Methods in (applied) folk linguistics

Getting into the minds of the folk

Dennis R. Preston
Oklahoma State University

This paper deals with data gathering and interpretation in folk linguistics, but, as the parenthetical title suggests, it is not limited to any prejudged notion of what approaches or techniques might be most relevant to the wide variety of concerns encompassed by applied linguistics.

I will conceive of folk linguistics broadly here, including not only the comments that nonlinguists make about linguistic topics but also the reactions they have to varieties of language and language use, including overt as well as subconscious responses. In other words, findings from the social psychology of language (i.e., attitude studies) are taken to be a part of folk linguistics, along with data derived from more conscious discourse or from operational tasks. This is a different position from one that assigns the label “folk linguistics” to more conscious responses and “language attitudes” to relatively more unconscious ones (e.g., Niedzielski & Preston 2003: xi). There is also no pretense here that folk linguistics makes use of unique interpretive strategies. Its quantitative and qualitative data have been looked at from common statistical, discoursal, content analytic, cultural studies, and other points of view.

I also most emphatically use the term folk in folk linguistics to refer to all persons except academic linguists, just as linguists would be folk in a study of folk botany, folk chemistry, etc…. I definitely do not use the term to refer to rural, marginalized, less educated, or romanticized (‘quaint’) groups. We’re all folk when we step into the world of traditional knowledge and ways of behaving outside our own technical training. Even then, folk knowledge may be at work when more subconscious modes prevail, although, as in the language attitudes of linguists, for example, they may be suppressed from overt comment or behavior by professional knowledge.

There have developed, however, techniques of collection and interpretation that are associated with folk linguistics, and I will characterize them here as traditional, operational, experimental, and discoursal, though I will be hard-pressed to always distinguish between them.

Traditional folk linguistic data come from a rich repository of belief embedded in tales, sayings, rituals, and practices. Büld (1939) looks at dialect variation in part on the basis of folk routines that involve imitations. Giles et al.’s obviously titled “Talk is cheap,” but “my word is my bond” (1991) takes its motivation from traditional sayings. A particularly important area of traditional folk linguistic belief derives from standard language pressures on speakers of nonstandard varieties and may be studied in educational materials, media punditry, and other popular culture outlets. Finegan (1980), Lippi-Green (1997), and Milroy & Milroy (1999) provide excellent examples for both the history
and relatively current status of varieties of English vis-à-vis so-called Standard English as revealed in such sources, and Lodge (1993) is particularly good in providing an historical account of such matters for French.

At the level of practice, folk linguistics may be indistinguishable from the ethnography of language, in which investigation of the status of languages and language varieties has come to be known as language ideology; this initiative seeks to show that how language is treated in the traditional practices of a speech community may clearly illustrate underlying beliefs and belief systems (e.g., Schieffelin et al. 1998).

In discussing operational methods I will focus on several research techniques associated with perceptual dialectology, a major sub-branch of folk linguistics that was intensively pursued in the mid-20th century among the Dutch and Japanese and is now undergoing a world-wide revival.

The section on experimental approaches presents a review of some of the early methods used by social psychologists to study language attitudes (e.g., the matched-guise technique) and closes with a discussion of more recent work that uses subtle manipulation of auditory (e.g., sound resynthesis) and visual stimuli and records respondent reactions in timed and even eye-tracked and brain-imaged monitoring.

The final section on discoursal methods will characterize attempts to record and analyze commentary on language and its use by nonlinguists, and it concludes with a review of a variety of more recent discourse analytic investigations of such conversations.

1. Traditional approaches

The methodologies of data collection and interpretation in the study of traditional beliefs and practices with regard to language are of two sorts: the folkloristic-literary and the anthropological-cultural.

1.1 Folkloristic-literary traditional approaches

In the folkloristic-literary approach, the data are acquired in various ways — from fieldwork interviews and questionnaires to introspection, to the extraction of data from various public sources, including now the rich resources of the internet. The interpretations of these data are often but not always those common to literary or cultural studies: the interpreter, with a rich background of cultural and historical knowledge of the speech community (or communities), provides insightful comments on and generalizations about the meaning and significance of the data.

Reddy (1979), for example, examines the conduit metaphor of language, which, according to Lakoff & Johnson (1980: 10), structures language about language as follows:

IDEAS (OR MEANINGS) ARE OBJECTS
LINGUISTIC EXPRESSIONS ARE CONTAINERS
COMMUNICATION IS SENDING

This lay notion of language is derived from such folk metaphors as the following:

It’s hard to get that idea across to him.
Your reasons came through to us.
It’s difficult to put my ideas into words.
Try to pack more thoughts into fewer words.
You simply can’t stuff ideas into a sentence any old way.
Don’t force your meanings into the wrong words.
[etc…] (Lakoff & Johnson 1980: 11)
Although Lakoff & Johnson go on to illustrate how the conduit metaphor does not align with academic notions of language by illustrating how it ignores contextual features and alternative interpretations of messages based on speaker identities, the exercise itself shows how one may derive a dominating cultural folk definition from folk metaphor.

Perhaps more interesting to applied concerns would be Lakoff & Johnson's investigation of the folk linguistic notion of *argument* as expressed in the metaphor *argument is war*, which they derive from the following folk metaphors:

- Your claims are *indefensible*.
- He *attacked* every weak point in my argument.
- His criticisms were *right on target*.
- I *demolished* his argument.
- I've never *won* an argument with him.

[etc…] (Lakoff & Johnson 1980: 4)

They also suggest, however, that such cognitively-controlling metaphors are not cultural universals. For example, they imagine that an equally compelling metaphor in another culture might be one in which argument is seen as dance:

- ...[T]he participants are seen as performers, and the goal is to perform in a balanced and aesthetically pleasing way. In such a culture, people would view arguments differently, experience them differently, carry them out differently, and talk about them differently. But we [italics in original] would probably not view them as arguing at all; they would simply be doing something different. (Lakoff & Johnson 1980: 5)

It will not take an applied linguist long to see that any interest in cross-cultural communications, whether embedded in second and foreign language instruction, language and politics, or many other applied concerns, would be enhanced by such comparative folk linguistic information.

