Language and Political Behavior: A Methodology for Utilizing the Linguistic Component of Socioeconomic Status

Mitchell A. Seligson; Susan Berk-Seligson

THE WORKSHOP

Language and Political Behavior: A Methodology for Utilizing the Linguistic Component of Socioeconomic Status*

Mitchell A. Seligson, Political Science
Susan Berk-Seligson, Linguistics, *University of Arizona*

This paper attempts to show the relevance of sociolinguistic methodology for social and political research. In particular, it reveals that, at least for the data analyzed, a sociolinguistic measure of socioeconomic status (SES) is a more powerful predictor of political behavior than are the traditional measures of SES (e.g., income and education). The paper explicates in detail the three stages necessary for carrying out this kind of research, examining data collected in Costa Rica. The findings strongly encourage the incorporation of sociolinguistic measures into survey research.

> *The rain in Spain stays mainly on the plain.*
> —Eliza Doolittle to Henry Higgins

The fact that socioeconomic status (SES) is componential in nature has been known by the social sciences for some time now. This knowledge, together with the desire to arrive at measures which have more and more predictive power, has led researchers to channel their efforts toward uncovering the various components which comprise SES. Sociolinguistics has made an important contribution to our understanding of SES by revealing one additional facet of it: language.[1] In particular, sociolinguistic studies

[1] It should be noted that the relationship between language and social status was pointed out before sociolinguistics ever became an established area of inquiry in its own right. In particular, anthropologists (Bright and Ramanujan, 1962; Ferguson, 1959; Gumperz, 1958; Gumperz, 1964) were forerunners in noting linkages between linguistic variation and social variation.

have shown that the choice of a linguistic variant, be it phonological or grammatical, depends, among other things,[2] upon the social status of the speaker (Anshen, 1969; Bernstein, 1971; Fishman, Cooper, Ma, et al., 1971; Fisher, 1958; Labov, 1966, 1972; Levine and Crockett, 1966; Sankoff and Cedergren, 1971; Shuy, 1967; Shuy, Wolfram and Riley, 1968; Trudgill, 1972; Wolfram, 1969). In other words, language variables correlate with a speaker's SES. Thus, whenever there is a choice of a "prestige" versus a "stigmatized" variant of a linguistic form (e.g., "am not" versus "ain't"; "doing" versus "doin' "), in general the higher the SES of the speaker, the lower the frequency with which he or she will use stigmatized variants.

With the finding that linguistic choice varies with SES, contemporary definitions of SES can be broadened to include the linguistic component, the expectation being that a greater percentage of variance in SES-related variables will be explained by including measures of linguistic performance. We are therefore suggesting that speech be added to the more traditional measures of SES (i.e., income, occupation, education, etc.).

Sociolinguistic insights, unfortunately, have not been widely appreciated by the social sciences as a whole, despite the fact that some scholars (e.g. Hymes, 1970) have been encouraging their broader application; thus, the findings of sociolinguistics have remained within the rather limited domain of researchers who are directly interested in language and its contextual framework. Why sociolinguists' observation that language varies with SES has been ignored by other social sciences is understandable. Their hesitancy in this regard probably stems from: (1) a general feeling of inadequacy toward eliciting and analyzing linguistic material, emanating from an awareness that linguistic fieldwork requires special prior training; (2) a concern for the expenses that would be incurred by the purchase of sophisticated audio equipment (e.g., highly sensitive recorders, microphones, etc.); and (3) the fact that in comparative political research language usually has been viewed as a barrier to research (e.g., problems of translation equivalence [Iyengar, 1976]) rather than as a facilitator of it. The present study shows how these obstacles can be overcome by modifying and simplifying the field procedures used by sociolinguists.

Objectives

The objective of this study is to demonstrate that the insights gained in sociolinguistic research can be beneficially applied to political science

_____

[2] Language also varies with the level of contextual formality in which the utterance is spoken, as is discussed below, and with the sex of the speaker.

research, particularly to survey research which utilizes SES as a predictor of attitudes or behaviors. We attempt to do this by taking the reader through a three-stage process of data collection and analysis. We have chosen to illustrate our methodology with foreign data gathered in a developing area (Costa Rica) rather than with data collected in the United States or Western Europe. We wish to present the reader with an illustration of a complex research situation so that he/she may become familiar with as many of the stages as may be required for conducting this type of research. However, researchers studying areas which have already undergone considerable sociolinguistic research will not have to repeat all the stages presented here.

The goal of this three-stage process is to develop an index of linguistic performance which can then be used as a predictor variable in political research. We use this index as a predictor in regression analysis in which four modes of political participation are the dependent variables. We find that in most cases the linguistic index is a better predictor of political participation than are the standard measures of SES. We believe that the strength of the linguistic index lies in its unobtrusive nature; whereas respondents can "fudge" their answers to questions on income, education, and the like, sociolinguistic performance is much less susceptible to such manipulation.

Method

Stage One: Selecting the Variables and Formality Level

The first and perhaps the most important step to be taken in devising an index of sociolinguistic performance is selecting the linguistic variables and the level of formality at which they are to be elicited, both of which concerns will be discussed in this section of the paper. The proper selection of the variables is of critical importance; if the items do not vary socially, then they will not have the power to say anything about the SES of the respondents. Undoubtedly, collaboration with a linguist would be necessary for the political scientist at this stage of research, since it would be the role of the linguist to choose the variables, create ways of eliciting them from the respondents, and train fieldworkers to discriminate aurally between the variants of a given sound.[3] Before we discuss the field procedures we used

---

[3] Phonological variables (i.e., those dealing with the sound system of a language) rather than syntactic or morphological ones (i.e., grammatical variables on the sentence level and word level, respectively), are particularly suitable for sociolinguistic

in Costa Rica, we need to explore two important factors which induce variation in speech styles, namely level of formality and geographic variation.

Formality Levels and Speech Style. Extensive research has been conducted by William Labov (1966; 1972) on the relationship between speech style and the level of formality employed by a speaker. Labov has found that different speech styles vary according to the formality of the speech situation (e.g., a university lecture versus cocktail party converation versus a family quarrel), and that these styles reflect the degree of attention that a speaker pays to his speech. The hypothesis here is that the higher the level of contextual formality, the more attention will a speaker pay to his speech. Speech styles, it should be noted, are best thought of as "co-occurrence patterns" (Fishman, Cooper, Ma, et al. 1971), that is, patterns in which the level of formality of one sort of linguistic variable (e.g., a grammatical one) is in consonance with the level of formality of other sorts of variables (i.e., phonological or lexical)[4] which occur alongside it in the stream of speech. Thus, for example, in the context of a child arguing with his mother about having to clean up his room, the following sentence would be an appropriate one for a very casual speech style:

Aw, Mom, are ya gonna make me clean my room?
Well, I ain't doin' it!

Whereas, the following sentence, illustrative of a rather formal speech style, would sound unnatural in the situation for which it is being proposed:

Oh Mother, are you going to force me to tidy up my
room? Well, I am not doing it!

Co-occurrence pattern in the above example refers to the fact that "gonna" is consonant with the informality of "Mom" (as opposed to "Mother"), with "ain't" (as opposed to "am not"), and "doin'" (as opposed to "doing").

It is believed that an individual's heightened attention to, and consciousness of, speech accounts for the usage of formal speech: in a word, carefulness. The result of this greater degree of attention is the closer approximation of that speech to whatever is considered by the given speech community to be the "standard," or "prestige" dialectal variety. Con-

---

research because, in contrast to the latter, the former are high in frequency and fairly immune from conscious suppression (Labov, 1964, p. 176).

[4] Lexical variables are those having to do with the choice of a word for a given meaning.

versely, the more informal the speech context, the greater will be the use of the nonprestige, or "stigmatized" variety. The questionnaire we employ sets up several quasi-artificial speech situations, each of which is meant to be an analog to a real one.[5] Thus, following Labov, if we define five formality levels for our study, ranging from most formal to least formal (i.e., most casual), we can create their analogs in the form of the following speech tasks or contexts: (1) wordlist reading, (2) paragraph reading, (3) word naming, (4) speech produced during the course of conversation between interviewer and respondent, and (5) excited emotional speech, produced during the relating of some critical incident that occurred in the life of the respondent (e.g., the time he narrowly escaped serious injury, or thought he was close to death). Reading a list of words, which is representative of the more formal speech situation, will make the speaker quite careful of and attentive to his pronunciation of the words on that list, whereas reading the same words when they are embedded within a paragraph will draw the speaker's attention away from that particular set of words. Naming words (e.g., "What do you call the things that hens lay?") will cause the speaker to pay even less attention to his pronunciation of them, in his effort to "get the right answer" in such a fill-in-the-blank exercise. Similarly, in a more open, freer context (e.g., ordinary conversation with the interviewer), a speaker will be even less conscious of his pronunciation, and under a situation of excitement or emotion, the attention he pays to the pronunciation of any given word should decline even further.

