

# Attribute Networking: A Sociolinguistic Technique for Modeling Subjective Social Space<sup>\*</sup>

ROBIN DODSWORTH  
*The Ohio State University*

## 0. Introduction

Recent approaches to studying the social meaning of linguistic variation have given a central role to speakers' subjective experiences and locally recognized social categories, relying heavily on ethnographic methods. While such approaches have made great strides toward explaining individual- and community-level linguistic performance, they sacrifice some of the objectivity and systematicity of "first-wave" approaches, in Eckert's (2002) terms.

This paper introduces a network-based technique, *attribute networking*, as an attempt to represent subjective conceptions of local social structures in a way that facilitates quantitative analysis, as the social networks framework does. The approach aims to find empirical justifications for delineating subjective social categories and abstracting from them. It also takes steps toward modeling perceived connections between local, concrete social facts and broader structures (cf. Milroy and Milroy 1992, Eckert 2000). It is not intended to replace but to complement existing methods. The technique is first described and then illustrated with a study of /l/ vocalization in Worthington, Ohio, a Columbus suburb.

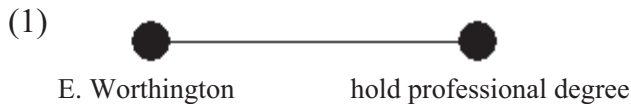
## 1. Attribute Networking

The use of networks presented here draws on Bearman, Faris, and Moody's (1999) network-based representation of a series of events in the Chinese Revolution. In their model, nodes represent single events, and ties between nodes represent temporal relationships. Attribute networking, the present technique, uses networks to model community members' stated conceptions of local social processes and categories. Its networks are not *social* networks: nodes represent socially meaningful characteristics of people and places in the community, and a tie

---

<sup>\*</sup> Many thanks especially to Don Winford for commenting on numerous drafts and for his continuing encouragement. Thanks also to others whose suggestions have greatly improved the paper: Jim Moody, Galey Modan, Keith Johnson, Mary Beckman, and Scott Schwenter.

between two nodes denotes a perceived association between the two attributes that the nodes represent. For instance, (1) indicates that according to at least one informant, living on the east side of Worthington and holding a professional degree are associated with one another. (1) does *not* mean the informant believes *everyone* in East Worthington holds a professional degree, nor that everyone in Worthington who holds a professional degree lives in East Worthington.



In this respect, attribute networks are similar to social networks: a tie between two people in a social network does not indicate that they socialize exclusively with each other, nor that they spend all or even most of their time together.

The social perceptions represented in the network are gathered from ethnographic interviews during which informants are asked to talk about the community’s social space. A single attribute network is constructed to represent as much information as possible from each ethnographic interview. The process is illustrated here with an excerpt from an interview with Ann, one of the Worthington informants. The boldface words in the excerpt map onto attributes in the matrix in (2). A matrix represents the same information as a network.

There are a lot of people that are educators that live [in **Worthington proper**], a lot of professors at OSU...it’s mostly **professional** people. Now, that said, in the **outskirts** of Worthington which is considered Columbus land but Worthington schools, there are a lot, and where most of the **apartments** are, I would say that that’s where a lot of the **single parent** families live, and they would not necessarily be the professionals but more of the clerical or even factory workers...That would be the **non-professionals**, a lot of, in those areas...

(2)	W proper	Professionals	Outskirts	Non-professionals	Single parent	Apartments
W proper	0	1	0	0	0	0
Professionals	1	0	0	0	0	0
Outskirts	0	0	0	1	1	1
Non-professionals	0	0	1	0	1	1
Single parent	0	0	1	1	0	1
Apartments	0	0	1	1	1	0

A 1 represents a stated association between two attributes (i.e., a link or tie), and a 0 represents the lack of a stated association. Some potential attributes from Ann’s interview, such as being a clerical worker, are not represented in the matrix; the level of detail used in representing the spoken discourse as a matrix can vary. In Ann’s case, it seemed that clerical and factory workers exemplified the category “non-professional.”

