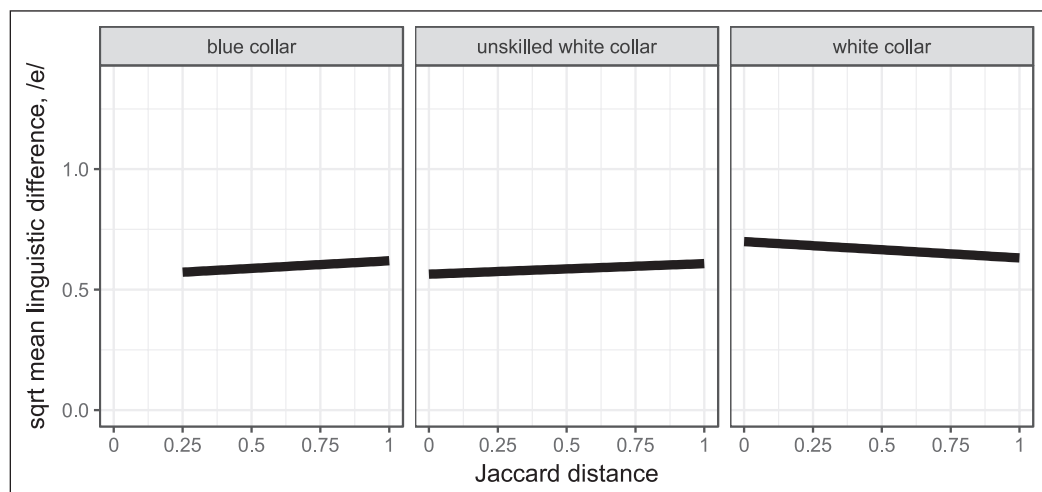


Figure 8: Generation 2 /e/.

have white collar jobs, though many of them do. The basic pattern, which appears subtle but is significant in the QAP model, is that the correlation between Jaccard distance and linguistic difference is less positive for white collar speakers than for blue collar and unskilled white collar speakers. The main difference between the occupational groups is that when *j* is white collar, dyads with low Jaccard distance can nevertheless have high linguistic difference.

We can gain insight into the semiotics associated with this pattern, and potentially with the role of the individual in a contact setting such as Raleigh, by looking at individual dyads in the context of the aggregate pattern. All observations above 1.22 on the y-axis in

Figure 8 correspond to dyads in which one person is one of six white collar speakers with unusually non-Southern /e/ and the other speaker (who may be white collar or not) has unusually Southern /e/, for that generation. For example, one dyad in which both speakers are white collar and grew up in central Raleigh has the minimum Jaccard distance of 0 (meaning the speakers attended exactly the same schools) but the relatively high linguistic difference of 1.3 for /e/. The more Southern-shifted speaker in this dyad is a male born in 1950 who works in corporate sales for the local and (Southern) regional construction industry, where the SVS carries positive cultural capital. The less Southern-shifted speaker is a female born in 1952 who works at various hospitals, and she talks every day with patients and other hospital workers from both Southern and non-Southern dialect regions. Even though the two speakers grew up in the same neighborhood and attended exactly the same schools at about the same time, and both speakers are white collar, under a third-wave sociolinguistic perspective (Eckert 2000; Eckert this volume) or any theoretical approach in which linguistic variants reflect, and/or partly constitute, socioeconomic structure, we would expect these two speakers to have different symbolic linguistic needs as adults. The speakers' contrasting roles in the local economy and perhaps also their different linguistic exposure as adults are likely to lead them to use SVS variants at different rates. It is important to note that both speakers are constrained by their age and network position; neither has a vowel system resembling someone who grew up in North Raleigh 10 years later. But the contrast between them nevertheless suggests that in this context, the individual has some agency in deciding to what degree to take part in Raleigh's ongoing shift away from the SVS. While our account of this or any single dyad is necessarily speculative, the more general point is that although network similarity promotes linguistic similarity overall (cf. the positive main effect of Jaccard distance in Table 3, model 14), we can also find evidence that the symbolic value of linguistic variants can moderate the homogenizing effect of network similarity. Eckert's perspective (this volume), that linguistic variants emerge and have value by virtue of the semiotic landscape, would in fact seem to require the interplay between aggregate network mechanisms and local semiotics, as the latter cannot arise and stabilize without the former.

Toward the later end of Generation 2, there are no white collar low outliers (cf. Figure 2). We can see this as evidence either that college-bound adolescents increasingly have little exposure to Southern linguistic variants, or that white collar speakers in any job – unlike blue collar speakers – are penalized for sounding too Southern. In either case, in Raleigh's changing semiotic landscape, Southern linguistic variants are increasingly at odds with a young, professional, educated persona.