Studies of this type have shown that the percentage of prestige (and, alternately, stigmatized) phonological variants varies with *both the formality level of the speech context as well as the SES of the speaker*. This two-dimensional variation means that although *all* speakers will "do their best" to pronounce words "correctly" on a wordlist reading task, there will be a variation in the percentage of items pronounced in the prestige form, and this variation will be largely a function of SES. Thus, the higher the SES of the speaker, the higher the percentage of words which he will pronounce in the nonstigmatized manner. Of course, the very sight of a selected list of words on a printed page should make for little variance. Nevertheless, variation does indeed exist. Some people, for example, are used to deleting, that is, omitting certain sounds in particular phonetic contexts (e.g., En-

---

[5] It is important to recognize that some argue that the interview context itself is not a completely natural speech situation, and that methods such as eavesdropping would gather data of a much more natural sort. Such an approach, however, besides being ethically reprehensible, is nowhere as systematic as a standardized interview format.

glish preconsonantal and word final r, as in "beard," "beer" are generally not pronounced by New Yorkers; neither is [k] in the frequently heard Costa Rican pronunciation of the consonant cluster [ks], as in *exacto* [eksakto] "exact"); the same people may continue to delete them even when seeing on a printed page words containing them. Thus, when contextual formality is held constant, variance in pronunciation is found largely to be accounted for by a speaker's SES.[6] What we need to do, then, is discover that level of formality which will produce the greatest speech variation without requiring a test situation too complex or too costly to include in a questionnaire whose focus is primarily political.

Geographic Variation and Speech Style. Of considerable influence on the choice of linguistic variants are geographical dialects and, to a lesser extent, argots and occupational jargons, all of whose simultaneous copresence with social dialects in a given geographical area will serve to complicate further the process of variable selection. Researchers who plan to study a sample that is spread out over relatively large expanses of territory should be careful not to include among their variables sounds which set off one geographical linguistic area from another, since this might introduce a possible confusion between socially and geographically determined linguistic variation. Fortunately, it is easy to avoid this kind of confusion

---

[6] Other factors (discussed further in the appendix) which may cause speech to vary and should also be held constant in the research situation are: (1) the particular phonological environment surrounding the given variable (i.e., the phonetic nature of the segments that precede or follow it, or whether it is contiguous to a word or syllable boundary; for example, we would normally pronounce the last sound of the word "London" as [n] if the word occurred at the end of our sentence, but we would be very likely to pronounce it as an [m] if we said "London Bridge" without pausing between the two words, because the [b] would affect the point of articulation of the preceding nasal sound); (2) the part of speech of the word in which the variable is embedded (e.g., in isolating languages such as English, where word order is relatively fixed, nouns, verbs, and adjectives may be less prone to deletion of segments than, for example, pronouns, since the former carry a heavier semantic load and hence convey more information than do the latter. Thus in American English, the [h] of "him" and "her" is often deleted, as in "Let 'im ('er) come in," but [h] is not as likely to be deleted in words such as "house," "hurry" or "heavy"; (3) the morphemic status of the variable (e.g., whether it is inflectional or not; for instance the change of [e] to [i] in Costa Rican Spanish is probably more frequent in a noninflectional ending, such as *calle* ("street"), than it is as a third person singular present subjunctive ending, such as *Que se calle* ("Let him be quiet"). G. Sankoff and Cedergren (1971) are among the early ones to observe that syntactic and phonological conditioning are coupled with stylistic level and SES in accounting for the occurrence of a given phonological variant.

since there are numerous regional dialect studies and "linguistic atlases," a number of them dividing the United States into major dialect areas (Atwood 1962; Babington and Atwood, 1961; Kurath, et al., 1939; Kurath, 1949; Kurath and McDavid Jr., 1961; McMillan, 1946; Pederson, 1965; Wood,, 1960, 1963) and others doing the same thing for European nations (Gilliéron and Mongin, 1905; Gilliéron and Roques, 1912; Jaberg and Jud, 1928–1940; Kirwin, 1968; Wagner, 1958; Wilson, 1958; Mather and Speitel, 1975).[7] Such atlases divide up geographical areas along the lines of dialect boundaries by using as variables a combination of phonological, grammatical, and lexical elements.[8]

Geographical atlases do more than merely map out geographical distribution. They provide descriptions of social variation within a language as well; therefore, studies to be conducted within a dialect area can use the atlases as a guide in selecting the sociolinguistic variables for study. Thus, for example, those working on *The Linguistic Atlas of the United States and Canada* have established three classes of informants ((1) old-fashioned, rustic, poorly-educated speakers; (2) younger, more modern, better-educated speakers; and (3) cultured, well-educated speakers) and take note of which classes of speakers use which linguistic forms.[9] In short,

---

[7] Although the aim of the following discussion is to demonstrate the utility of dialect atlases for social science research, it must be pointed out that traditional atlases do have several limitations, (e.g., biased sampling procedures). For a critique of traditional American linguistic geography, see Pickford (1956), and Underwood (1974).

[8] Thus, the English of New York City and the Hudson Valley is defined by a number of features, among them being the following ones (note that brackets [ ] represent sounds at the phonetic level of analysis and slashes / / indicate phonemic representation): there is no pronunciation distinction between "mourning": "morning," and "hoarse": "horse"; /—r/ is "lost" except before vowels; the vowels in words such as "pan," and "man," are raised and lengthened [æ .], as, too, is the vowel in words such as "lawn," "off" [.]; the use of prepositions results in sentences such as the following: "He lives *in* Fulton Street," "We stood *on* line"; and, lexically, "pot cheese" is used in place of "cottage cheese," and "Dominie" instead of "preacher." In the North of the United States a distinction is made between mourning:morning, hoarse:horse (i.e., between / / and / /, respectively), the first element in the diphthong of "fine" is centralized (phonetically [ɪ] or [ɪ], phonemically / y/); in grammar, "dove" is the preterit of "dive," and "clim" is that of "climb"; and lexically, "pit" is used for "see," and "johnnycake" for "cornbread." These examples are based on McDavid, Jr. (1958).

[9] Thus in the New York City and Hudson Valley area the variables mentioned in note 8 are considered to be characteristic of all three social classes. In the Northern area, on the other hand, whereas all the classes make the /o/ vs. / / distinction

then, for the social scientist interested in conducting survey research in a nonextensive geographical area, geographical atlases can be helpful in roughly delineating those variables which are seen to vary with social status.

Linguistic atlases are good starting points in the process of selecting and "de-selecting" variables among the entire range of possibilities. Unfortunately, such atlases have been devised primarily for North American and Western European nations, and are hard to find among Third World countries. What, then, does the social scientist who is about to embark upon research in Latin America, Africa, or Asia do? He/she can follow a series of steps, along the lines that we have taken in our Costa Rican study, steps which should guarantee that the linguistic variables which one finally employs will in fact vary with SES. By describing what we did and how we did it, we hope to make clear the procedure which should be followed in carrying out research in those developing parts of the world.

Field Procedures in Costa Rica. In order to begin to find out what characterizes lower-class speech in Costa Rica, we drew on three sources: (1) fictional literature which portrays lower-class types (Dobles, 1966; Echeverría, 1953; Fallas, 1963, 1967; Lyra, 1966); (2) the writings of linguists and philologists (Agüero, 1962, 1964; Arroyo Soto, 1971; Azofeifa, 1947; Salguero, 1967; Wilson, 1970) and (3) our conversations with primary school teachers, who are notorious for trying to eradicate from the speech of their pupils various types of stigmatized linguistic forms. Using these sources of information, then, we derived a list of linguistic forms which seemed to separate one group of speakers from another. It should be reemphasized that this stage of the procedure can be greatly simplified for those nations for which linguistic atlases are available.