In another such approach, one relying heavily on media documents, Finegan (1980) gives an account of the public reception to the 1961 publication of *Webster's Third New International Dictionary*, a publication that, unlike its predecessors, used popular culture or distinctly non-literary sources in its definitions and provided usage citations that sanctioned or only minimally warned against long-standing shibboleths (e.g., *ain't*). Finegan examines and interprets commentary from nonlinguists from a wide range of media sources and identifies three major themes — accusations that the dictionary and its supporters were (1) *scientific*, (2) *permissive*, and (3) *democratic*.

For the first, he cites an English Professor, A. M. Tibbets: "After a time the objective, 'scientific' study of language corrupts a man. He becomes less a moralist … and more a pedant … playing with tape recorders and other gimcrackery…” [ellipses in Finegan]. He also quotes a modern languages professor who apparently found it “as unsatisfactory to divorce style and taste from linguistic arguments … as it is to restrict oneself to describing the chemical composition of the colors used by Michelangelo without reference to their effect” (both quoted in Finegan 1980: 123). The folk notion Finegan derives from these and other nonlinguist comments is one of the “…alleged mechanical approach of natural and social, scientists” (ibid.), an approach obviously filled with “gimcrackery” and no room for taste.

The accusation of *permissiveness* is based on even stronger folk notions and ones further removed from academic or esthetic concerns. Finegan quotes an American Bar Association publication, where it was noted that “…opening the floodgates to every word that is used, no matter how or by whom, and regardless of its propriety, is like the printing of paper money backed by no sound value” (ibid.). That linguistic permission may lead to perdition is not a new idea in the Anglophone world. Graddol & Swann (1988) cite former British government minister Norman Tebbit:
If you allow standards to slip to the stage where good English is no better than bad English, where people turn up filthy … at school … all those things tend to cause people to have no standards at all, and once you lose standards there’s no imperative to stay out of crime. (ellipses in Graddol & Swann 1988: 102)

Finally, Finegan notes that even democracy is decried by the dictionary detractors. He cites the well-known writing handbook author Sheridan Baker:

Good English has to do with the upper classes — and there’s the rub — with the cultural and intellectual leaders, with the life of the mind in its struggle to express itself in its intellectual best. Linguistic relativism has a fervently democratic base. ‘Science’ is only an antiseptic label for the deep social belief that we ought not to have classes at all, even among our words (quoted in Finegan 1980: 124).

I give quite a lot of space to Finegan’s work since it illustrates nicely that the high-minded and well-educated critics of Webster’s Third New International Dictionary were, in spite of their erudition, folk commentators so far as linguists are concerned. I have, however, not called them just folk commentators, for the ideas they embed in their published commentary are important themes in (at least US) folk linguistics.

1.2 Anthropological-cultural traditional approaches
In anthropological-cultural approaches, the data are more often acquired through long-term observation of, or participant observation in, local behavior. Irvine (2001) makes the motivation for such investigation clear:

Participants in some community of discourse are not entirely objective observers of each other’s behaviors. Yet, their own acts are deeply influenced by their perceptions and interpretations of those behaviors. … Some of the most important and interesting aspects of ideology lie behind the scenes, in assumptions that are taken for granted — that are never explicitly stated in any format that would permit them also to be explicitly denied. (Irvine 2001: 25)

The main illustration of this approach to folk linguistic data collection and interpretation offered here is taken from descriptions of Javanese speaking styles and Wolof registers; the contrast between the two is detailed in Irvine (1998). In Javanese, the highest status speech styles, and even awareness of speech style differentiation, is attributed to the traditional elite class (the priyayi) (Irvine 1998: 56). These styles are calm, orderly, and not ones that display emotion (Irvine 1998: 57), but, more importantly, they are that way due to the speaker’s recognition of the addressee’s status:

The language levels … are thought of as a means of guarding the addressee’s equanimity, of avoiding angering him or her, and of expressing politeness by deferring to the addressee’s wishes and effacing one’s own. [It] is conduct that is stylized, depersonalized … because that is the behavioral environment that such a respected being’s ‘nature’ supposedly requires. (quotation marks in original, Irvine 1998: 57)

In contrast, consider the speaker-centered stylistic differentiation among the Wolof:

The Wolof metapragmatic terminology firmly identifies the two registers — ‘noble speech’ and ‘griot speech’. … In the folk theory that relates ways of speaking to kinds of speakers, the registers take the form they do because persons of the high and low ranks … are ideologically accorded certain temperamental characteristics, such as affectivity and excitability. Thus ‘noble speech’ is a flat-affect speech, while ‘griot speech’ is a high-affect, theatrical, hyperbolic style …. (quotation marks in original, Irvine 1998: 57)
At first glance, one might find that these two systems share a great deal — a deferential, flat, emotion-free high-status style and a more emotional lower-status one, but, as Irvine points out above, the Javanese higher style incorporates behavioral characteristics accorded to the addressee while the Wolof varieties reflect assumed behavioral characteristics of the speaker. Any applied linguistics work (e.g., foreign or second language or nontraditional native language instruction, language policy and planning) that did not take into account the different ideologies reflected in these speech communities might find itself attempting to modify strongly-held and usually unconscious belief systems.

Much of the interpretive work in such studies relies heavily on Silverstein’s notion of **indexicality** (e.g., 2003). Put simply, indexicality marks those parts of speech that signal speaker and contextual identities and associations. To borrow from Irvine’s Wolof example above, if one speaks in the lower, griot style, they index their lower status, although a higher-status speaker might want to index a lower-status situation or aura by using the griot style metaphorically (Irvine 1998: 57).

Irvine & Gal take the process of indexing one step further in an associative process they call **iconization**: “the attribution of cause and immediate necessity to a connection (between linguistic and social groups) that may only be historical, contingent, or conventional” (2000: 37). In other words, low-status Wolof speakers use exciting griot speech because they are (unlike nobles) excitable people, and, perhaps more importantly, griot speech is seen as exciting because the typical speakers of it are excitable.