The positive main effect of *j*'s occupation (Table 3, model 14) indicates that blue collar speakers show less linguistic difference overall from their alters in Generation 2. This is because they are more linguistically Southern than white collar speakers; returning to Figure 2 (the apparent-time graph for /e/), there is little linguistic difference between a low outlier born in 1965 and an average speaker born in 1951. For that reason, it is not surprising that linguistic difference is lower on average when *j* is blue collar than when *j* is white collar, even after omitting the blue collar outlier dyads with both very low Jaccard distance and very low linguistic difference, as noted. What we need to explore, from a social indexical perspective, is why linguistic difference is higher when Jaccard distance is high, when *j* is blue collar. We again turn to individual speakers for insight. Among the blue collar speakers in Generation 2, three have notably the most Southern mean /e/ (below $-.4$, in contrast with the other blue collar speakers, whose means range from $-.09$ to $.44$). As a consequence, these three speakers are members of most dyads with the highest linguistic difference in which *j* is blue collar. All three are male, and all have school histories that result in high Jaccard distance vis-à-vis most other speakers. One

of them attended school in the Garner/Fuquay region to the south of Raleigh, and one attended private religious schools rather than public schools. For both speakers, it appears that the combination of a relatively peripheral school network position and a blue collar profession results in the unusually strong retention of Southern /e/. Plenty of white collar speakers attended private schools but have less Southern /e/, and the other blue collar speakers, who are more central in the school network, also have less Southern /e/. For blue collar speakers, then, we surmise that the symbolic value of SVS variants at work, together with unusually consistent childhood exposure to the SVS as the result of network position, result in the retention of SVS variants and thus linguistic distance from others.

The third blue collar speaker with exceptionally Southern /e/ is a different story. His Jaccard distance from others is high not because he is peripheral in the school network but because he attended an unusually diverse set of schools including two elementary schools, two middle schools, and two high schools in central and North Raleigh. His varied school background undoubtedly gave him exposure to both Southern and non-Southern dialects, in contrast with the experiences of the two other exceptional blue collar speakers. He also went to college, unlike the others. However, he owns and runs a blue collar business that he took over from his father. The business is not only in a “dirty work” industry (Hughes 1951/1971; Ashforth & Kreiner 1999) but is also physically located to the south of town in a traditionally working class area. We again speculate, as we did in the case of the white collar dyad above, that his economic role motivates his retention of SVS variants. In particular, talking Southern is probably good for his relations with his employees, many of whom may be from surrounding small towns and probably none of whom attended college. This speaker’s school attendance during childhood and adolescence necessarily gave him strong, long-term exposure to non-Southern dialects, giving him the opportunity to adopt a less Southern vowel system, as most of his peers did. He did not take this opportunity. While we have no direct evidence that his economic or family ties promote his use of SVS variants, we do know that his large linguistic difference vis-à-vis others is a function of something other than school network peripherality (because he is not network-peripheral). The overall high linguistic difference between blue collar speakers and their alters thus appears to result from the combination of network distance and the semiotic value of Southern variants.

7.2 Generation 2 /ai/

Similar to /e/ nucleus raising and fronting, /ai/ glide strengthening in Generation 2 shows positive main effects of both *j*’s occupation and Jaccard distance, as well as their interaction (Table 4, model 14). Whereas the slope associated with Jaccard distance was more positive for blue and unskilled white collar speakers than for white collar speakers in the case of /e/, this is true *only for women* in the case of /ai/ (Figure 9). This is underscored by the fact that QAP models containing the sex and age difference variables, but not *j*’s occupation, find positive main effects for sex (Table 4, models 11 and 12).

One side of the pattern is the absence of a positive effect of Jaccard distance among white collar women. This results in part from the fact that there are several white collar women born around 1955 in central Raleigh who are high outliers for /ai/, producing an unusually strong (non-Southern) glide relative to other speakers born near the beginning of Generation 2. Their Jaccard distance from other central Raleigh speakers is low, but their linguistic difference from many of those others is relatively high, resulting in a flat slope for white collar women in Figure 9. Given that monophthongal /ai/ is arguably the most iconic Southern vocalic feature, and possibly the most geographically widespread SVS feature (Labov et al. 2006), we are not surprised to find some middle-aged white collar women resisting it.