The variables chosen for preliminary study may be classified under three major headings: accent shift (i.e., a shift in the placement of primary stress within a given word), vowel change (i.e., the substitution of a vowel for another vowel in a word), and consonantal change (i.e., the sub-

---

in pairs such as hoarse:horse, only class 1 speakers, and, to a lesser extent class 2 speakers, centralized the first element in the diphthong of "fine." Similarly, only classes 1 and 2 use "dove" and "clim," whereas all the social classes use the lexical variants mentioned in note 8.

stitution of a consonant for another consonant in a word). For a complete discussion of the variables see Berk-Seligson and Seligson (1978).[10] Of those variables, 12 were selected for use in the present study, primarily on the basis of ease of elicitation and the degree of variation they reflected across SES gradations. They are presented, along with concrete instances of them as they are embedded in vocabulary items, in the Appendix of this paper.

The principal purposes of this stage in our analysis are to determine, first, if the variables chosen are SES-related, and to discover which level of speech formality appears to be the most effective for further research. To achieve these objectives it is imperative that a socially heterogeneous sample be selected, for if there is little or no SES variation in the data, the test will be meaningless. The sample, however, need not be large, for if the linguistic variables chosen are appropriate ones, the relationships between the standard SES measures and language should emerge readily.

The variables listed above were elicited from a sample of 48 male respondents living in an area bordering on Costa Rica's national capital. The

---

[10] The following series of examples should illustrate the three categories. Each example is presented in the form of phonetic segments (along with accent marks where stress placement is the variable) flanking an arrow, the latter pointing toward the nonprestige variant, accompanied by lexical items, written first in Spanish orthography, then in phonetic spelling reflecting the nonprestige variant and, lastly, in the form of their English nearest equivalents. The lexical items are, then, specific manifestations of the variables.

Thus, included under the category of accent shift are the following: aú → áu (*ataúd* [atáud] "coffin"), aó → áo (*ahogo* [áogo] "suffocation"), aí → ái (*maíz* [máis] "corn"), aé → áe (*maestro* [máestro] "teacher"), éa → eá (*océano* [oseáno] "ocean"), eí → éi (*increíble* [iŋkréible] "incredible"), and oí → ói (*oído* [óido] "hearing or having heard"). Different types of vowel change were found to include o → e (*oscuro*, standard Costa Rican Spanish for *obscuro*, [eskuro] "dark"), o → u (*poeta* [pueta] "poet"), e → i (*peor* [pior] "worse"), and i → e (*policia* [polesia] "police", "policeman"). Consonantal change, perhaps the most widespread and certainly the most varied phenomenon of the three, is reflected not only in the alteration d → l, r → l, r → ∅ (i.e., phonetic "zero", or lack of any phonic substance), h → f, f → h (Costa Rican Spanish regularly employs [h] where the Spanish of most South American countries uses [x], a velar fricative), b, b → g, g, ∅ → g, and k → ∅ (examples of which are included in the Appendix of this paper and consequently are omitted here) but also in: (1) other alternations involving liquids (i.e., l and r), such as l → r *fácil* [fasir] "easy"), d → r (*admirar* [armirar] "admire"); (2) nasal alteration, such as m → n or ŋ (*himno* [inno] or [iŋno] "hymn"), g → n or ŋ (*ignorante* [innorante] or [iŋnorante] "ignorant"); and (3) other alternations involving stops, such as p → k (*aceptar* [asektar] "accept"), g → k (*gangrena* [kaŋgrena] "gangrene"), t → d (*Atlántico* [adlantiko] "Atlantic"), and others.

particular district, Sabanilla de Montes de Oca, was selected for sampling because of its uniquely varied demographic composition. In particular, Sabanilla is the residence of people with such diverse occupational background as peasants and university professors (the latter live in the area because of the proximity to one of the nation's two universities). Education and income levels are equally varied: from zero years of schooling (8 percent of the sample) to university education (4 percent), from subsistence salaries of 35 dollars per month (6 percent) to salaries of over 325 dollars per month (8 percent). Thus, Sabanilla provided a remarkably ideal testing ground for obtaining SES-varying data. It is because Sabanilla is a neighborhood in transition—from what was entirely a rural coffee-farming area less than a generation ago, to a suburban development today—that it houses the heterogeneous population that it does, and reflects the various strata of Costa Rican society (namely, peasants, blue collar workers, white collar workers, and professionals). Furthermore, sampling from a wider area in order to check for geographic variation was not necessary in our research since Costa Rica constitutes, with three minor exceptions, a geographically homogeneous speech community (Agüero, 1962, 1964). The areas where the exceptions occur (Puntarenas, Limón and Guanacaste) were not included in the subsequent stages of this research project.

Our preliminary linguistic study elicited linguistic variables at three levels of formality and confirmed that at each level of formality linguistic variation was correlated with standard measures of SES (as measured by income, occupation, education), various aspects of the house (e.g., facilities, such as piped-in water and electricity; type of flooring, door, windows, walls, toilet facility, number of bedrooms, etc.), and artifacts (e.g., various electrical appliances, type of wall decoration, type of furniture, etc.). The correlations accounted for an average of 25 percent of the variance.[11] Although significant multiple R's were found at the two most

[11] Linguistic scores for the respondents (the linguistic input into SES language correlations) were arrived at in the following manner. Prestige responses were scored 1, and nonprestige variants were given 0. The coding system for the two most casual formality levels, however, had to be different from that of the three more formal ones since the former two were elicited via open-ended procedures. The system decided on was one of tabulating from the tape recordings all the instances in which the variables under study occurred, using 100 percent of the corpus (i.e., data) produced on formality levels 1 and 2, a procedure which was feasible, given the fairly limited size of the corpus collected. On each informant's questionnaire, for the most casual contexts, the various phonological categories were listed, and the total number of mentions of lexical items containing the variables was tabulated, along with the dichotomous coding of "prestige" versus "stigmatized." Every person's score on each of levels 1

casual levels of formality, we decided not to elicit variables at these two levels in the subsequent stages of the research. This decision was a pragmatic one directly related to the complexities involved in administering and coding casual speech. The most formal level involved merely checking off of precoded responses to closed-ended questions, while the casual levels required transcribing open-ended speech. Of the three, word-naming proved to have the highest correlation with SES (it explained 26 percent of the variance, in contrast to 20 percent and 19 percent accounted for by the other two), and is, therefore, the most highly recommended among the three. An additional pragmatic reason for recommending its use as an elicitation procedure is that it can be used with illiterate respondents (a substantial portion of these respondents we planned to include in the political study in stage three were known to be illiterate), whereas other procedures (e.g. "word-list" reading) cannot.

The first stage in our analysis has demonstrated that the linguistic variables we have chosen are appropriate for proceeding to the next stages of research. Had our variables not demonstrated positive statistically significant correlations with standard SES measures, additional spade work would have had to be done. In addition, we have found that the linguistic component of SES is not so strongly related to the standard SES measures that it is simply another measure of them. Finally, we have learned that

---

and 2 was arrived at by dividing the number of stigmatized mentions by the total number of variables mentioned.

For over 90 percent of the variables there was more than one stigmatized variant per variable (e.g., *abuela*, has the alternate nonprestige forms [agwela] and [awela]; *maíz* may be nonprestige [máis] or [mays], the latter form being a more extreme case of the former in that not only has the stress shifted to another vowel, but that the now unstressed vowel has been reduced to the point of losing syllabicity, creating thereby a diphthong). Attempts had been made to rank the variants in order of distance from the prestige norm; however, such attempts proved to be unsuccessful. Whereas it had first occurred to us to do the ranking ourselves along the lines of linguistic criteria, we quickly dismissed the notion, realizing that our rank order would probably not coincide with that of a native speaker. With this rationale in mind, we had several Costa Ricans, philologists and linguists by profession, rank order the variants, placing them along the continuum of "most correct" to "least correct." Surprisingly there was no consensus on the orderings, so that the notion of weighting the responses had to be dismissed. What replaced it was the dichotomous breakdown described above. Any attempt on our part to give the variants weighted scores would have been arbitrary from the point of view of native intuition, and, apparently, achieving a valid rank ordering would have involved a study in itself.

the word-naming level of formality is an efficient, effective way to elicit linguistic data. We can now proceed to stage two of our analysis.

Stage Two: Testing the Salience of the Sociolinguistic Variables

Having shown that the series of linguistic variables that were chosen for pretesting are correlated with SES, it remains to be demonstrated that they are socially salient for Costa Ricans, that is to say, that they influence a listener's subjective view of a speaker who uses them in his/her speech. Tests constructed with the aim of procuring a respondent's reactions toward persons (including him/herself) using sociolinguistic variables are called "subjective reaction tests."