To illustrate this from my own research, speakers of Southern American English are said to speak slowly because they are slow (but hospitable); speakers of Northern American English (especially East Coast urban Northern) speak fast because they are fast (and rude and inhospitable). Through the process of iconicity, therefore, slow speech itself is seen as dull-witted but friendly and fast speech reflects a quick mind but not a very sympathetic attitude. So it is that much nonstandard speech is said to be illogical on no basis of internal (or even mathematical) logic — ‘two negatives make a positive,’ as English-speaking folk linguists like to observe — since the speakers who utter such constructions are assumed to be illogical. That Labov presented a widely-reprinted and cited article on the separation of standard language and logical expression in 1969 has had little or no influence on this very strong correlation, at least in US folk linguistics, but I suspect in many other places as well.

Applied linguists will do well to take into account these ethnographically-derived folk linguistic notions, perhaps particularly in native language education environments where placement of children in schools and even into special education programs may reflect such strongly held beliefs about intelligence and educability on the basis of iconic relations between language and ability.

### 2. Operational approaches

Although there are many tasks that might shed light on folk conceptions of language, I will focus here on perceptual (or folk) dialectology — attempts to learn where nonprofessionals believe language differs geographically. In the oldest of these methods, respondents characterized surrounding areas on the basis of their similarity to or difference from their own speech, and a number of techniques were developed in the Netherlands and Japan in the mid-Twentieth Century to infer folk dialect boundaries from such responses. In the Netherlands, the **Pfeilchenmethode** (‘little-arrow method,’ e.g., Rensink 1955) was used, in which an arrow was drawn from one area to another when a respondent in the first said that the second area’s speech was the same as their own. Figure 1 shows an application of this technique; the dark, thick lines are the traditional (i.e., professionally-determined) dialect divisions, and the perceptual areas may be derived from clusters of sites connected by arrows. Interestingly, there are very few cases where the perception arrows cross the linguistically-
determined dialect boundaries, but there are many cases within those regions where respondents do not draw arrows, i.e., where they do not recognize the speech of a nearby site as similar.

This approach is realized in Rensink (1955) in a map of all Dutch perceptual areas and further elaborated on by Daan (1970), who developed a single map based on both perception and production data for Dutch-speaking areas. Kremer (1984) is another interesting little-arrow study of the perception of varieties by German and Dutch speakers within and across national borders.

In Japan slightly different methods were developed in realizing maps from respondent judgments, shown here in a study of alpine Japan. Mase (1964) first mapped responses to two questions:
Which sites sound the same and which a little different from the home site of the respondent? Figure 2 shows how he maps the results.

The respondent at #57 in Figure 2 identifies #58 and #59 as the same but indicates that #62, #63, #56, #55 and several sites in Nagawa are a little different. Speakers from #58 and #59 agree that their own regions are similar to one another and #57 and that the same sites in Nagawa are a little different. Although it is not shown in Figure 2, respondents from surrounding areas classify #57, #58, and #59 together. In short, the perceptual dialect area made up of these three sites is based on more complex reciprocal perceptions of similarity, on similar perceptions of minor degrees of difference, and on the perception by surrounding areas of their similarity to one another. Other Japanese techniques for determining respondent notions of speech similarity and difference include Grootaers (1964), Nomoto (1963), and Sibata (1959).

The mapping of folk notions of similarity and difference has emerged in newer studies, in which the respondents are asked to determine differences between areas such as states or other regional or political zones pre-selected by the researcher. For example, Preston (1996, 317: 20) used a four-point scale for ratings of “degree-of-difference” (1=same, 2=slightly different, 3=different, and 4=unintelligibly different) for the fifty US states, New York City, and Washington, D.C. Figure 3 is a map derived from such ratings by residents of southeastern Michigan showing these four degrees of similarity-difference.

These more recent studies make use of sophisticated statistical techniques to represent the degree of difference or similarity perceived by different groups of respondents or the internal variability of such characteristics as sex, age, and ethnicity. Similar degree-of-difference studies have been done in other areas of the US and in France (Kuiper 1999), Turkey (Demirci & Kleiner 1999), French-speaking Canada (Evans 2002), Switzerland (L’Eplattenier-Saugy 2002), and Spain (Moreno Fernández & Moreno Fernández 2002).

Figure 3. Mean degree of difference ratings for MI respondents (N=147) (Preston 1996: 318)
Tamasi (2003) provides an optional technique borrowed from cognitive anthropology, called *pilesort*; respondents distribute into as many piles as they like the regions whose dialects they feel are the same. Maps of these pile sorts based on cluster analyses allow specification of clusters at different levels of agreement. Figure 4 which represents a .25 agreement level for her respondents from Georgia clearly shows that several states — California, Hawai‘i, Alaska, Florida, and Texas — are either viewed as single dialect areas or that there is insufficient agreement to align them with other areas. The remainder of the country was sorted into six zones at this level of agreement.

Another approach to perceptual dialectology was drawn from the cultural geographical technique known as *mental mapping* (e.g., Gould & White 1974) in which respondents are asked to draw their own maps, in this case, of where they think different dialects exist. The word *dialect*, however, is assiduously avoided in folk linguistic research since, at least in Anglophone areas, it always suggests nonstandardness, not simply areal distribution. Figure 5 is a typical hand drawn map of a speech area, in this case by a respondent from Chicago, Illinois, who was asked to outline and label on a US map containing only state lines the areas of the US where people spoke differently.

Although the task is operational, one approach to interpretation mirrors the cultural studies approach outlined above for naturally occurring data. Figure 5 shows that Californians are “High Class Partying Slobs” and the speech has a “stuck-up sound,” and Hawai‘ian English is identified as “Slang.” The entire state of New York is circled, although the comment associated with it refers to only a part of New York City (“Real Bad slang terrible around Bronks”). The singling out of Chicago, Illinois (where there is *Normal talk*) and Detroit, Michigan (where the respondent identifies a
large African-American population) also provides opportunities for historical, social, and cultural interpretation.