Table 4: Coefficients from QAP models for /ai/ in Generation 2. Coefficients are multiplied by 100.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	.7													
j's birthyear		-2	-2	-2	-2	-1	-2	-6	0	1.3	-1.7	-1	-2	-3
age difference			-1.5						.6				6.1	
same occupation				5.1						8.3				27.1**
j's occupation					3.3									
same sex						-9.8					-16.0*	-24.9*		
j's sex							14.4	12.6				-6.9	18.7	76.1**
Jaccard distance									-4					
age difference*same occupation										-6				
age difference*j's occupation											1.1			
age difference*j's sex								.4						
age difference*Jaccard distance													-8.1	
same occupation*Jaccard distance														-23.7*
j's occupation*Jaccard distance												16.5		
j's sex*Jaccard distance	.01	0	0	.01	0	.01	0	0	0	.01	.01	.02	0	.01

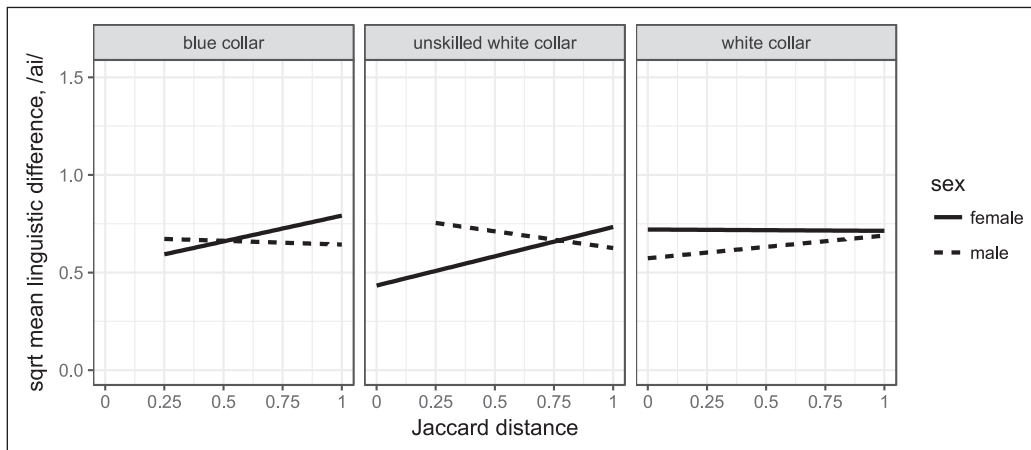


Figure 9: Generation 2 /ai/.

Among blue collar women, a single speaker is largely responsible for the positive slope in Figure 9, and the contrast between that speaker and another speaker brings us again to the question of semiotics and the individual. There are only two blue collar female speakers in Generation 2. Of these two speakers, one has the most Southern-shifted pre-voiced /ai/ of anyone in Generation 2, and the other has unexceptional /ai/ in the context of the Generation 2 sample. The more Southern-shifted speaker was born in 1958, worked for over 30 years in a warehouse, and grew up mostly in northeast Raleigh, distant from both central Raleigh and the affluent northwestern section that saw heavy migration during this period. In most of this speaker's dyads, therefore, linguistic difference and Jaccard

distance are both high. In this sense, the speaker follows the pattern described for /e/, in which blue collar speakers who are network-peripheral are the most likely to retain SVS variants. The other blue collar female in this generation was born in 1956 and grew up in East Raleigh, which means her Jaccard distance from the more affluent areas of Raleigh is high. But because her mean /ai/ is unexceptional for her age, she accounts for the instances of average linguistic difference but high Jaccard distance when *j* is blue collar and female. If both of these speakers are blue collar and network-peripheral, why does only one of them retain an unusually Southern-shifted /ai/?

Our explanation, speculative but consistent with third-wave sociolinguistic perspectives, is that the more Southern-shifted speaker encountered significantly greater need for the symbolic capital of the SVS in her long-term warehouse job. She did heavy manual work in the company of men, and symbolically displaying either toughness or a working-class orientation (or, more likely, both, as they are not independent) was probably important. Like the Burnout girls at Belten High (Eckert 2000), she had more to gain from using the local vernacular, and so her retention of SVS forms is likely a strategic symbolic choice. This is not to say that the less Southern-shifted speaker was discouraged from talking Southern at work. Over the course of her life, she worked in restaurants, as a bartender, and in clothing retail prior to taking up her current (manual) trade. While some of these customer-facing jobs might have given her exposure to non-Southern varieties, all of them were at local establishments in which talking Southern would have been perceived as normal and friendly during the 1970s and 1980s. But these jobs probably would not have required her to symbolically display toughness or a working-class identity in order to be respected by male co-workers. On the contrary, having customer-facing jobs probably made linguistic flexibility a greater priority.