The test we used is a cross between the "matched guise technique" developed by Canadian psychologist Wallace Lambert and his associates (Lambert, 1967; Lambert, Anisfeld and Yeni-Komshiann, 1965 Lambert, Frankel and Tucker, 1966; Markel, Eisler and Reese, 1967; Tucker and Lambert, 1972), and a subjective reaction test devised by Labov (1966). It involves the use of a recording of a standard passage taped by three different people, each version varying from the next by only one feature: a different proportion of stigmatized to prestige forms. Subjects, after listening to the tapes twice (the speakers being place in random order) are asked to rate each of the speakers on a number of traits, each of which is listed on a sheet of paper in terms of a seven-point scale. The traits may be as diverse in nature as intelligence, wealth, cleanliness, honesty, and generosity. This rating format parallels the scheme set up by Osgood in his semantic differential technique. The listeners are then instructed to express their feelings about each speaker on four social distance questions, namely:

(1) I would befriend this person.
(2) I would accept this person as a relative by marriage.
(3) I would accept this person as a neighbor.
(4) I would help this person if he were in need.

In addition, listeners are to match the speakers with occupations which they consider to correspond best to them, and, finally, to state which of the three speakers they felt "spoke the best" and which one "spoke the worst," even though they are not provided with any definition of "speaking well" or "speaking badly."

We conducted a subjective reaction test of the sort described above in Costa Rica, using the same phonological variables that had been used in the linguistic study in stage one (Berk-Seligson, 1976). Several hypotheses were proposed. First, it was believed that listeners would be more positive toward speakers whose speech conformed phonologically to a prestigious

dialect, and, conversely, more negative toward those whose phonological variants were to a greater extent stigmatized. Positiveness and negativeness were operationalized as each of 13 pairs of polar-opposite traits (specifically: good/bad, friendly/unfriendly, strong/weak, hardworking/lazy, intelligent/dumb, trustworthy/untrustworthy, active/passive, honest/dishonest, worthy of esteem/worthy of scorn, generous/stingy, genteel/brusque, rich/poor, compliant/bossy). Secondly, it was hypothesized that in responding to social distance questions listeners would place themselves closest to those speakers whose linguistic variants were predominantly prestige rather than stigmatized. Thirdly, it was hypothesized that listeners would assign occupational categories to speakers in accordance with the proportion of stigmatized prestige forms in the latter's speech (i.e., that they would assign the highest status occupations to speakers whose phonological form more closely adhered to the prestige dialect, and progressively lower status occupations to speakers whose speech was more stigmatized). Fourthly, it was predicted that listeners would be able to judge accurately the degree of "correctness" (i.e., adherence to a prestige norm) of the speech of speakers whose use of prestige and stigmatizd phonological forms varied.

This type of subjective reaction test (in contrast to ones that ask respondents to provide perceptions of their own speech) can be conducted with large groups using paper and pencil questionnaires. Thus, with a good amplification system, the recording can be played in large halls or auditoriums, making the test a highly suitable one for student populations.[12] And, as is the case with most group-type paper and pencil questionnaires, research costs are greatly minimized, so that large samples can be obtained relatively inexpensively.

For the subjective reaction test carried out for our purposes, the sample ($N = 450$) consisted primarily of 9th grade high school students (56.9 percent) and 5th and 6th grade elementary school students (30.6 percent), the rest of the respondents (12.4 percent) consisting of 7th and 8th grade night school students and adults living in Sabanilla de Montes de

---

[12] It should be pointed out, however, that children are not very sensitive to sociolinguistic variation, and that awareness of socially significant linguistic variants does not set in until about the age of 15, and does not attain adultlike norms until age 16, or even as late as 18, as in the U.S. (Labov, 1964).

[13] The schools from which the sample was drawn were the Colegio Anastasio Alfaro, an all-girl public high school; the Liceo de Costa Rica, an all-boy public high school; the Escuela José Figueres, an elementary school; and Colegio Calasanz, a night school.

TABLE 1

Mean Score for Each Speaker by Adjective Pair

Adjective Pair	Prestige Speaker	Sig[b]	Intermediate Speaker	Sig[b]	Stigmatized Speaker
bueno/malo (good/bad)	2.07	<.001	3.04	ns	2.89
simpático/antipático (friendly/unfriendly)	2.98	<.001	3.78	ns	3.56
fuerte/débil (strong/weak)	3.56	<.001	4.11	ns	3.96
trabajador/vago (hard-working/lazy)	2.22	<.001	3.07	ns	2.87
inteligente/tonto (intelligent/stupid)	2.07	<.001	3.42	<.001	3.97
de confianza/desconfiable (trustworthy/untrustworthy)	2.94	<.001	3.55	ns	3.51
activo/pasivo (active/passive)	3.35	.011	3.72	.02	3.94
honrado/deshonesto (honest/dishonest)	2.09	<.001	2.82	.04	2.62
estimado/despreciable (worth of esteem/worthy of scorn)	2.52	<.001	3.24	ns	3.30
generoso/tacaño (generous/stingy)	2.73	<.001	3.53	<.001	3.10
gentil/brusco (genteel/brusque)	2.34	<.001	3.24	ns	3.26
rico/pobre (rich/poor)	3.91	<.001	4.46	<.001	5.25
suave/mandon (compliant/bossy)	3.01	.001	3.42	.05	3.20

[a] Note that the lower the mean, the more favorable the rating.
[b] These significance levels are t-tests between the indicated pairs of speakers. Hotelling's T² test performed on all three groups (test, retest) reveals that p < .01 in every case and is < .001 in 10 of the 12 adjective pairs.

TABLE 2

Mean Score Regarding Speakers
on Social Distance Scale[a]

	Prestige Speaker	Sig[b]	Intermediate Speaker	Sig[b]	Stigmatized Speaker
Accept as relative through marriage	1.38	<.001	1.55	.03	1.60
Accept as friend	1.10	<.001	1.22	ns	1.25
Accept as neighbor	1.01	<.001	1.19	ns	1.23
Help if in need	1.06	.03	1.10	ns	1.09

[a] Note that responses were coded: 1 = yes, 2 = no. Thus, the lower the score, the more positive were the listeners toward a given speaker; conversely, the higher the score, the more negative were the listeners.

[b] Significance levels are t-tests. Hotelling's T^2 test performed on all three speakers reveals that for the first three social distance significance is < .001, but that for the "help if in need" measure no overall statistically significant results are obtained.

Oca.[13] Thus, the ages represented in the sample ranged from 10 years to 47 years, 72 percent of the respondents being 13 years old or older.

The findings of this study demonstrated that the hypotheses were borne out; however, they also showed that, across the board, listeners could distinguish sharply between the "prestige speaker"[14] on the one hand, and the "intermediate" and "stigmatized" speaker on the other, but not very distinctly (and sometimes not at all) between the intermediate and the stigmatized speaker in relation to each other. A look at Tables 1 and 2 makes this conclusion clear.

Table 1 reveals that the prestige speaker is considered to be significantly better, friendlier, stronger, more hardworking, more intelligent, more trustworthy, more active, more honest, more worthy of esteem, more generous, more genteel, wealthier and more compliant than either of the other two speakers. However, in comparing the ratings for the intermediate speaker and the stigmatized speaker, on only 5 of our 13 pairs of traits are the ratings in the predicted direction (*intelligent/dumb, active/passive,* worthy of *esteem/*worthy of *scorn, genteel/brusque,* and *rich/poor*), and even among the ones that are as predicted, only 2 out of the 5 are statistically significant (.05 or better). Nevertheless, despite the overall trend toward not distinguishing between those speakers, listeners made clear distinctions between them when it came to traits that in some way were associated with SES (i.e., *rich/poor, intelligent/dumb*[15]), demonstrating in this way that they did in fact view them as socioeconomically differentiated, but apparently not differentiated in personality characteristics.

In a parallel fashion, Table 2 demonstrates that for each category of social distance, listeners placed themselves closest (i.e., scores approaching 1 rather than 2) to the prestige speaker, farthest away from the stigmatized speaker, and somewhere in-between in relation to the intermediate speaker. However, as in the case of the adjective pairs, listeners placed a much greater distance between the prestige speaker and the intermediate speaker,

---

[14] The "prestige speaker" is the one who used only prestige variants; the "stigmatized speaker" used only stigmatized variants; and the "intermediate speaker" used an equal proportion of both.