Although such individual maps may be ethnographically interesting, perhaps they are idiosyncratic; in further perceptual work, many maps have, therefore, been combined to reveal speech community findings. These combinations were first done by tracing all respondent boundaries for a single area onto a map and tallying the correspondences. Figure 6 shows such a map of the outlines of the Tohoku dialect area (in the northernmost part of Honshu, the main island of Japan) as drawn by 60 respondents from Aichi Prefecture (also in Honshu, but quite far south of Tohoku, between Tokyo and Osaka). Figure 7 shows the results of combining the most frequent corresponding lines. All sixty respondents agree that Tohoku dialect is spoken in the area in solid black farthest north, but in the area just to the south of this total agreement, the east is regarded as more Tohoku-like (50 to 59 respondents) as opposed to the west (only 40 to 49).

This technique of combining numerous hand drawn lines to determine a perceived speech area boundary has been used in a number of studies, including Preston (1981, 1986, 1989a, 1989b), Dailey-O’Cain (1999), Kuiper (1999), Evans (2002), Benson (2003), and Montgomery (2007).

A slightly different approach to generalizing the lines drawn by respondents is taken by Inoue (1996), who, in determining the folk dialect boundaries of England, finds which counties have been included by respondents in their line-drawing task and then prepares a mental map on that basis. If political or other zones within the area to be studied are small enough or can be easily divided into subzones, this approach may be a viable alternative to tracing the precise pathways of individual lines and determining where they cluster. In their study of mental maps of Korean dialect division, Long & Yim (2002) found this second approach preferable since nearly all their respondents drew lines along prefectural boundaries.

In still another technique developed by Preston & Howe (1987), outlines of areas were traced onto a digitizing pad and percentages of respondent agreement about boundaries were calculated.
Figure 8, for example, shows the results of such a computational procedure for 147 hand drawn maps of US dialect regions provided by respondents from southeastern Michigan when a fifty percent agreement criterion is used. This computational approach has been used, with adaptations to local territories, by Long (1999), Long & Yim (2002), Montgomery (2007), Cramer (2010), and Evans (2010a).

Some more recent work in the mental mapping tradition focuses on boundary agreements from the perspective of a larger number of respondents from a single area. This allows testing of typical sociolinguistic demographic differences within a group of respondents. Demirci (2002), for example, contrasts hand drawn maps of Turkish dialect areas by gender and finds them significantly different.

A final approach to folk perception of regional varieties has to do with areal identification. Preston (1996) asked respondents from southeastern Michigan to listen to tape-recorded speech samples, carefully selected to avoid regional words or grammar, of middle-aged, European-American, college-educated males. Figure 9 shows the sites where the voices were recorded.

Figure 10 shows a cluster analysis (Euclidean distance, single linkage method ['nearest neighbor']) of the results of this task.

Figure 6. A map showing 60 Aichi respondents' boundaries for the dialect area of Tohoku (Long 1999: 180)
Figure 7. A composite map of 60 Aichi respondents’ hand drawn maps of the Tohoku dialect area (Long 1999:183)

Figure 8. A computationally generalized map of US dialect areas represented in at least fifteen percent of 147 hand drawn maps (at a fifty percent level of agreement) provided by southeastern Michigan respondents (Preston 1996:305)
These cluster analysis results (which are essentially the same as linkages shown in a Tukey post-hoc test on the same data) might suggest considerable folk acuity in placing speech samples. The northernmost voices (Coldwater and Saginaw) are linked first (i.e., joined with a '+', farthest to the left in Figure 10), showing a strong association; this pair is then linked to South Bend, the next voice south, then this group of three is linked to Muncie, the next voice to the south, but then these northern and midland four are linked to New Albany. In a professional dialect geography, New Albany would first be linked to sites south of it (Bowling Green and Nashville) before it would be linked to the northern configuration.
There is also a southern grouping, but it is not as strong as the northern one, as revealed by the fact that the linkages are farther to the right. First Nashville and Florence are linked; then they are tied to Bowling Green, although, as suggested above, a traditional dialectology would probably have first linked New Albany, Bowling Green, and Nashville and perhaps then those three to Florence. The striking fact for students of US varieties, however, is that Dothan, the southernmost voice, is not linked to the southern cluster of Bowling Green-Nashville-Florence. That southern cluster is linked first to the large northern group before all eight are finally linked to Dothan. Perhaps Dothan is phonetically so southern (it is the only /r/-less voice, although only variably so) that all other southern varieties are linked to everything to the north before it is included. That would not satisfy professional dialectologists, since many southern features (e.g., /aɪ/ monophthongization, /ɪ/-/ɛ/ conflation before nasals) would be shared by all the voices from New Albany to Dothan. The perceptual grouping teaches us which features are salient and how very distinct the southernmost variety of US English is for nonlinguists.

More recent work in area identification has used other collection, display, and interpretation strategies. Montgomery (2007), for example, asks respondents from various sites in the north of England to identify voice samples from around the country by marking on a map where they think the voice is from. He then shows, in what he calls a ‘starburst’ diagram, the relationship of each folk placement to the actual site of the sample voice. This technique does away with the forced choice linearity used in Preston (1996).


What can the applied linguist get from such studies? First, if the practitioner works in an unfamiliar speech community, beliefs about all varieties, regional and otherwise, will be helpful in general. Specifically, for example, he or she might avoid features of or even reference to regions that have a local reputation for nonstandardness or inappropriateness. Of course, the applied linguist might also like to nurture a healthier linguistic relativity, but that would be extremely hard to do without knowledge of local beliefs about varieties. Of equal importance, one would certainly want to know which varieties were heard as similar to one another or as representative of a region. In fact, how could a language planner or policy maker do without information that people from region B said their variety was ‘the same’ as that spoken in Region A while those in Region A said theirs was completely different from that of B? That is exactly the situation Wolff (1959) encountered in Nigeria when a B population wanted to incorporate the A’s territory into their own but the A’s wanted to maintain independence.

Many of the task-oriented programs I have characterized here might also have been classified as experimental, and I turn now to that tradition in folk linguistics.

3. Experimental approaches

When Hoenigswald pleaded for more attention to folk linguistics, he did not ignore “…how people react to what goes on…” (1966: 20), and that puts the subject matter of language attitudes firmly in the folk linguistic camp.