## *Attribute Networking*

An individual attribute network is understood to represent the speaker's perceptions of the community's social space. Clearly this way of eliciting and encoding subjective information carries a set of complications. Unlike the excerpt from Ann's interview, speakers often do not talk about their conceptions of social space in a way that translates easily into sets of nodes and ties. Many speakers described sets of social attributes but either did not link them to one another very explicitly, or did so with considerable hedging. Speakers also discussed sets of social characteristics that were not easily represented with discrete nodes. Other complications, including the likelihood that no speaker is likely to mention all the attributes that he/she considers socially salient, are minimized by intercalating the individual networks from every interview into an aggregate network (AN).

The AN is the union of individual networks. Shared nodes appear only once. A tie between nodes X and Y is valued according to the number of individual networks in which it appears. Therefore, the AN unites individuals' conceptions of the social space to form a system in which diverse views are represented, points of agreement among informants are apparent, and the elements of social space can be seen in relation to one another. The AN has the ability—valuable in the Worthington study—to represent disagreement among community members' conceptions of the social space because any attribute can be linked to a set of other attributes that conflict with one another. A second advantage of the AN is that it need not contain discrete categories because all nodes may be connected. In the Worthington AN, nodes that represent relative affluence and nodes that represent relative poverty are co-members of a connected set of nodes, just as members of a social network may be more or less central but still linked, directly or indirectly. This characteristic facilitates a treatment of social categories as fluid, overlapping entities with dynamic, negotiable boundaries, as in Mendoza-Denton's (1997) study in which social categories are “parts of a single coherent system that is always fluid, always changing, precariously equilibrated, and constantly innovating on itself” (37).

Third, and most relevant to Milroy and Milroy's (1992) search for a general social theory, the AN can show links between individuals' conceptions of the local community's social structure and macro-level phenomena, as well as between separate dimensions of social identity. As an example of the former, some Worthington informants linked career-motivated transience to the perceived lack of community involvement in West Worthington. Mobility, a macro-social phenomenon, thus interacts with community involvement, a particularly important aspect of Worthington identity.

The AN is used to identify the most broadly salient aspects of social identity in the community. The goal is to find nodes or subsets of nodes that are structurally important as determined by quantitative criteria. Structurally important nodes are likely to represent characteristics that many community members consider socially meaningful, because they are mentioned by multiple informants and/or are linked to multiple nodes. Other nodes are likely to be connected to one another only through these important nodes, except in

particularly dense networks. For those reasons, the characteristics associated with important nodes are good candidates for social variables in quantitative analysis. However, isolated nodes may not provide an accurate picture of the range of social identities in the community, especially if identities are assumed to be dynamic and fluid. An alternate strategy (left for future work) is to consider dense or central subsets of nodes and evaluate their meaning with respect to the rest of the aggregate network.

## **2. An Illustration of Attribute Networking: /l/ Vocalization in Worthington**

### **2.1. Worthington**

Worthington, Ohio, was established by settlers from New England in 1803, predating Columbus by nine years. Although Worthington is considered a northern suburb of Columbus, an ongoing period of urban expansion has led Columbus to completely surround Worthington. During the 1970s, Columbus and Worthington agreed to define the boundary of the latter's affluent school district such that it encompassed not only the entire city of Worthington but also some of the surrounding areas of Columbus. Since then, countless residential neighborhoods have emerged in those areas and their populations have exploded. Unlike 20 years ago, the majority of students attending Worthington City Schools live in Columbus. Yet Worthington, now a city of close to 15,000 people compared to over 700,000 in Columbus, remains politically and socially distinct despite being swallowed by urban growth. Ninety-four percent of its residents identify as white, in contrast to 67.9% in Columbus, and its 2000 median annual household income of close to \$70,000 dwarfed the Columbus median of under \$40,000. Perhaps the most important component of Worthington's distinctiveness is what many residents call its "colonial feel," or the collective consciousness of its New England roots. Old Worthington, the space occupied by the original 1803 village, is governed by a set of strictly enforced architectural guidelines geared toward maintaining a colonial atmosphere. A village green lies at the center of Old Worthington, and several original New England-style buildings still stand as museums or churches. Regular public events, such as the annual Founder's Day, refer to the city's heritage.