[15] Although, on the surface, intelligence would appear not to be a status-linked trait, in practice Costa Ricans usually associate degrees of intelligence with corresponding educational levels (i.e., if a person has had many years of schooling he is assumed to be very intelligent; conversely, if he has had very little schooling, he is assumed to be unintelligent). Thus, even though intelligence and education are by no means necessarily associated, popular opinion in Costa Rica links the two, particularly in the case of judging people at the lower and upper extremes of educational attainment.

on the one hand, than between the intermediate speaker and the stigma-
tized speaker on the other. Thus, with the exception of the "help if in
need" category, the differences between the mean scores for columns one
and two (.17, .12, and .18) of Table 2 are far greater than they are for
columns two and three (.04, .03, and .04), looking from top to bottom.
Why the "help if in need" category is an exception to the pattern is not
self-evident. A possibility is that there may be a special quality inherent in
it which neutralizes whatever social differences may be perceived in regard
to the speakers, and that this quality is some sort of humaneness. Thus,
whether the speaker is an upper SES type, a lower SES type, or someone
somewhere in-between, listeners in Costa Rica seem to be nearly equally
willing to help him if he were in need. Perhaps similar findings would ap-
pear cross-culturally, further indicating that this question might be a poor
measure of social distance.

Why the listeners did not discriminate as sharply between the inter-
mediate and the stigmatized speaker as they did between the prestige
speaker and the latter two is not clear. We suspect, however, that what
may be responsible for this undifferentiation is a threshold phenomenon: a
threshold of "prestigeness" which has not been reached. In other words,
the proportion of 50 percent stigmatized forms to 50 percent prestige
forms may not be sufficiently weighted toward the prestige side for listeners
to set apart the two speakers who use stigmatized forms. The proper ratio
would have to be arrived at empirically, for there is no way of knowing if
it should be 60:40 or 75:25, or something even higher. An alternative ex-
planation rests not on quantitative grounds, but on a qualitative notion;
namely, that the inclusion among the variables of some that are stereotypi-
cal of lower-class speech (on the order of "toidy-toid street" in Brooklyn-
ese) in the recordings made by both the intermediate and the stigmatized
speaker may have heavily influenced the listeners toward perceiving the
former to be very similar to the latter, despite the fact that one used twice
as many stigmatized variants as the other.[16]

We found that listeners were able to match up the speakers with cor-
responding occupations (specifically, the listeners were asked: Which of
the three men is a professor? Which one is a laborer? Which one is a

---

[16] Another explanation which might account for the listeners' inability to dis-
criminate sharply between the two nonprestige speakers is what is known as "para-
language," that is, features of speech which are not part of the sound segments them-
selves, but which overlay them (e.g., intonational features such as pitch, stress and
was in, and that we were interested in knowing which words were the ones
oranges (*naranjas*) are called *chinas* and buses (*cazadoras*) are called

mason?), as the following figures demonstrate. The highest status occupation, professor, was assigned to the prestige speaker by 68.1 percent of those who responded. The intermediate occupational rank, that of mason, was correctly matched up with the intermediate speaker by 66.7 percent of the respondents. And 62.3 percent matched the occupation of laborer with the stigmatized speaker.

The fourth hypothesis, that listeners would be able to judge the "correctness" of a speaker's speech according to the degree to which he used prestige and stigmatized forms, is borne out as well. To the question, "Which of the speakers spoke best?" 77.1 percent of those who responded did so correctly (i.e., chose the prestige speaker), and in reply to the question, "Which of the speakers spoke worst?" 66.3 percent were accurate (i.e., chose the stigmatized speaker). Furthermore, looking at these percentages and those cited in support of the third hypothesis, we see that the prestige speaker is more often correctly identified (in terms of occupation and "best" speech) than are the other two speakers. This again is consistent with the findings uncovered in reference to the first two hypotheses, namely, that the prestige speaker stands out for the listeners, and is set apart from the other two.

Viewing the findings of the subjective reaction test in their totality, we may conclude that they demonstrate the social salience of our original set of linguistic variables, and that our decision to use them in the subsequent political study was a well-founded one.

Stage Three: Inclusion of Sociolinguistic Variables in Political Surveys

The final stage of the research is the one most directly related to political research. It is here that we finally are able to determine the utility of gathering sociolinguistic data in a political survey. Our central concerns in this phase were to design a series of questions which would: (1) be easy to administer by nonlinguists, (2) absorb as little interviewing time as possible, and (3) be as unobtrusive as possible. The word-naming level of formality had previously been shown to be a particularly promising one because of its high correlations with the standard SES measures and because of its ease of administration (see our discussion of stage one).

The problem of eliciting the variables in an unobtrusive manner was solved by first informing the respondent that a given object may be called by different names according to the particular Latin American country one used in Costa Rica. For example, the interviewer stated that in Puerto Rico juncture, relative loudness of speech, vocal qualities, such as huskiness, breathiness, etc.)

guagas. This procedure resulted in a very unobtrusive method of eliciting the desired lexical items, since the respondent, quite happy to tell an outsider the names of commonplace objects and actions, would concentrate on providing the proper word without focusing on its pronunciation. This general procedure served an ancillary function, beyond providing linguistic data: it was a good way of establishing rapport with the respondent early in the interview because the respondents saw the exercise as a kind of game. Pretesting of the elicitation questions is of course necessary. It was discovered during this pretesting stage that the questions had to be rephrased in such a way as to ensure the occurrence of the word being elicited. For instance, to the originally phrased question. "What do you call your mother's mother?" many of the people interviewed during the pretesting replied "María" or whatever their grandmother's name happened to be. The question had to be reworded to "What does *one* call *one's* mother's mother?" in order to ensure the eliciting of "grandmother." Another problem occurred with the question, "Name some animals that live in the jungle and roar a lot." Many peasants at first did not think of lions; instead they provided a list of animals whose Spanish names were not even known to us (e.g. tapir and ocelot). However, after adding the probe, "Animals with a large mane," the word "lion" emerged.

In a study where differences in pronunciation are a target of attention, it is necessary for the interviewers to be given linguistic training in order that they be able to discriminate between what are often initially difficult to perceive differences. Here again the assistance of a linguist is indispensible. Minimal pair drills (drills using pairs of words which are identical except for one sound, e.g., 'bet' vs. 'pet') are a good exercise for this kind of linguistic training. It might appear to some that achieving an adequate degree of linguistic sensitivity through a short training period would be an unlikely expectation. However, this study provides evidence to support our contention, for despite the fact that the three interviewers involved in the study had markedly varying degrees of linguistic training (one was a linguist, another was a political scientist trained by the latter, and the third was a native Spanish speaker with two hours of formal training and five hours of supervised field experience), there was no statistically significant (t-test) difference between the mean responses obtained by the three. Of course, the training procedure was greatly simplified by the fact that the study focused on only 12 words and for this reason it was possible to achieve high competence in a relatively short training period. Bearing this in mind, we may now turn to a discussion of the data.

The Data

During the course of the latter part of 1972 and the first half of 1973 a probability sample of 531 peasants was interviewed as part of a study designed to tap the political attitudes of Costa Rican peasants (Seligson, forthcoming-a).[17] The linguistic portion of the questionnaire used (see Table 3) was designed to elicit 12 words: *huevo* (egg), *abuela* (grandmother), *afuera* (outside), *jornalero* (day worker), *vomitar* (to vomit), *ataúd* (coffin), *maíz* (corn), *león* (lion), *oscuro* (dark), *inyección* (injection), *matarlo* (to kill it/him), and *problemas* (problems). These items were treated as a test to see for how many of the 12 each respondent could answer "correctly" (i.e., provide the prestige variant).[18] The respondent received a score of +1 for each correct response. The index of a given respondent was computed by summing up the total number of his correct nonmissing responses and dividing by his total number of nonmissing responses.[19] The resulting index, which ranged from 0 percent to 100 percent had an overall mean of 69 percent, with .6 percent of the sample at the low end, 5.5 percent at the high end, and a standard deviation of .19. The distribution of the responses was a normal one (skewness = —.6; kurtosis = —.16).