The grandfather of experimental approaches to language attitudes is the matched-guise technique (e.g., Lambert et al. 1960); in it the researcher offers two audio samples from the same person who speaks two different languages or varieties; in the 1960 version, the speaker was fluent in French and English, but the stimulus voices from the same speaker were separated by other voices so that the respondents had no idea they were hearing the same speaker twice. Respondents then rated each voice sample on a Likert scale for such characteristics as “fast-slow,” “friendly-unfriendly,” etc…. Ideally, these pairs should be elicited from the community where the listening evaluation will
be done. Very shortly after the original Lambert et al. 1960 study, samples from different speakers were introduced (technically no longer matched-guise tests), but the technique remained the same. Tucker & Lambert (1969), for example, compared white and African-American responses to several US English varieties (differentiated by both ethnicity and region). A great deal of such work is summarized and exemplified in Ryan & Giles (1982).

The most consistent finding in these studies, revealed primarily by factor analytic treatments of the paired opposite rankings, was that respondents tended to rate voice samples along two main dimensions: status (as revealed by rankings on such pairs as “intelligent-unintelligent”) and solidarity (as revealed by rankings on such pairs as “friendly-unfriendly”) (Ryan et al. 1982: 8).

In much of the work carried on in this tradition, the attempt has been to see how different varieties are evaluated for status and solidarity and how different respondents (and subgroups of respondents) differ in their evaluations. In the following I will provide an example of such work derived from the research on perceptual dialectology reported above. Figures 5 and 8 show, for example, that US respondents faced with the task of drawing dialect areas are responding to something other than geographical speech boundaries. First, as some of the labels in Figure 5 make clear, not all areas are regarded as equal on what social psychologists would call the status dimension (“Southern talk the worst English in American”); second, not all areas were singled out as distinctive as often as some others; Figure 8 shows, for example, that of the 147 map-drawers from southeastern Michigan, 138 (94%) drew a boundary around a southern speech region and only 90 (61%) around the home area. The area around New York City (area 3 in Figure 8) came in third with 80 respondents (54%). Such findings, along with researcher knowledge of area linguistic stereotypes, suggest very strongly that at least the status dimension of language attitude studies was at work here as well.

To investigate this further, and to include the solidarity dimension, respondents from several areas in the US were asked to rank the 50 states, New York City, and Washington, D.C. for correctness and pleasantness. A northern group (all from southeastern Michigan) found their own state to be uniquely correct but high pleasantness was shared with several other states, and the pleasantness core was not as high as the correctness one. A southern group (principally from Alabama) found no area to be especially correct but rated their own state and a few bordering ones as uniquely pleasant (and found some northern states to be particularly unpleasant) (Preston 1996: 310–317).

My initial interpretation of these findings suggested that respondents had two sorts of language investment. They were either status oriented (like Michiganders) and paid less attention to solidarity factors, or they were solidarity oriented (like Southerners) and paid less attention to status (Preston 1996: 317). What if

1. the more cognitively robust areas, as determined by mental mapping, were presented (rather than states and two cities), and
2. a community-determined list of opposite attributes was presented for the respondents to use for evaluation?

In a follow-up study carried out only in southeastern Michigan, a full matched-guise protocol (except that single-person voice stimuli were presented) was used. First, southeastern Michigan respondents were shown a simplified version of Figure 8, the generalized mental map of US dialects from a Michigan point of view. Second, they were asked to characterize these speech areas in any way they would like. The most frequent labels elicited were then arranged as paired opposites:

- slow – fast
- polite – rude
- snobbish – down-to-earth
Methods in (applied) folk linguistics

educated – uneducated
normal – abnormal
smart – dumb
formal – casual
bad English – good English
friendly – unfriendly
nasal – not nasal
speaks with – without a drawl
speaks with – without a twang (Preston 1999b: 363)

The results of this study for the north (area 2 in Figure 8) and the south (area 1 in Figure 8) are shown in Table 1.

Table 1. Ratings of the North and the South for twelve pairs of descriptors on a scale of 1 to 6 (* indicates the only two adjacent scores that are significantly different and ‡ indicates negative ratings.) (Preston 1999b: 366)

<table>
<thead>
<tr>
<th>South</th>
<th>Mean</th>
<th>North</th>
<th>Mean</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rank</td>
<td>Attribute</td>
<td></td>
<td>Rank</td>
</tr>
<tr>
<td>1</td>
<td>Casual</td>
<td>4.66</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Friendly</td>
<td>4.58</td>
<td>2</td>
</tr>
<tr>
<td>3</td>
<td>Down-to-earth</td>
<td>4.54</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>Polite</td>
<td>4.20</td>
<td>4</td>
</tr>
<tr>
<td>5</td>
<td>Not nasal</td>
<td>4.09</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>*</td>
<td>6</td>
<td>Down-to-earth</td>
</tr>
<tr>
<td>6</td>
<td>Normal [Abnormal]</td>
<td>‡3.22</td>
<td>7</td>
</tr>
<tr>
<td>7</td>
<td>Smart [Dumb]</td>
<td>‡3.04</td>
<td>8</td>
</tr>
<tr>
<td>8</td>
<td>No twang [Twang]</td>
<td>‡2.96</td>
<td>9.5</td>
</tr>
<tr>
<td>9</td>
<td>Good English [Bad Eng.]</td>
<td>‡2.86</td>
<td>9.5</td>
</tr>
<tr>
<td>10</td>
<td>Educated [Uneducated]</td>
<td>‡2.72</td>
<td>11</td>
</tr>
<tr>
<td>11</td>
<td>Fast [Slow]</td>
<td>‡2.42</td>
<td>12</td>
</tr>
<tr>
<td>12</td>
<td>No drawl [Drawl]</td>
<td>‡2.22</td>
<td></td>
</tr>
</tbody>
</table>

These findings shed much greater light on the assessment of status and solidarity with regard to region. These northerners find themselves superior in having no drawl or twang and also superior on the status scales of “normal,” “smart,” and “good English.” We knew that from the state ranking study, but this more detailed study shows that these northern raters actually find southern speech superior on the solidarity scales of “casual,” “friendly,” “down-to-earth,” and “polite.” This reveals a linguistic insecurity that the simple state ranking studies of pleasant and correct did not: northerners find their speech lacking in the solidarity function (Preston 1999b).