Yet these traditions preserve not historical knowledge but collective imagination and distinctive community identity. A Worthington resident in her 80s who has served as the local historical society's curator explained that Worthington's historical consciousness, of which high property values are partly a byproduct, has been cultivated only since the mid-twentieth century. The choice of Worthington as a testing ground for attribute networking rests primarily on two facts. First, Worthington consciously constructs and maintains an identity that makes it a cohesive, distinct community. Second, the fact that it is (arguably) the most economically and socially homogeneous Columbus suburb precluded the possibility of the study presupposing familiar class-based categories, forcing the search for locally perceived social structures (cf. Rickford 1986).

## **2.2. Finding and Interviewing Community Members**

The linguistic data presented here is extracted from 21 ethnographic interviews, each lasting approximately one hour. The speakers lived either in Worthington or in surrounding areas inside the school district and ranged in age from 15 to over 80. All speakers younger than 30 grew up in Worthington and all speakers over 30 had lived in Worthington for at least 20 years. Informants were recruited using a “friends of friends” process. During the interviews, informants were asked to talk about what types of social groups or divisions, or what types of people, existed in Worthington (cf. Mendoza-Denton 1997:71).

## **2.3. The Worthington Aggregate Network**

Intercalating the individual networks from the 21 interviews yielded an AN with 138 nodes. The network consists of one large component containing 100 nodes (pictured in the appendix) as well as 14 small components. The large component can be viewed as having two sections, connected only through the nodes labeled “live in Worthington proper” and “sense of community.” The larger and smaller sections roughly correspond to living outside and inside the Worthington city limits, respectively. Strikingly, only 17 ties have values greater than 1—recall that a tie’s value indicates the number of speakers who have stated an association between the nodes it connects—and the highest value is 5 out of a possible 21 (the number of speakers). The lack of high-valued ties may indicate disagreement: if all informants had been in perfect agreement as to the community’s social structure, then many ties would have values near the maximum of 21, the only low values resulting from differences in the ways that informants stated their perceptions or from differences in what informants remembered or were willing to say. In that case, the ability to represent conflicting views is critical.

Several quantitative measures were employed to identify the structurally important nodes in the aggregate network.

1) *Betweenness centrality*: the relative number of geodesics, or shortest paths between two nodes, that a node lies on (Wasserman and Faust 1994). This property identifies nodes that are responsible for uniting other nodes and holding the network together. For example, given that the node representing living in Worthington’s outskirts lies on the geodesic between the nodes representing more crime and paying Columbus taxes, the latter two attributes can be assumed to be related to one another only insofar as they are both properties of living in the outskirts. The two nodes with the highest betweenness centrality corresponded to living in Worthington proper and having a sense of community.

2) *Degree*: the number of ties incident to a node, or the number of nodes it is adjacent to. Nodes with high degrees are likely to represent attributes that more than one informant has mentioned, so degree is an indicator of how broadly recognized a given attribute is. The nodes with the highest degree corresponded to living in an apartment, living in Worthington proper, and living in the outskirts.

3) *Incidence with at least one tie valued greater than 1*. A tie’s value reflects the number of individual networks it appears in, or the number of informants who

have expressed the association. The higher a tie's value, the more likely it reflects a perceived social fact that is recognized throughout the community. A node incident to a high-valued tie therefore represents an attribute that multiple informants have mentioned in connection with another attribute. The node incident to the greatest number of valued ties, and also to the highest-valued tie, corresponded to living in an apartment.

4) *Cutpoints*: nodes whose deletion would create more components (distinct sets of connected nodes). Therefore, a cutpoint holds together at least two sets of nodes that would be disconnected from one another without that node (Wasserman and Faust 1994). Cutpoints may connect densely linked clusters of nodes that represent separate ideas. An interesting cutpoint in the aggregate network is the node representing having a sense of community. It connects the two sections of the largest component; one section roughly corresponds to living in Worthington proper, and the other roughly corresponds to living in the outskirts. Cutpoints are never at the extreme periphery, though not every cutpoint is a particularly central node.