The Generation 2 /ai/ pattern can be seen as further evidence that network position and the semiotic landscape simultaneously influence the distribution of Southern variants among this age group. As in the case of /e/, individuals' network positions matter, but they appear to have some agency in deploying semiotic linguistic resources in the context of rapid contact-induced linguistic change.

7.3 Generation 3 /e/

By the time the Generation 3 speakers (born between 1967 to 1996) were in school, the contact-induced retreat from the SVS had been underway for at least a decade, and some exposure to non-Southern dialects was inevitable, though not equally across neighborhoods and schools. During this second post-contact generation, dialect change continued and then began to level off. The regression results reveal an evolving intersection of the effects of age, occupation, network, and sex. For /e/, one of the vowels showing significant network and occupation effects in Generation 2, occupation and network remain significant as main effects in Generation 3, and their interaction remains significant (Table 5, model 14) but the model accounting for the most variance contains the significant interaction between age difference and *j*'s sex (Table 5, model 11). What is happening (Figure 10) is that blue collar women are much more different from their alters, as a group, than blue collar men, while white collar women and white collar men are not different from each other. The apparent significant effect of Jaccard distance in model 14 is in fact due to the difference between blue collar men and women.

Blue collar females' very high linguistic difference from their alters overall largely reflects the fact that the two older blue collar females – neither of whom has a peripheral school network position – have by far the most Southern-shifted /e/ in Generation 3. In contrast, the youngest blue collar female is only slightly more Southern than the white collar speakers of the same age. So between Generations 2 and 3, the significant

Table 5: Coefficients from QAP models for /e/ in Generation 3. Coefficients are multiplied by 100.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
j's birthyear	-5													
age difference		1.1***	1.1***	1.1***	1.1***	1.1***	1.1***	-2.5	.9**	-1.1*	3.1***	1.1***	1.1***	1.1***
same occupation			-12.9						-16.2*				-49.1	
j's occupation				-2.3						-10.6*				-61.7***
same sex					-8.5***									
j's sex						-11.7					2.5	-5.6		
Jaccard distance							-9.5	-41.8*				-1.1	-40.4	-185.2***
age difference*same occupation									.3					
age difference*j's occupation										.8***				
age difference*j's sex											-1.3***			
age difference*jaccard distance								3.8*						
same occupation*jaccard distance													37.6	
j's occupation*jaccard distance												-6.5		61.4***
j's sex*jaccard distance														
Adjusted R ²	.01	.04	.06	.04	.05	.06	.04	.05	.06	.05	.08	.06	.06	.05

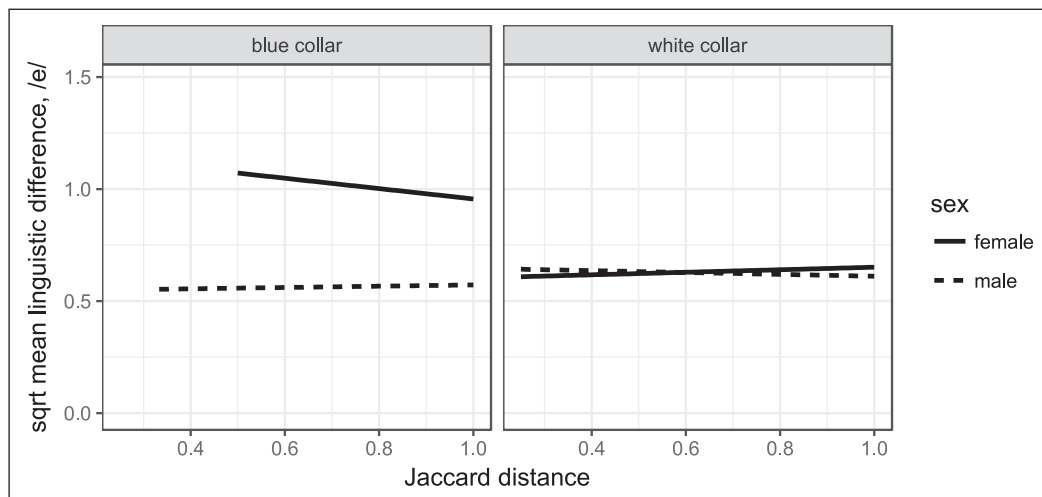


Figure 10: Generation 3 /e/.