Findings

In this section we examine the predictive power of our sociolinguistic index by comparing it to some measures of SES that are more commonly used to predict political behavior. If the sociolinguistic index in fact measures a unique component of SES, then we should find that it has some independent predictive power of its own when it is entered into a multiple regression equation. Furthermore, if the linguistic component of SES is not to be considered a trivial one, then in some cases language should be at

[17] Complete details of the sample design are contained in Seligson (1974, pp. 248–67 and 1977a).

[18] The particular stigmatized items under consideration in this study are by no means peculiar to Costa Rican Spanish. Most of them reflect phonological processes that Spanish is undergoing and that are attested in other Latin American countries. See, for example, Menéndez Pidal (1966), Canfield (1962), and Lapesa (1965).

[19] While the items are treated here as an index rather than as a scale, the mean inter-item correlation (tetrachoric r) is .33 and Cronbach's alpha is .71. Missing data were encountered in 10 percent of the cases, 8.9 percent having one sociolinguistic item of the twelve reported as missing and 1.1 percent of the cases with two missing items. No respondent had more than two missing items.

TABLE 3

Sociolinguistic Variables

Spanish orthography	English gloss	Pronunciation of Prestige form	%	Pronunciation of Stigmatized form	%	Missing
inyección	injection	[inyeksyón]	90.6	[inyesyón]	9.2	.2
maíz	corn	[maís]	88.3	[maís or mays]	11.5	.2
huevo	egg	[webo]	86.6	[gwebo]	12.4	.9
afuera	outside	[afwera]	85.3	[ahwera]	13.7	.9
oscuro	dark	[oskuro]	84.7	[eskuro]	13.7	1.5
matarlo	to kill it	[matarlo]	77.6	[matalo or matallo]	21.3	1.1
vomitar	to vomit	[bomitar]	76.1	[gomitar]	24.5	.0
problemas	problems	[problemas]	75.5	[poblemas]	20.7	.2
abuela	grandmother	[abwela]	55.0	[agwela]	44.6	.4
león	lion	[león]	51.2	[lión]	48.2	.6
jornalero	day-worker	[hornalero]	32.6	[fornalero]	66.9	.6
ataúd	coffin	[ataúd]	17.9	[ataúl]	81.7	.4

least as strong a predictor of behavior as some of the other standard measures.

We have selected a "worst case" data set for our test. That is to say, we have set up a test under the most difficult conditions, hoping that if our sociolinguistic index proves effective in this instance it should be of considerable utility under more favorable circumstances. The data set is of the "worst case" type in two ways. First, the sample represents a narrow segment of the range of SES variation in Costa Rica. Hence, we are working with a truncated distribution in which SES variation is severely restricted. Second, the political variables chosen for analysis have been found in cross-national studies to have low (although consistently positive) correlations with standard SES measures.

Our procedure is to compare the beta weights of the linguistic index along with six other commonly used measures of SES in four separate regression equations. The SES predictors used are: (1) average weekly income, (2) highest number of years of school completed, (3) presence or absence of electricity in the residence, (4) type of bed respondent sleeps

on (none, wooden slat, manufactured cot), (5) type of floor in residence, (earth, wood, cement, tile), and (6) type of toilet facility (none, latrine, indoor plumbing). Occupation is not included as a variable since in this sample it is a constant; all respondents are rural cultivators. Predictors 3 to 6 refer to the condition of the respondent's residence in terms of those characteristics which distinguish Costa Rican peasant homes socioeconomically. Thus, a residence with no electricity, a slat bed, a dirt floor and no toilet facility is considerably poorer than one with electricity, a manufactured bed, a tile floor and indoor plumbing.

Each regression equation uses as the dependent variable one of four distinct modes of political participation. These modes, corresponding closely to the extensive cross-national research conducted by a team headed by Verba, Nie, and Kim (1971), are discussed and analyzed in detail elsewhere (Booth and Seligson, 1975, 1976, 1977b, 1978, 1979; Seligson and Booth, 1976; forthcoming; Seligson, 1978; forthcoming-b) and hence will be described here only superficially. These four modes were isolated using principal components analysis on a matrix of 14 variables drawn from the survey. Independent confirmation of their existence comes from a separate study conducted in Costa Rica at approximately the same time (Booth, Mondol, and Hernández, 1973). The first mode we call "Communal Project Participation." It refers to the extent to which an individual has actively become involved in and has contributed to some type of community project, such as the construction of a local school. This mode is distinguished from "Communal Organizational Activism," which measures the extent to which an individual attends meetings of local committees, such as the school board, community development association, etc. The third mode, which we call "Interaction with Local Government," measures an individual's degree of contact with his local municipal government (*municipalidad*). Finally, there is the "Voting" mode, which is simply a measure of frequency of voting in elections. Scales of each dimension were created so that they could be correlated with the SES predictors. (See Booth and Seligson, 1979).

The results of the stepwise regression analyses are shown in Table 4. The seven predictors (the linguistic index and the six conventional SES indicators) are listed among the rows and the four modes of political participation are placed in the columns. The values shown are the beta weights.

The most important finding revealed in Table 4 is that for three of the four modes of political participation the sociolinguistic index is the best predictor. Only in the voting dimension is the language component of SES a weaker predictor than the more commonly used measures. The value

TABLE 4

Modes of Political Participation and SES Predictors: Beta Weights and F-Tests for the Costa Rican Sample

(Final Step)

Dependent Variables (Political Participation):

SES Predictors	Communal Project Participation		Organizational Activism		Interaction With Local Government		Voting	
	Beta	Sig.	Beta	Sig.	Beta	Sig.	Beta	Sig.
Sociolinguistic Index	.21	<.001	.15	.002	.15	.001	.03	ns
Income	.12	.012	-.02	ns	.13	.006	-.09	.05
Education	.02	ns	-.09	.05	.00	ns	.15	.002
Electricity	-.02	ns	.00	ns	-.12	.024	.06	ns
Bed Type	.09	ns	.05	ns	.01	ns	.01	ns
Floor Type	-.08	ns	-.07	ns	.00	ns	.06	ns
Toilet Type	-.07	ns	.00	ns	.02	ns	.03	ns
Multiple R[a]	.27		.15		.22		.18	

[a] All multiple R's significant at the .004 level or better.

of examining the linguistic component of SES is clearly substantiated by these findings; for the first two modes of participation examined, the inclusion of the linguistic index doubles the amount of variance explained by the traditional SES measures and it significantly increases the explained variance for the third mode.[20]

We hasten to point out that the multiple R's for all four modes are quite small; seven predictors acting together explain only a small portion of the variance on each mode. While it would take us too far afield in this paper to attempt to explain the reasons for these low correlations, they were not unexpected because of our "worst case" methodology. Cross-national research (Huntington and Nelson, 1976, pp. 81–85) has consistently found a weak relationship between SES and participation even when cross-section samples are used. In the five-nation study conducted by Verba, Nie and Kim (1971, pp. 53–60), for example, the highest average correlation between SES (measured by education) and participation reached only .23, SES being correlated with the communal activity mode. The correlation with voting was the lowest, averaging for all five nations only .01. Interestingly, the findings in our Costa Rica study parallel those of the Verba, Nie, and Kim research in that our strongest correlations were also found with community activities and the weakest relationship was with voting. The consistency of these findings justifies our analysis of these relationships despite their weakness.

Moreover, we have shown that the sociolinguistic index can significantly increase the predictive power of the traditional SES measures. Our

[20] A matrix of simple Pearson r correlations of the seven variables (i.e., the six SES indicators plus the linguistic index) was examined in order to check for the possibility of multicollinearity among the predictors. The linguistic index correlates at a fairly low level ($r = .10$ to $.19$) with each of the conventional SES measures except education, with which it has a moderate correlation ($r = .39$). The strength of relationship among the conventional indicators ranges from low ($.13$) to moderate ($.45$). Hence, multicollinearity appears not to be a problem for the subsequent data analysis. Verification of the absence of multicollinearity was provided by examining the tolerance levels at each step in the regression run. These remained consistently high (always above .65 and typically above .80) indicating that the predictors are largely independent. Furthermore, the stepwise regression program utilized may remove variables already entered into the equation based on an F-Test. After it is entered on the first step the linguistic index is never removed. A further test for multicollinearity was performed by re-running the regression analyses displayed in Table 4 by first entering all of the traditional SES predictors and then, on the final step, permitting the entrance of the linguistic index. Even when the linguistic index is entered last it still becomes the best predictor of each of the first three modes of participation, thereby supporting the independent contribution that the linguistic index makes to the equation.

other research has shown, however, that political participation in peasant society is also related to other factors, particularly ecological ones such as a region's remoteness, levels of service infrastructure and socioeconomic development (Booth and Seligson, 1975, 1976, and 1978). Thus, if the variable of remoteness (i.e., distance of the dwelling unit from the county seat and the national capital) is added to the regression equation, the multiple R for organizational activism increases from .15 to .25; for interaction with local government it goes from .22 to .34; and for voting it increases from .18 to .30. Nevertheless, what is important about the findings presented in Table 4 is that the linguistic index is the best SES-related predictor of three modes of political participation in Costa Rican peasant society.