5) *Continuous coreness*: treats nodes as having varying degrees of coreness based on the strength of their ties. A node whose ties have high values thus has a greater degree of coreness than a node with low-valued ties (Borgotti and Everett 1999). The 11 nodes selected by this metric are nearly a proper subset of the nodes incident to ties valued greater than 1.

All calculations were performed by the network software UCINET (Borgotti et. al. 2002). (3) shows the number of nodes selected by each metric, and the criteria for node selection in each case.

(3) Measures of node importance	Criteria for node selection	No. of selected nodes
Normalized betweenness centrality	At least one standard deviation above the mean	11
Normalized degree	At least one standard deviation above the mean	10
Incident to a tie valued greater than 1	All	21
Cutpoints	All	21
Continuous coreness	At least one standard deviation above the mean	11

(4) Structurally important attributes	What each attribute refers to
Live in Worthington proper	Space inside the W political boundary
Live in the outskirts	Outside the W political boundary, inside school district
Live in Old Worthington	Space of the original 1803 village
No community involvement	No participation in W organizations
Live in Colonial Hills	Neighborhood in W proper
Live in a "grand" house	(Unclear)
Live in an apartment	(Self-evident)

The attributes in (4) were selected by all five metrics and were therefore taken to be the most broadly recognized among the informants. For that reason, it was hypothesized that they would interact systematically with sociolinguistic variables in Worthington. The first four attributes were used as the basis for constructing independent variables for a VARBRUL statistical analysis of /l/ vocalization. The other three attributes were either too vague to permit confident categorization of speakers or not applicable to any of the informants. The first independent variable corresponded to location of residence and had three variants: 1) Old Worthington, 2) Worthington proper but outside Old Worthington, and 3) the outskirts. The second independent variable encoded whether the speaker was generally involved in the community, according to his/her own estimation. Finally, independent variables were established for sex and age.

#### **2.4. /l/ Vocalization**

Post-vocalic (dark, coda) [ɫ] differs from pre-vocalic [l] in that it involves greater retraction of the tongue body and a delayed raising of the tongue tip (Sproat and Fujimura 1993). In several varieties of English, the tongue tip may be raised only minimally, resulting in little or no alveolar contact and producing a “vocalized” variant (Hardcastle and Barry 1985). /l/ vocalization has been documented in several varieties of English (e.g., Ash 1982, Carver 1993, Horvath and Horvath 2002) and is common in southern and central Ohio.

A study of /l/ vocalization in Worthington was performed using 724 tokens of coda /l/ extracted from the 21 interviews. The set of tokens includes /l/ occurring syllable-finally as in *all* or *almost*, as the first segment in a coda consonant cluster as in *cold*, and as a syllabic segment as in *little*. Just as [l] and [ɫ] are not categorically distinct (Sproat and Fujimura 1993), the vocalized and unvocalized variants of post-vocalic /l/ clearly encompass overlapping sections of a continuum. It follows that categorizing tokens as vocalized or unvocalized is not entirely straightforward unless objective acoustic criteria are used (and even they will be somewhat arbitrary). In an effort to sidestep the danger of the analyst’s category boundary between [ɫ] and the vocalized variant being either too narrow or too broad to capture important variation patterns, three linguists, all native English speakers who do not vocalize /l/, were asked to categorize the tokens as closer to [ɫ] or closer to another unspecified sound (i.e., vocalized). A token was coded as vocalized if at least two of the three listeners judged it to be. 114 of the 724 tokens were coded as vocalized.