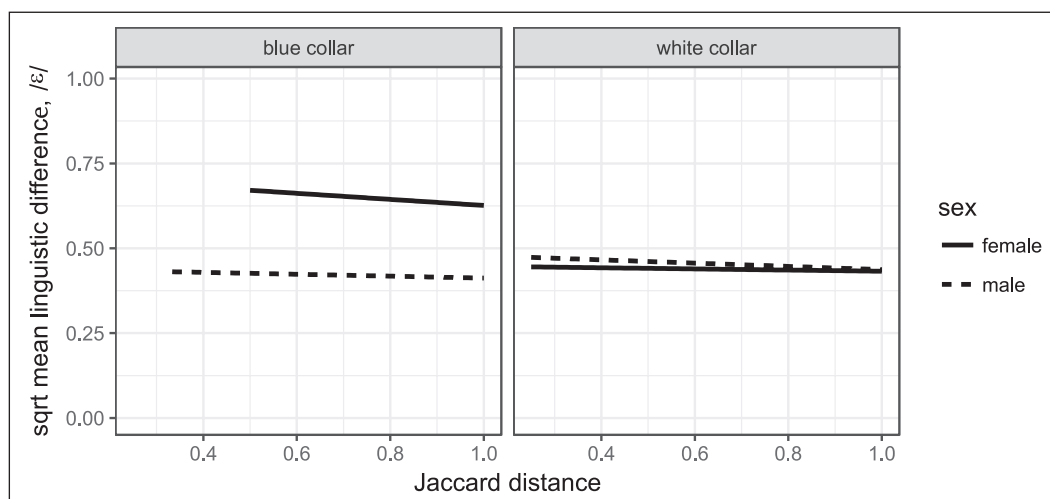


Figure 11: Generation 3 /ε/.

network effects in the case of /e/ weakened, but the interacting effects of occupation and sex – which were significant in Generation 2 – remained and then weakened as the SVS further disappeared from Raleigh. We take this as evidence that although by Generation 3, most Raleigh natives had consistent childhood *exposure* to non-Southern dialects, they continued to have varied, socially driven motivation for embracing or rejecting SVS variants. In particular, blue collar women and white collar women had contrasting symbolic needs, regardless of network position.

The case of /ε/ is remarkably similar to that of /e/ in Generation 3. Jaccard distance and *j*'s occupation again show significant main effects as well as a significant interaction (Table 6, model 14). In Figure 11, which is nearly identical to Figure 10, the blue collar females as a group are more different from their alters than all other groups. As in the case of /e/, the two oldest blue collar females have the most Southern-shifted /ε/ of anyone in Generation 3, whereas the youngest blue collar female shows an unexceptional mean value for /ε/, higher than some and lower than other speakers her age. Therefore, the two mid vowels, which underwent more dramatic Southern shifting in Raleigh to start with, both show evidence of interacting effects of network position, occupation, sex, and age in Generation 3. The stark difference between the older blue collar females and males is absent among white collar speakers of the same age.

Table 6: Coefficients from QAP models for /ε/ in Generation 2. Coefficients are multiplied by 100.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14
j's birthyear	-.3*													
age difference		.3**	.3**	.3**	.3**	.3**	.3**	-.7	.4**	-.2	.9**	.3**	.3**	.3**
same occupation			-5.3						-4.3				-18.7	
j's occupation				-1.4						-3.3				-24.8**
same sex					-3									
j's sex						-2.2					1.6	-1.7		
Jaccard distance							-5.4	-14.0				-4.9	-16.8	-74.7**
age difference*same occupation									-1					
age difference*j's occupation										.2*				
age difference*j's sex											-3*			
age difference*jaccard distance								1.0						
same occupation*jaccard distance													13.9	
j's occupation*jaccard distance														24.2**
j's sex*jaccard distance												-5		
Adjusted R ²	.02	.02	.03	.02	.02	.02	.02	.02	.03	.02	.02	.02	.03	.02

8 Conclusion: Viewing individual agency through network position

We have found significant effects of Jaccard distance and its interaction with occupation and sex: when two speakers are far apart in the community network, they're more likely to be linguistically different, and in some cases, blue collar females are most likely to retain the SVS. We see this as an important set of results for two reasons. The first is that although Jaccard distance is a very rough measure of network distance, it can nevertheless reveal the influence of network position (in this case, as indexed by school attendance), in shaping the distribution of linguistic variants. The second reason is that the bipartite approach to social network structure, together with QAP analysis, has allowed us to look at the intersecting and changing effects of network, occupation, and sex, something that has proven elusive in previous sociolinguistic research (Milroy & Milroy 1992). Our school-based bipartite network approach has the further advantage of representing each speaker's position in the overall community network, in contrast with previous sociolinguistic network studies, and it also avoids self-reported data and is replicable across communities with public school (or other similar) systems.