Language attitude research, if fully considered here, might make this survey twice as long, but there is no doubt about folk respondents’ awareness of categories beyond region, standardness, and solidarity. In newer approaches, however, linguistic detail rather than the overall speech style used in earlier social psychological work (e.g., Giles & Bourhis 1976) has come to the fore. Folk respondents have been shown to be sensitive to specific features in language varieties (e.g., Graff, Labov & Harris 1986; Purnell, Idsardi & Baugh 1999), and in much of this work sophistication from the speech sciences and acoustic phonetics has played a large role, since phonological features have been the most frequently investigated.

Plichta & Preston (2005) selected a well-known southern US speech stereotype (/ay/ monophthongization) and resynthesized a sample of the word *guide* so that it increased in monophthongization in seven regular steps from a fully diphthongal form ([ɑɪ]) to a fully monophthongal one ([ɑː]).
The seven voice samples (one male and one female) were played three times at each of the seven steps for a total of forty-two judgments. In each case, the respondent was to assign the word to one of the nine sites shown in Figure 9. The sites were numbered one through nine (Saginaw to Dothan) so that a numeric base could be used to ascertain if degree of monophthongization was perceived by the respondents (from all over the US, N=96) as an increasingly southern feature. Table 2 shows the results.

An ANOVA post-hoc test shows that each of these mean scores is significantly different from every other one, revealing considerable sensitivity to very minor phonetic changes and very clearly showing an association between monophthongization and the respondents’ perception of its southern-ness.

At each of these seven steps, however, there was also a significant difference (based on independent t-tests) between the male and female voice scores (Figure 11).

<table>
<thead>
<tr>
<th>Step</th>
<th>Mean</th>
<th>Region</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.85</td>
<td>1 Saginaw</td>
</tr>
<tr>
<td>2</td>
<td>3.17</td>
<td>2 Coldwater</td>
</tr>
<tr>
<td>3</td>
<td>3.87</td>
<td>3 South Bend</td>
</tr>
<tr>
<td>4</td>
<td>4.89</td>
<td>4 Muncie</td>
</tr>
<tr>
<td>5</td>
<td>5.99</td>
<td>5 New Albany</td>
</tr>
<tr>
<td>6</td>
<td>6.58</td>
<td>6 Bowling Green</td>
</tr>
<tr>
<td>7</td>
<td>7.02</td>
<td>7 Nashville</td>
</tr>
<tr>
<td>8</td>
<td>7.02</td>
<td>8 Florence</td>
</tr>
<tr>
<td>9</td>
<td>7.02</td>
<td>9 Dothan</td>
</tr>
</tbody>
</table>

Table 2. Mean scores based on regional values assigned each step of the increasingly monophthongized versions of /a/ (Plichta & Preston 2005:121)

Figure 11. Assignment of seven-step monophthongized male and female samples of guide to the nine sites of Figure 9 (Plichta & Preston 2005:121)
The female voice, at the same degree of monophthongization as the male one, always has a lower (i.e., “more northern”) score. Women are regarded in sociolinguistic literature as more standard speakers (e.g., Trudgill 1972), but this finding tells us that this is also a folk stereotype, for respondents regarded the female voice as “more northern,” an areal fact caricaturistically regarded as “more standard” in US folk linguistics (e.g., Preston 1996 and above).

Perhaps what is most important about this survey is its complete separation from overt folk linguistic knowledge. Here are two respondent comments on the task:

I don’t do surveys ordinarily. … Unfortunately, after doing the sample. [sic] I realized that I am clueless. My ignorance could only compromise your findings. Had I a printed version of the word before it was sounded, I might have had more of a snowball’s chance.

I’ve just done the web linguistic survey and I kept going even though I was quite clueless. How should anyone other than travelling salesmen have any familiarity with these regional dialects? Who’s a ‘good respondent’ as opposed to one like me who’s mostly guessed in a tripartite way: north, middle, south? I had one stop in Indianapolis once during which I was so struck by a sixteen year [old] waitress’s accent that I’ve always remembered the difficulty of understanding her, but I sure haven’t remembered her voice.

I hope your survey produces something of value, but I’m curious how you might get there if other respondents are as clueless as I am! (Plichta & Preston 2005: 126)

Much phonetic work in folk perception is exactly like this: respondents cannot give a conscious account of the features involved, but experimental procedures show that they are very sensitive to them.

In some cases, rather than revealing folk sensitivity to variety, as the above does of the relationship between sex and standardness, experiments have also shown folk inability to make such distinctions. Niedzielski (1999) reports on forty-two southeastern Michigan respondents who were asked to listen to the recorded voice of a local speaker (whose Michigan identity was indicated); they were told to concentrate on the vowel they heard in particular words and to compare that vowel to a set of three resynthesized vowels (from the same speaker’s data). They were then asked to choose the one that best matched the original. The vowel space of the speaker is shown in Figure 12.

<table>
<thead>
<tr>
<th>F2 in Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>3000</td>
</tr>
<tr>
<td>2500</td>
</tr>
<tr>
<td>2000</td>
</tr>
<tr>
<td>1500</td>
</tr>
<tr>
<td>1000</td>
</tr>
<tr>
<td>500</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>F1 in Hz</th>
</tr>
</thead>
<tbody>
<tr>
<td>300</td>
</tr>
<tr>
<td>400</td>
</tr>
<tr>
<td>500</td>
</tr>
<tr>
<td>600</td>
</tr>
<tr>
<td>700</td>
</tr>
<tr>
<td>800</td>
</tr>
</tbody>
</table>

Figure 12. Vowel space of the Detroit female speaker on the test tape (modified from Niedzielski 1999: 65)
This speaker is influenced by a vowel rotation in the Inland North of the US known as the *Northern Cities Shift*: the F1 for her /æ/) is at about 700 Hz; the norm for female speakers of American English (according to Peterson & Barney 1952: 183) should be considerably lower, around 860 Hz. Her /a/ is also fronted to F2 1775 Hz, while the Peterson & Barney norm is 1220 Hz. The raising of /æ/ and fronting of /a/ are usually considered the first two steps of the shift (Labov 1994: 184).