The tokens were also coded for the social factors described above as well as the following linguistic factors:

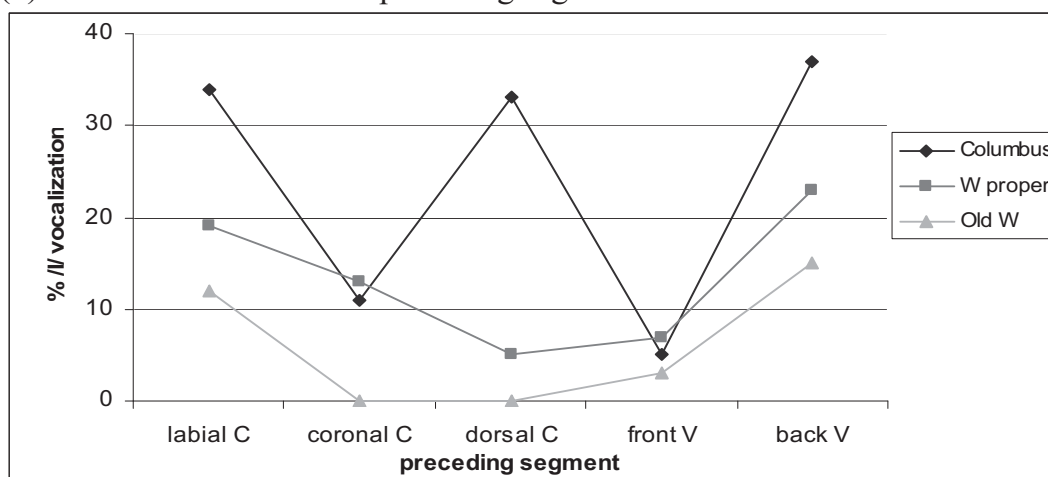
- 1) morpheme-final vs. morpheme-internal
- 2) preceding segment: labial, coronal, or dorsal consonant, or vowel (all vowels were initially coded as distinct)
- 3) following environment: pause, labial, coronal, or dorsal consonant, or vowel (vowels were not differentiated); note that when preceding a pause, /l/ could not occur in a consonant cluster.

In the first VARBRUL run, the significant factor groups were preceding segment, following segment, and location of residence. Preceding vowels were subsequently grouped as front or non-front. Another run using only the three significant factor groups yielded the results in (5)–(7). The application value is the vocalized variant, the input is .122, and  $\chi^2/\text{cell}$  is .7799.

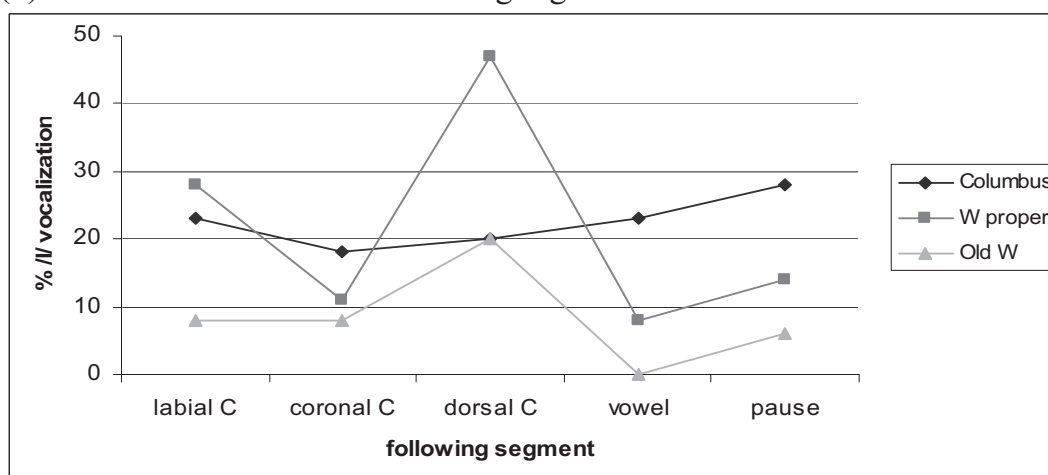
(5) Significant factor groups and weights

Preceding segment	Weight	Following segment	Weight	Location of residence	Weight
Labial consonant	.651	Labial consonant	.654	Old Worthington	.296
Coronal conson.	.393	Coronal conson.	.441	Worthington proper	.503
Dorsal consonant	.337	Dorsal consonant	.775	Outskirts	.619
Front vowel	.284	Vowel	.391		
Back vowel	.695	Pause	.524		