Conclusions

It should be kept in mind that the sociolinguistic index used for this study has tapped only one of the three components of language, namely the phonological. At least one other of the three elements, the morphosyntactic, or grammatical, is also known to vary with SES (e.g., see Wolfram, 1969). The third component of language, the semantic, embodied in the lexicon, might vary with SES as well, but little research has been conducted in this area. If an index of perfcrmance including all three components of language were available, a more powerful prediction of political behavior might be obtained, assuming that lexical and grammatical performance are not completely multicollinear with phonological production. In a follow-up study (N = 753) we conducted in Costa Rica in 1976 (Seligson and Wachong, 1975), we included three grammatical items along with phonological ones.[21] We find that although the phonological and grammatical items co-vary, they are far from multicollinear. However, the grammatical items are not found to be as good predictors of political participation as are the phonological. Consequently little appears to be gained by their inclusion in an overall index. Grammar is apparently more subject to conscious control on the part of the respondent than is phonology and therefore offers less fertile terrain for future analysis. However, Wolfram (1969) finds that grammatical variation ·may provide a clearer indication of SES

[21] The grammatical variables were: *haya* [aya] vs. [ayga], the latter having been historically a nonstigmatized form (3rd person singular present subjunctive of *haber*, "to have"); *trajeron* [traheron] vs. [trahyeron] (3rd person plural preterit of *traer*, "to bring," whose prestige form is irregular, and whose stigmatized ending follows the regular pattern of -er verb endings); and *peor* vs. *más malo*, "worse" (similarly, prestige *mejor* has its frequently heard counterpart, *más bueno*, "better").

where sharp class differences emerge (e.g., between the lower and middle classes). Since our data from the follow-up study is limited to the peasant sector alone we are unable to verify Wolfram's findings.

To sum up, we have treated linguistic variation, in particular variation at the phonological level, as a component of SES, and have found it to be a better predictor of political behavior than other, standard measures of SES, including income, education, and household artifacts. Therefore, social scientists who are interested in developing more sensitive measures of SES should consider including in their questionnaires linguistic variables since, as our research shows, the payoffs may be considerable.

Manuscript submitted 8 March 1976
Final manuscript received 17 October 1977

APPENDIX

I. Accent Shift
  aí → ái (e.g., *maís, país, raíz*)
II. Vowel Change
  e → i (e.g., *león, peor, peón*)
  o → e (e.g., *oscuro*)
III. Consonantal Change
  A. Fricatives and Stops
    h → f (e.g., *jornalero, frijoles, juego*)
    f → h (e.g., *afuera, fulano, fuego*)
    b, ƀ → g, ǥ (e.g., *vomitar, abuela*)
    ∅ → g (e.g., *huevo, hueco, huelga*)
    k → ∅ (e.g., *inyección, exacto, correcto*)
  B. Liquids (l, r)
    d → l (e.g., *ataúd, admirar*)
    r → l (e.g., any infinitive followed by a direct or indirect object
       pronoun: *matarlo, comprarla, venderlo*)
    r → ∅ (e.g., *problema*)

Some readers may wonder whether the fact that syllable, word, and pause boundaries promote certain kinds of variation in Spanish (e.g., stop vs. fricative, such as [b ~ ƀ, d ~ đ, g ~ ǥ]), might have affected variation in the pronunciation of our variables. The answer would have to be no, because despite the fact that such boundaries do play an important role in Spanish phonological variation, their role has been held constant since (1) all the words were elicited in the same phonetic context (i.e., in isolation, which is to say, after a pause boundary) so that there was no source of variation at the inter-word level, and (2) the position of a given variable within a word was held constant, in that its place relative to other segments was fixed. What we are saying, then, is that for a variable such as [o → e], undoubtedly there will be a much higher likelihood of [e] occurring when it does not follow a pause boundary, and particularly if preceded by [e] (e.g., Se pon*e* *e*scuro, "It gets dark"). However, our respondents were not given the opportunity to embed the lexical items

in discourse. All they were allowed to say was a single word, without a larger phonetic context.

Furthermore, some readers might claim that speed of speech must have been determining which variant was produced for a given variable. Although not disputing the contention that speed of speech may play an important role in pronunciation variation in Spanish, if we correlate rate of speed with stylistic level, as does James Harris (1969), then we see that the speed with which a given respondent pronounced a particular word was a function not of some characteristic peculiar to him, but of the stylistic level at which he was speaking, and *this was standardized for all the respondents* in the study. According to Harris' scheme, we may say that the respondents were responding in an "andante" style (1969, p. 7) (i.e., "moderately slow, careful, but natural"). Whether we use the notion of "speed" or not in defining stylistic level, we find, nonetheless, that the use of stigmatized forms goes up as speed goes up, and, concomitantly, carefulness goes down, so that, as our purely linguistic "pretest" has shown, at any stylistic level, upper SES respondents score higher (use fewer stigmatized forms) than do lower SES respondents.

REFERENCES

Agüero, Arturo. 1962. *El español de América y Costa Rica*. San José: Editorial Antonio Lehmann.
———. 1964. *El español de Costa Rica/y su atlas lingüístico*, Madrid: Ediciones Cultura Hispánica.
Anshen, Frank. 1969. Speech variation among Negroes in a small southern community. New York: Unpublished Ph.D. dissertation, New York University.
Arroyo Soto, Victor Manuel. 1971. *El habla popular en la literatura costarricense*. Ciudad Universitaria "Rodrigo Facio." Costa Rica: Tesis de grado, No. 18.
Atwood, E. Bagby. 1962. *The regional vocabulary of Texas*. Austin, Texas: University of Texas Press.
Azofeifa, Isaac Felipe. 1947. *Pronunciamos nuestra lengua y Conversaciones sobre literatura costarricense*. San José: Librería Atenas.
Babington, Mima and E. Bagby Atwood. 1961. *Lexical usage in southern Louisiana*. Publication of the American Dialect Society 36.
Berk-Seligson, Susan. 1976. Costa Rican listener response to socially-based phonological variation: A subjective reaction test. Unpublished manuscript, University of Arizona.
——— and Mitchell A. Seligson. 1978. The phonological correlates of social stratification in the Spanish of Costa Rica. *Lingua*, 45 (No. 1, 1978): 1–28.
Bernstein, Basil. 1971.*Class, code and control: Vol. 1, Theoretical studies towards a sociology of language*. London: Routledge and Kegan Paul.
Booth, John A., Miguel Mondol V. and Alvaro Hernández C. 1973. *Estudio de tipología de comunidades. Tomo II*: San José: Dirección Nacional de Desarrollo de la Comunidad-Acción International Técnica.
Booth, John A. and Mitchell A. Seligson. 1975. Dimensions of political participation among Latin American peasants. Paper delivered to the Southwestern Political Science Association meeting, San Antonio, Texas, March 27–29, 1975.

────── and ──────. 1976. Peasant political participation: An analysis on two Costa Rican samples. Paper delivered to the seminar on the Faces of Participation in Latin America: A New Look at Citizen Action in Society, San Antonio, Texas, November 12–13, 1976.

────── and ──────. 1978. Images of political participation in Latin America. In John A. Booth and Mitchell A. Seligson, eds., *Political participation in Latin America, Vol. 1: Citizen and state*. New York: Holmes and Meier.

────── and ──────. 1979. Peasants as activists: A reevaluation of political participation in the countryside. *Comparative Political Studies*, (Spring).

Bright, William and A. K. Ramanujan. 1962. Socio-linguistic variation and language change. Proceedings of the Ninth International Congress of Linguistics.

Canfield, Delos Lincoln. 1962. *La pronunciación del español en América*, Bogotá: Instituto Caro y Cuervo.

Dobles, Fabian. 1966. *Historias de Tata Mundo*. San José: Editorial Costa Rica.