Niedzielski examined the respondents’ classification of the /æ/-word ‘last.’ The formant frequencies for the three resynthesized tokens that the respondents were given to choose from in the matching task are shown in Table 3.

Table 3. Formant values of tokens offered to respondents to match with the vowel in the speaker’s pronunciation of ‘last’ (Niedzielski 1999: 74)

<table>
<thead>
<tr>
<th>Token #</th>
<th>F1</th>
<th>F2</th>
<th>label</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>900</td>
<td>1530</td>
<td>hyper-standard</td>
</tr>
<tr>
<td>2</td>
<td>775</td>
<td>1700</td>
<td>canonical</td>
</tr>
<tr>
<td>3</td>
<td>700</td>
<td>1900</td>
<td>actual token (see Figure 11)</td>
</tr>
</tbody>
</table>

The results of this matching experiment are shown in Table 4.

Table 4. Respondent matching results for the vowel in ‘last’ (adapted from Niedzielski 1999: 72)

<table>
<thead>
<tr>
<th>token</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>/æ/</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>token</td>
<td>10%</td>
<td>90%</td>
<td>0%</td>
<td>42</td>
</tr>
<tr>
<td>n=</td>
<td>4</td>
<td>38</td>
<td>0</td>
<td>42</td>
</tr>
</tbody>
</table>

Not one of the respondents chose token #3, the variant that matched the one first produced by the speaker. Instead, they overwhelmingly chose the lower, more central token, #2. A few respondents even chose the hyper-standard token.

This work shows a considerable mismatch between perception and acoustic reality. The respondents reported that they heard a fellow Michigan speaker (importantly identified as one) use the canonical (or ‘pre-shift’) forms of vowels rather than the shifted ones actually used. Why are these respondents so inaccurate in this task?

When these respondents are presented with data from a speaker who they think is a fellow Michigander, the stereotype of Michigan English as standard emerges (see, for example, Preston 1996 and the status-evaluation of Michigan speech by locals reported above). As a result of this folk stereotype, the respondent selects the ‘standard’ vowel in the task. It appears that the linguistically secure can alter their perceptual and even production systems somewhat easily since they cannot conceive that their own performance (or that of others like them) would stray from a standard (i.e., their norms). Michiganders are so linguistically secure that they even seem to acoustically recalibrate the vowels of those around them and avoid notice of change or difference.

Another technique in folk linguistics, and one underutilized in my opinion, has to do with language variety imitations and their reception. Evans (2010b), for example, shows how a speaker of northern US English rather accurately imitates some of the detailed vocalic features of a more southern variety in a task in which he was instructed to simply read a short text in that variety.
When his performance was played along with that of three authentic locals (and several non-local filler voices), he tied one of the locals and bested the other two in ratings of local identity elicited from local speakers. This folk imitative proficiency is rather surprising, for the usual linguistic assessment of such ability is low:

Although one can achieve a certain amount of insight working with bilingual informants, it is doubtful if as much can be said for ‘bidialectal’ informants, if indeed such speakers exist. We have not encountered any nonstandard speakers who gained good control of a standard language, and still retained control of the nonstandard vernacular. Dialect differences depend upon low-level rules which appear as minor adjustments and extensions of contextual conditions, etc. It appears that such conditions inevitably interact, and although the speaker may indeed appear to be speaking the vernacular, close examination of his speech shows that his grammar has been heavily influenced by the standard. He may succeed in convincing his listeners that he is speaking the vernacular, but this impression seems to depend upon a number of unsystematic and heavily marked signals. (Labov 1972: 215)

Purschke (2010) asked traditional Hessian speakers to suppress Hessian elements in their speech and asked non-Hessian speakers to imitate Hessian. Although the Hessians did not succeed in sounding non-Hessian, their imitations were of what Purschke calls “New Hessian,” the Frankfurt area variety that has many distinctive Hessian features but, from the point of view of older vernacular speakers, might represent a non-Hessian or more widely distributed variety. That was exactly the variety that non-Hessians had access to, and their imitations of Hessian were very much like what the authentic Hessians did when asked to speak non-Hessian (164). When asked to identify the locality of the speakers, however, Hessian respondents spotted the imitators nearly perfectly while non-Hessian respondents (from Schleswig-Holstein) could not distinguish between the two (168).

Many more interesting folk linguistic experiments might be listed here, but applied linguists will surely see the relevance of such findings as those reported above in their various enterprises. I will not dwell on the importance of imitative ability (and the success or failure of imitations in convincing local speakers of their authenticity); that is too obvious for workers in any area of variety instruction in the native language and all areas of foreign and second language instruction. Of equal importance to practitioners, however, are surely those experimental findings that show how folk belief may interfere with the abilities of speakers to produce and process language and language variety. If Niedzielski’s Michiganders do not hear their own vowels, how can we work with them on pronunciation, no matter what our goal in that area might be? In an obviously applied study, Rubin (1992) showed that undergraduates rated a Midwestern US speech sample as “more accented” when they were led to believe that the speaker was an Asian, and they actually comprehended less of the same speech sample when they were so misled. Put crudely, this is “I don’t understand you because you look like somebody I can’t understand.” This folk linguistic finding should be required reading for any whose business it is to make sure that non-native speakers can be understood. Maybe, as Rubin concludes, work should be done with the listeners as well.