(6) Interaction between preceding segment and location of residence



(7) Interaction between following segment and location of residence





As shown in the cross tabulations in (6) and (7), Columbus speakers follow a strikingly different pattern from Worthington speakers, in particular with respect to their high rate of vocalization before dorsal consonants. The following segment seems not to matter for Columbus speakers, in contrast with its strong effect for both groups of Worthington speakers. In summary, Worthington speakers vocalize /l/ less frequently than speakers in the surrounding Columbus area, and their vocalization is conditioned differently by linguistic factors. Further, speakers in Old Worthington vocalize /l/ less frequently than speakers outside of Old Worthington but within the city limits.

### 3. Discussion of Results

Although the network technique identified a social characteristic that interacts with /l/ vocalization, it cannot directly explain the interaction. Even so, the aggregate network embeds each attribute within a structured system of attributes, revealing direct and indirect connections among attributes, thereby taking a step toward explaining their contextual importance. There is a node labeled “no community involvement”—one of the seven most structurally important nodes—as well as a node labeled “community involvement.” Both of these nodes are adjacent to nodes representing locations: having no community involvement is adjacent to “outskirts,” “west of the Olentangy River,” and “Worthington Hills” (a neighborhood in the outskirts); “community involvement” is adjacent to “Worthington proper.” Thus the relationship between community involvement and place of residence is well represented in the network, reflecting the fact that many speakers addressed the links between geography and community. One of the most elaborate discussions of this theme was provided by Rita, who lives in Old Worthington, has strong community involvement, and vocalizes /l/ rarely:

I guess I see the boundaries of Worthington as, um, the communities that do get involved, the parts of the communities that do get involved.

...

And now I think you know there are no boundaries between Worthington and Polaris [a mall north of Worthington]. I mean it's all, there's no *physical* definition. And the way things uh, in Columbus, and therefore in Worthington, are growing and developing and expanding as, as far as they can go I don't think there *is* gonna be anything that that defines one community to the next. Um, and I guess that's just, that's just city development, I don't know. I don't know, pretty soon we're gonna look like Tokyo, where, one, one city just, just flows right into the next.

...

We're too big. We're too big, already. *Because* we don't know each other. *Because* we're not, involved, you know because, because you can now say “Oh well this is the Worthington *school* district,” but you pay Columbus *taxes*, but are you, are you affiliated with Worthington? Are you, you know, are you *involved*? And I think maybe it's because they're too far out and, and don't feel a connection, I don't know.

Other speakers draw similar connections between urban growth and the loss of community identity and cohesion. Dana, a resident of the outskirts, described her neighborhood as “community sprawl,” lacking cohesion, and noted that people

living inside the city limits of Worthington are highly conscious of that boundary. Dana herself does not feel like a true member of the Worthington community. Thus Rita, Dana, and others perceive a connection between geography and one's status and involvement in the community.

My analysis of these facts is that resistance to /l/ vocalization is part of the effort to maintain Worthington's identity as a community distinct from Columbus, a task that is increasingly difficult as Columbus annexes and develops areas surrounding Worthington. Thus the pattern of linguistic variation presented here is a product of and a tool in the construction of a range of stances regarding Worthington's community identity in the face of urban sprawl. By not vocalizing /l/ (unlike a great many speakers in central Ohio), speakers like Rita resist the disappearance of the Worthington community. Ethnographic observation (not described here for lack of space) suggests that there exist two broad groups: people who want to preserve a certain amount of exclusivity and closeness in the Worthington community, and people who recognize this desire in others but either do not take part in it or actively dismiss it. Resistance to /l/ vocalization among people with strong civic involvement mirrors their unwillingness to let go of Worthington's identity as a small, relatively affluent town distinct from Columbus and its sprawl. Many of the speakers who do not share in this resistance are like Dana in that they live outside the city boundary and therefore do not benefit from the prestige or affluence associated with Worthington except through the school system. Within the two broad groups, of course, there is variation both in ideology and in use of the linguistic variables. There is also ideological variation among Old Worthington residents; some of them resist any efforts to exclude Columbus residents from the Worthington community, and those who work to preserve Worthington's identity have varying reasons for doing so. Accordingly, individuals use linguistic variables in unique ways to construct their own styles. I leave these topics for future work.