More important for the present analysis, however, is that assessing the interacting effects of age, sex, occupation, and network position reveals some apparent symbolic use of the SVS that cannot be reduced to differences in exposure. In Generation 2, two central Raleigh speakers of the same age with identical network position (which points to very similar linguistic exposure during childhood) were found to be linguistically very different, and we proposed that the difference arises from individual symbolic choice. Similarly, the linguistic difference between two blue collar females, both network peripheral, in Generation 2 appears to arise from their differing symbolic needs over the course of their occupational trajectories. The most surprising individual is perhaps the blue collar, Generation 2 male speaker with exceptionally Southern /e/. His network position means that he had strong exposure to non-Southern variants during childhood and adolescence, and he graduated from a four-year college that undoubtedly gave him more contact with non-Southern speakers. Nevertheless, in the context of the business he runs in a "dirty work" industry, he uses Southern variants apparently in order to construct a blue collar identity, and/or to accommodate to his employees. While we can't know his or the other exceptional speakers' motivations, the important point in the context of the present analysis is that their exceptional vowel systems (whether very Southern or very non-Southern, relative to their peers) are not uniquely the products of exceptional network positions. Instead, we speculate that their occupational histories have made particular linguistic variants symbolically valuable to them.

We have similarly proposed that Raleigh blue collar women's retention of the SVS is a strategic symbolic choice, similar to Burnout girls using the most extreme Northern Cities vowels (Eckert 2000). In that sense, our results are consistent with those of previous sociolinguistic network studies, which typically find that certain aspects of interactional life and personal identity promote the symbolic use of linguistic variants associated with a place-based or ethnic identity (Ash & Myhill 1986; Milroy 1987; Lippi-Green 1989; Sharma 2017).

Nevertheless, in our results, *position in the overall network* is also significant during the first generation of contact-induced change (Generation 2), both as a main effect and in interaction with other factors depending on the vowel. This indicates that *exposure* to the SVS also drives the differential retention of, or resistance to, SVS variants. Agent-based models, as in Stevens et al. (this volume), similarly indicate that contact among speakers (personal identity aside) promotes the diffusion of linguistic variables. We can best discern the role of the individual in linguistic change when we can see individual speakers within the context of information about the community network structure, which allows

for inferences about their linguistic exposure, together with information about age, class, and other social factors. Our goal is not to replace the ethnographic methods that have been useful in previous sociolinguistic network analysis, but to introduce new aggregate network methods that can be used together with previous methods to construct a unified picture of the sociolinguistic landscape and the individual's place in it.

Funding Information

This work was supported by the National Science Foundation (BCS-1323153).

Competing Interests

The author has no competing interests to declare.

References

- Al-Rojaie, Yousef. 2013. Regional dialect leveling in Najdi Arabic: The case of the deaffrication of [k] in the Qaṣīmī dialect. *Language Variation and Change* 25. 43–63. DOI: <https://doi.org/10.1017/S0954394512000245>
- Al-Wer, Enam. 2007. The formation of the dialect of Amman: From chaos to order. In Catherine Miller, Enam Al-Wer, Dominique Caubet & Janet CE Watson (eds.), *Arabic in the city: Issues in dialect contact and language variation*, 55–76. Routledge.
- Ash, Sharon & John Myhill. 1986. Linguistic correlates of inter-ethnic contact. In David Sankoff (ed.), *Diversity and diachrony*, 33–44. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/cilt.53.06ash>
- Ashforth, Blake E. & Glen E. Kreiner. 1999. “How can you do it?": Dirty work and the challenge of constructing a positive identity. *The Academy of Management review* 24. 413–434. DOI: <https://doi.org/10.5465/amr.1999.2202129>
- Bailey, Guy. 1997. When did Southern American English begin? In Edward Schneider (ed.), *Englishes around the world 1: Studies in honor of Manfred Görlach*, 255–75. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/veaw.g18.24bai>
- Baker, Adam, Diana Archangeli & Jeff Mielke. 2011. Variability in American English s-retraction suggests a solution to the actuation problem. *Language Variation and Change* 23. 347–374. DOI: <https://doi.org/10.1017/S0954394511000135>
- Bortoni-Ricardo, Stella M. 1985. *The urbanization of rural dialect speakers: A sociolinguistic study in Brazil*. Cambridge: Cambridge University Press.
- Britain, David. 1997. Dialect contact and phonological reallocation: ‘Canadian raising’ in the English Fens. *Language in Society* 26. 15–46. DOI: <https://doi.org/10.1017/S0047404500019394>
- Chambers, J. K. 1992. Dialect acquisition. *Language* 68. 673–705. DOI: <https://doi.org/10.1353/lan.1992.0060>
- Cheshire, Jenny. 1982. *Variation in an English dialect: A sociolinguistic study*. Cambridge: Cambridge University Press.
- Cheshire, Jenny, Sue Fox, Paul Kerswill & Eivind Torgersen. 2008. Ethnicity, friendship network and social practices as the motor of dialect change: Linguistic innovation in London. *Sociolinguistica* 22. 1–23.
- Dodsworth, Robin. 2018. Community detection and the reversal of the Southern Vowel Shift in Raleigh. In Jeffrey Reaser, Eric Wilbanks, Karissa Wojcik & Walt Wolfram (eds.), *Language variety in the New South: Contemporary perspectives on change and variation*, 241–256. Chapel Hill, NC: University of North Carolina Press.
- Dodsworth, Robin & Mary Kohn. 2012. Urban rejection of the vernacular: The SVS undone. *Language Variation and Change* 24. 221–45. DOI: <https://doi.org/10.1017/S0954394512000105>