4. Discourse approaches

Not so many years ago at a meeting of the American Association for Applied Linguistics, Emanuel Schegloff announced in his plenary address that “conversation analysis is applied linguistics.” That seems to me to overstate the scope of conversation analysis and underestimate that of applied linguistics, but the discourse bug has already bitten many who have applied interests. Folk linguistic discourses, although not necessarily identified as such, have been analyzed in, for example, Kalaja & Barcelos (2003), who provide a number of different discourse-centered approaches to beliefs about
second language acquisition. The authors use interpretive strategies that, for the most part, can be related to the unpacking of folk belief from the actual words of respondents. That is, they are interpreted within the broader framework of the situation of speech and even of the surrounding culture, much as outlined in the section above on traditional approaches. In more general folk approaches to conversation, Jara Murillo (2006), for example, applies quantitative techniques to the topical themes drawn from conversations with Costa Ricans about their own language.

Here I will offer more linguistically oriented approaches to the metalinguistic content of talk about talk, i.e., attempts to derive from the structure of the discourse some clues to its content. Linguists are well-equipped to look beyond what is said and to uncover what is presupposed, and such presuppositions often involve deeply-held folk beliefs. Preston (1994) reviews a number of strategies for the analysis of folk linguistic discourses that might reveal subconscious attitudes — topic selection in imitation, referential specificity in argument, speaker-hearer footing, discourse markers, and topic perspectives.

Since I cannot deal with all of these, let me illustrate this potential for revealing the subconscious in discourse by pointing out the possibility of extracting pragmatic presuppositions, which are related to lexical and structural triggers (e.g., Levinson 1983: 181–85). For example, 'started' in 'Bill started smoking' presupposes that there was a time in the past when Bill did not smoke (e.g., Levinson 1983: 182). Although 'Bill didn't flunk Algebra' doesn’t presuppose that Bill flunked anything, 'What Bill didn’t flunk was Algebra' suggests he flunked something (e.g., Levinson 1983: 182–83). When discourses turn to language (rather than smoking and flunking), such presuppositions prove even more interesting.

In the following exchange, a Taiwanese fieldworker (C) discusses African American English with an African American friend (D).

1 C: We uh — linguistics, in this field, uh — from the book I s- I mean, I saw from the book that — many linguists quite interest in black English. So could you tell me — a little bit about — your dialect?
2 D: Dialects.
3 C: Heh yeah
4 All: ((laugh))

5 D: Well, uh: — well — see the world’s getting smaller. There’s=
6 C: ((laughs)) I- I mea- do you have-
7 D: =not — even among all the ethnic groups we’re- we’re getting- getting less and less of dialectual in- inFLUence. (.hhh) Uh I’m- happen — not to be — from the South, ….. (Preston 1994: 286–87)

Without an account of presuppositions, I think this discourse is difficult to interpret, particularly the content of 5–7 D. The first key is in the presupposition(s) of “So could you tell me a little bit about your dialect” (1 C). “Your dialect” presupposes the existence of “dialect(s)” and that “you” are the speaker of one. D’s perception of C’s presuppositions leads to the otherwise difficult to understand assertions in 5–7 D:

The world’s getting smaller.
We’re getting less and less of dialectual influence (i.e., there are fewer and fewer dialects)
I happen not to be from the South.
"The world's getting smaller" is an explanation of why there are fewer dialects (education, media, mobility, etc…), but D's next assertion, that there are fewer dialects, responds to C's presupposition that there are such things (a definite description; e.g., Levinson 1983: 181). Finally, and more subtly, D confirms C's presupposition that dialects exist, but, for him (D), they exist only in such places as "the South." In other words, if C had been lucky enough to encounter a speaker from the South, he might have had his request for information about "your dialect" fulfilled.

How can we make sense of D's observation that he is not from the South unless it is in some way related to his response to C's request for information about D's dialect (and embedded in D's assertion that there are fewer dialects)? Recall that Michiganders, D included, find the South very salient as a regional speech area and that its salience is undoubtedly related to its incorrectness (see Figure 8 and Table 1); i.e., it is "a dialect."

Presuppositional work, however, may also explain why D "happens" not to be from the South. Why does he not simply say 'I am not from the South'? 'Happen' belongs to a group of implicative verbs (Levinson 1983: 181) and presupposes 'inadvertence,' 'lack of planning,' or 'by chance.' D “happens” not to be from the South because it is only a case of bad luck that C picked on a respondent who was not from the South (and could therefore not fulfill his request for "dialect" information).

Such work as this leads us to the building of a cultural model of D's (as well as C's) language ideologies, and I provide a great deal more on this conversation and various pragmatic approaches to its content in Preston (1994). Work on discourse, then, from many perspectives, but surely from both formal and informal pragmatic ones, reveals not only what speakers have said or asserted (the conscious) but also what they have associated, entailed, and presupposed (the subconscious).

There is no doubt that the subconscious is important in folk linguistics. I have quoted Irvine (above) on the importance of seeing language in action as well as talk about language; that the two may be in stark contrast is true. Kristiansen (2009), for example, finds that Danes from all over Denmark say that they like their home variety best, but that, when a carefully constructed matched-guise test is given, they all seem to prefer the emerging 'New Copenhagen' standard, a form that seems to be sweeping the country. A discourse model that looked only at what was said would find only one interesting folk linguistic fact about modern Danish varieties (and serve Danish applied linguists very badly). I encourage you to talk to people about talk and analyze what they have said and what they have meant.

Some readers may be unhappy to see their favorite approaches to discourse (interactive analyses, critical analyses, and the like) not highlighted here. I have not done so because 1) they already have wide representation and use in applied linguistics, and 2) I have tried to make more of relatively more formal linguistic approaches to discourse structure and meaning, areas I feel are underrepresented in applied work.

5. Conclusion
This methodological survey of data collection techniques and interpretive strategies in folk linguistics is necessarily broad and is meant to provide only an overview of research to date. It is not possible here to detail the many specific applications of such work for each reader's particular area of applied linguistics. The articles in this AILA Review, however, are all good examples of folk linguistic studies relevant to applied linguistic concerns, and I hope they and this survey of methods will inspire you to do some research of your own. Whether you improve on or replicate the methods presented here, I hope to have convinced you of the importance of knowing the beliefs about language and its varieties among the people you aspire to serve.
References


Author’s address

Dennis R. Preston
Oklahoma State University
English Department
205 Morrill Hall
Stillwater, OK 74078
USA

dennis.preston@okstate.edu