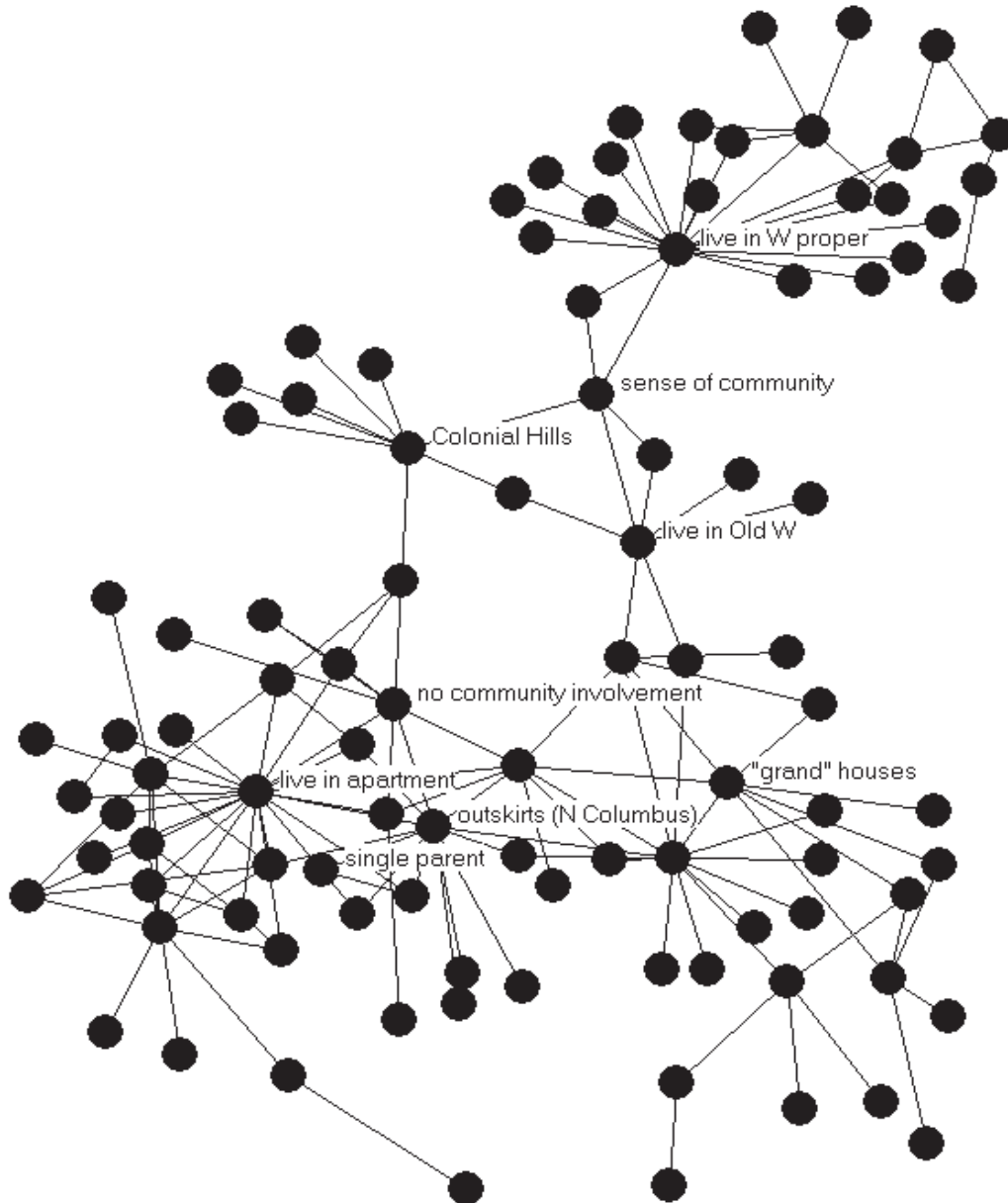
#### **4. Conclusion**

Attribute networking is potentially useful to sociolinguists because it can be used to represent subjective information in a way that facilitates quantitative linguistic analysis. Linguistic variation is analyzed with respect to social variables that are not analyst-imposed but rather derived directly from community members' conceptions (cf. Rickford 1986), following the ethnographic tradition. Divergence among community members' views of social space is easily represented, and widely recognized social boundaries and issues are likely to be identified. Finally, by virtue of being a network-based technique, attribute networking can show perceived connections between concrete, local social phenomena and broader, abstract categories and processes; in Worthington, for example, there is a strongly felt connection between geography and community status. The technique thus has the potential to contribute to current efforts to unite the particular and the universal (Meyerhoff 2002:543), as well as efforts to link observable social network patterns to abstract categories such as class (Milroy and Milroy 1992).

## Attribute Networking

Clearly this paper does not exhaust attribute networking's potential. The paper does not, for instance, exploit the fluid, interconnected nature of categories in the aggregate network, a property that makes attribute networking compatible with the community of practice concept. However, I hope to have demonstrated that it can be useful in exploring some questions of interest to variationists.

### Appendix: Largest component of the Worthington aggregate network



## References

Ash, Sherry. 1982. The vocalization of intervocalic /l/ in Philadelphia. *SECOL Review* 6(3):162-175.

Robin Dodsworth

- Bearman, Peter S., Robert Faris, and James Moody. 1999. Blocking the future: New solutions for old problems in historical social science. *Social Science History* 23(4):501-533.
- Borgotti, Steven P., and Martin G. Everett. 1999. Models of core/periphery structures. *Social Networks* 21(4):375-395.