- Dodsworth, Robin & Richard Benton. Forthcoming. *Language variation and change in social networks: A bipartite approach*. Routledge.
- Dodsworth, Robin & Richard Benton. 2017. Social network cohesion and the retreat from Southern vowels in Raleigh. *Language in Society* 46. 371–405. DOI: <https://doi.org/10.1017/S0047404517000185>
- Eckert, Penelope. 2000. *Linguistic variation as social practice: The linguistic construction of social meaning in Belten High*. Oxford: Blackwell.
- Eckert, Penelope. This volume. The individual in the semiotic landscape. *Glossa: A Journal of General Linguistics. Special Collection: Individuals, communities, and sound change*.
- Edwards, Viv. 1986. *Language in a black community*. Clevedon: Multilingual Matters.
- Edwards, Walter. 1992. Sociolinguistic behavior in a Detroit inner-city black neighborhood. *Language in Society* 21. 93–115. DOI: <https://doi.org/10.1017/S0047404500015050>
- Fagyal, Zsuzsanna, Samarth Swarup, Anna Maria Escobar, Les Gasser & Kiran Lakkaraju. 2010. Centers and peripheries: Network roles in language change. *Lingua* 120. 2061–2079. DOI: <https://doi.org/10.1016/j.lingua.2010.02.001>
- Forrest, Jon & Robin Dodsworth. 2016. Towards a sociologically-grounded view of occupation in sociolinguistics. *University of Pennsylvania Working Papers in Linguistics* 22. Online: <http://repository.upenn.edu/pwpl/vol22/iss2/5>.
- Galaskiewicz, Joseph & Stanley Wasserman. 1981. A dynamic study of change in a regional corporate network. *American Sociological Review* 46. 475–84. DOI: <https://doi.org/10.2307/2095266>
- Gal, Susan. 1979. *Social determinants of linguistic change in bilingual Austria*. New York: Academic Press.
- Garrett, Andrew & Keith Johnson. 2013. Phonetic bias in sound change. In Alan C. L. Yu (ed.), *Origins of sound change*, 51–97. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780199573745.003.0003>
- Guy, Gregory R. 1980. Variation in the group and the individual: The case of final stop deletion. In William Labov (ed.), *Locating language in time and space*, 1–36. New York: Academic Press.
- Hoffman, Michol F. & James A. Walker. 2010. Ethnolects and the city: Ethnic orientation and linguistic variation in Toronto English. *Language Variation and Change* 22. 37–67. DOI: <https://doi.org/10.1017/S0954394509990238>
- Hughes, Everett Cherrington. 1951. Work and the self. In *The Sociological Crossroads*, 313–323.
- Hughes, Everett Cherrington. 1971. *The sociological eye: Selected papers*. Chicago, IL: Aldine.
- Johnson, Daniel Ezra & Jennifer Nycz. 2015. Partial mergers and near-distinctions: Stylistic layering in dialect acquisition. *University of Pennsylvania Working Papers in Linguistics* 21. Online: <http://repository.upenn.edu/pwpl/vol21/iss2/13>.
- Kerswill, Paul & Ann Williams. 2000. Creating a new town koine: Children and language change in Milton Keynes. *Language in Society* 29. 65–115. DOI: <https://doi.org/10.1017/S0047404500001020>
- Khan, Arfaan. 2006. A sociolinguistic study of Birmingham English: Language variation and change in a multi-ethnic British community. Ms. Lancaster University.
- Krackhardt, David. 1988. Predicting with networks: Nonparametric multiple regression analysis of dyadic data. *Social networks* 10. 359–81. DOI: [https://doi.org/10.1016/0378-8733\(88\)90004-4](https://doi.org/10.1016/0378-8733(88)90004-4)