- Borgotti, S. P., M.G. Everett, and L.C. Freeman. 2002. Ucinet for Windows: Software for social network analysis. Harvard: Analytic Technologies.
- Carver, Craig. 1993. Regional speech variation in Wisconsin. In T. Frazer, ed., "Heartland" *English: Variation and Transition in the American Midwest*, 199-214. Tuscaloosa and London: University of Alabama Press.
- Eckert, Penelope. 2000. *Linguistic Variation as Social Practice: The Linguistic Construction of Identity in Belten High*. Malden, MA: Blackwell.
- Eckert, Penelope. 2002. Constructing meaning in sociolinguistic variation. Paper presented at the Annual Meeting of the American Anthropological Association, New Orleans, November.
- Hardcastle, William, and William Barry. 1985. Articulatory and perceptual factors in /l/ vocalizations in English. *Journal of the International Phonetic Association* 15(2):3-17.
- Horvath, Barbara M., and Robert J. Horvath. 2002. The geolinguistics of /l/ vocalization in Australia and New Zealand. *Journal of Sociolinguistics* 6(3): 319-346.
- Mendoza-Denton, Norma Catalina. 1997. Chicana/Mexican identity and linguistic variation: An ethnographic and sociolinguistic study of gang affiliation in an urban high school. Ph.D. diss., Stanford University.
- Meyerhoff, Miriam. 2002. Communities of practice. In J. K. Chambers, P. Trudgill, and N. Schilling-Estes, eds., *The Handbook of Language Variation and Change*, 526-548. Malden, MA: Blackwell.
- Milroy, Lesley, and James Milroy. 1992. Social network and social class: Toward an integrated sociolinguistic model. *Language in Society* 21(1):1-26.
- Rickford, John. R. 1986. The need for new approaches to social class analysis in sociolinguistics. *Language and Communication* 6(3):215-221.
- Sproat, Richard, and Osamu Fujimura. 1993. Allophonic variation in English /l/ and its implications for phonetic implementation. *Journal of Phonetics* 21(3): 291-312.
- Wasserman, Stanley, and Katherine Faust. 1994. *Social Network Analysis: Methods and Applications*. Cambridge: Cambridge University Press.

222 Oxley Hall  
1712 Neil Ave.  
Columbus, OH 43210-1298

dodswor@ling.ohio-state.edu