- Labov, William. 1972. The linguistic consequences of being a lame. In William Labov (ed.), *Language in the Inner City: Studies in the Black English vernacular*. Philadelphia, PA: University of Pennsylvania Press.
- Labov, William. 1991. The three dialects of English. *New Ways of Analyzing Sound Change* 5. 1–44.
- Labov, William. 2001. *Principles of linguistic change, vol. 2: Social factors*. Malden, MA: Blackwell.
- Labov, William, Sharon Ash & Charles Boberg. 2006. *The atlas of North American English: Phonetics, phonology and sound change*. Berlin: Walter de Gruyter. DOI: <https://doi.org/10.1515/9783110167467>
- Lippi-Green, Rosina. 1989. Social network integration and language change in progress in an alpine rural village. *Language in Society* 18. 213–34. DOI: <https://doi.org/10.1017/S0047404500013476>
- MacKenzie, Laurel. This volume. Perturbing the community grammar: Individual differences and community-level constraints on sociolinguistic variation. *Glossa: A Journal of General Linguistics. Special Collection: Individuals, communities, and sound change*.
- Milroy, Lesley. 1987. *Language and social networks*. Second Edition; Oxford: Blackwell.
- Milroy, Lesley & James Milroy. 1992. Social network and social class: Toward an integrated sociolinguistic model. *Language in Society* 21. 1–26. DOI: <https://doi.org/10.1017/S0047404500015013>
- Mizruchi, Mark S. 1993. Cohesion, equivalence, and similarity of behavior: A theoretical and empirical assessment. *Social Networks* 15. 275–307. DOI: [https://doi.org/10.1016/0378-8733\(93\)90009-A](https://doi.org/10.1016/0378-8733(93)90009-A)
- Rickford, John & Mackenzie Price. 2013. Girlz II women: Age-grading, language change and stylistic variation. *Journal of Sociolinguistics* 17. 143–179. DOI: <https://doi.org/10.1111/josl.12017>
- Rohe, William M. 2011. *The research triangle: From Tobacco Road to global prominence*. Philadelphia, PA: University of Pennsylvania Press. DOI: <https://doi.org/10.9783/9780812207514>
- Rosen, Nicole & Crystal Skriver. 2015. Vowel patterning of Mormons in Southern Alberta, Canada. *Language & Communication* 42. 104–115. DOI: <https://doi.org/10.1016/j.lang-com.2014.12.007>
- Sankoff, Gillian & Hélène Blondeau. 2007. Language change across the lifespan: /r/ in Montreal French. *Language* 83. 560–88. DOI: <https://doi.org/10.1353/lan.2007.0106>
- Sharma, Devyani. 2017. Scalar effects of social networks on language variation. *Language Variation and Change* 29. 393–418. DOI: <https://doi.org/10.1017/S0954394517000205>
- Siegel, Jeff. 2010. *Second dialect acquisition*. Cambridge: Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511777820>
- Stanford, James & Laurence Kenny. 2013. Revisiting transmission and diffusion: An agent based model of vowel chain shifts across large communities. *Language variation and change* 25. 119–153. DOI: <https://doi.org/10.1017/S0954394513000069>
- Stevens, Mary, Jonathan Harrington & Florian Schiel. This volume. Associating the origin and spread of sound change using agent-based modeling applied to /s/-retraction in English. *Glossa: A Journal of General Linguistics. Special Collection: Individuals, communities, and sound change*.
- Thomas, Erik R. 1997. A rural/metropolitan split in the speech of Texas Anglos. *Language Variation and Change* 9. 309–32. DOI: <https://doi.org/10.1017/S0954394500001940>
- Thomas, Erik R. 2001. *An acoustic analysis of vowel variation in New World English* (American Dialect Society 85). Durham, NC: Duke University Press.

- Tillery, Jan & Guy Bailey. 2004. The urban South: Phonology. In Bernd Kortmann, Edgar W. Schneider, Kate Burridge, Rajend Mesthrie & Clive Upton (eds.), *A handbook of varieties of English Phonology 2*. 325–337. Berlin, New York: Mouton de Gruyter.
- Trudgill, Peter. 1986. *Dialects in contact*. Oxford: Blackwell.

How to cite this article: Dodsworth, Robin. 2019. Bipartite network structures and individual differences in sound change. *Glossa: a journal of general linguistics* 4(1): 61.1–29. DOI: <https://doi.org/10.5334/gjgl.647>

Submitted: 04 March 2018 **Accepted:** 09 January 2019 **Published:** 18 June 2019

Copyright: © 2019 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



Glossa: a journal of general linguistics is a peer-reviewed open access journal published by Ubiquity Press.

OPEN ACCESS The Open Access logo, featuring the words 'OPEN ACCESS' followed by a circular icon containing a stylized padlock with an open keyhole.