Echeverría, Aquileo J. 1953. *Concherías*. San José: Editorial Trejos Hermanos.

Fallas, Carlos Luis. 1963. *Marcos Ramírez*. San José: Editorial Trejos Hermanos.

Ferguson, Charles A. 1959. Diglossia. *Word*, 15 (August 1959): 325–340.

Fisher, John L. 1958. Social influences on the choice of a linguistic variant. *Word*, 14 (April 1958): 47–56.

Fishman, Joshua A., Robert L. Cooper, Roxana Ma., et al. 1971. *Bilingualism in the barrio*. Bloomington: Indiana University Publications, Language Science Monographs.

Gilliéron, Jules and J. Mongin. 1905. *Étude de géographic linguistique: Scier dans la Gaule romaine de Sud et l'Est*. Paris.

────── and Mario Roques. 1912. *Études de géographic linguistique*. Paris.

Gumperz, John J. 1959. Dialect differences and social stratification in a North Indian village. *American Anthropologist*, 60 (August 1958): 668–682.

──────. 1964. Linguistic and social interaction in two communities. In J. J. Gumperz and Dell Hymmes, eds., *The ethnography of communication*. Special Publication: *American Anthropologist* 66 (6), Part 2: 137–53.

Harris, James W. 1969. *Spanish phonology*. Cambridge, Mass.: M.I.T. Press.

Hymes, Dell. 1970. Linguistic aspects of comparative political research. In Robert T. Holt and John E. Turner, eds., *The methodology of comparative research*. New York: The Free Press.

Iyengar, Shanto. 1976. Assessing linguistic equivalence in multilingual surveys. *Comparative Politics*, 8 (July 1976): 577–589.

Jaberg, Karl and Jakob Jud. 1928–1940. *Sprach-und Sachatlas Italiens und der Südschweiz*. Sofingen.

Kirwin, William, ed. 1968. *Regional studies in English: Newfoundland*. St. John's Department of English Language and Literature, Memorial University of Newfoundland.

Kurath, Hans, et al. 1939. *Handbook of the linguistic geography of New England*. Providence, R.I.: American Council of Learned Societies.

Kurath, Hans. 1949. *A word geography of the Eastern United States*. Ann Arbor, Mich.: University of Michigan Press.

────── and Ravin I. McDavid, Jr. 1961. *The pronunciation of English in the Atlantic States*. Ann Arbor: University of Michigan Press.

Labov, William. 1964. Phonological correlates of social stratification. In J. J. Gumperz and Dell Hymes, eds., *The ethnography of communication*, Special Publication: *American Anthropologist*, (December 1964) Part 2: 164–176.

——. 1966. *The social stratification of English in New York City*. Washington, D.C.: The Center for Applied Linguistics.

——. 1972. *Sociolinguistic patterns*. Philadelphia: University of Pennsylvania Press.

Lambert, Wallace E. 1967. A social psychology of bilingualism. *The Journal of Social Issues*, 23 (January 1967): 91–109.

——, Moshe Anisfeld, and Grace Yeni-Komshian. 1965. Evaluational reactions of Jewish and Arab adolescents to dialect and language variations. *Journal of Personality and Social Psychology*, 2 (July 1965): 84–90.

——, Hannah Frankel, and G. R. Tucker. 1966. Judging personality through speech: A French-Canadian example. *The Journal of Communications*, 16 (December 1966): 305–321.

Lapesa, Rafael. 1965. *Historia de la lengua española*, 6th edition, Madrid: Escelicer, S.A.

Levine, Lewis and Harvey J. Crockett, Jr. 1966. Speech variation in a Piedmont community: Post-vocalic r. In Stanley Lieberson, ed., "Exploration in Socio-linguistics," *Sociological Inquiry*, 36 (Spring 1966): 204–226.

Lyra, Carmen. 1966. *Cuentos de mi Tía Panchita*. San José: Litografía e imprenta Costa Rica.

Markel, N. N., R. M. Eisler, and H. W. Reese. 1967. Judging personality from dialect. *Journal of Verbal Learning and Verbal Behavior*, 6 (February 1967): 33–35.

Mather, J. Y. and H. H. Speitel, eds. 1975. *The linguistic atlas of Scotland*. Hamden, Conn.: Anchor Books.

McDavid, Jr., Raven I. 1958. The dialects of American English. In W. Nelson Francis, *The structure of American English*. New York: The Ronald Press Company.

McMillan, James B. 1946. A phonology of the Standard English of East Central Alabama. Chicago: Unpublished Ph.D. dissertation, University of Chicago.

Menéndez Pidal, Ramón. 1966. *Manual de gramática histórica española*, 6th edition, Madrid: Espasa-Calpe, S.A.

Pederson, Lee. 1965. *The pronunciation of English in metropolitan Chicago*. Publication of the American Dialect Society 44.

Salguero, Miguel. 1967. *A través del terruño*, San José: Editorial Costa Rica.

Sankoff, David and Henrietta Cedergren. 1974. The dimensionality of grammatical variation. Technical Report 438, Centre de Recherches Matématiques, Université de Montreal. Mimeographed.

Sankoff, Gillian and Henrietta Cedergren. 1971. Some results of a sociolinguistic study of Montreal French. In Regna Darnell, ed., *Linguistic Diversity in Canadian Society*. Edmonton, Canada: Linguistic Research Inc.

Seligson, Mitchell A. 1974. *The peasant and agrarian capitalism in Costa Rica*. Pittsburgh: Ph.D. dissertation, University of Pittsburgh.

——. 1977a. Prestige among peasants: A multidimensional analysis of preference data. *American Journal of Sociology*, 83 (November 1977): 632–652.

——. 1977b. Trust, efficacy and modes of political participation. Paper presented to the Southwestern Political Science Association meeting, Dallas, Texas, April, 1977.

————. 1978. Development and participation in Costa Rica: The impact of context. In John A. Booth and Mitchell A. Seligson, *Political participation in Latin America, Vol. 1: Citizen and state.* New York: Holmes and Meier.

————. forthcoming-a. *The demise of the yeoman: Agrarian capitalism and peasant insecurity in Costa Rica.* Madison: University of Wisconsin Press.

————. forthcoming-b. Unconventional political participation: Cynicism, powerlessness, and the Latin American peasant. In Mitchell A. Seligson and John A. Booth, eds., *Political participation in Latin America, Vol. 2: Politics and the poor.* New York: Holmes and Meier.

————, and John A. Booth. 1976. Political participation in Latin America: An agenda for research. *Latin American Research Review,* 11 (Fall 1976): 95–119.

———— and ————. forthcoming. Structure and levels of political participation in Costa Rica: Comparing the countryside with the city. In Mitchell A. Seligson and John A Booth, eds., *Political participation in Latin America, Vol. 2: Politics and the poor.* New York: Holmes and Meier.

————, and Elena A. Wachong. 1975. Reducing the structural incentives for high peasant fertility in Latin America. Research proposal funded by the Population Policy Research Program of the Ford and Rockefeller Foundations.

Shuy, Roger W. 1967. *Linguistic correlates of social stratification in Detroit speech.* Washington, D.C.: Office of Education Corporation Research Project No. 6-1347.

————, Walter A. Wolfram, and William K. Riley. 1968. *A study of social dialects in Detroit.* Washington, D.C.: U.S. Department of Health, Education and Welfare, Office of Education: Bureau of Research, Final Report Project No. 6-1347.

Trudgill, Peter. 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. *Language in Society,* 1 (October 1972): 179–195.

Tucker, G. Richard and Wallace E. Lambert. 1972. White and Negro listeners' reactions to various American-English dialects. In Joshua A. Fishman, ed., *Advances in the sociology of language,* Vol. II. The Hague: Mouton.

Underwood, Garry. 1974. American English dialectology. *Linguistics,* no. 128 (May 1974): 19–40.

Verba, Sidney and Norman H. Nie. 1972. *Participation in America: Political democracy and social equality.* New York: Harper and Row.

Wagner, Heinrich. 1958. *Linguistic atlas and survey of Irish dialects.* Dublin: Dublin Institute for Advanced Studies.

Wilson, H. Rex. 1958. *The dialects of Lunenburg County, Nova Scotia.* Ann Arbor: University of Michigan Press.

Wilson, Jack L. 1970. A generative phonological study of Costa Rican Spanish. Ann Arbor: Unpublished Ph.D. dissertation, University of Michigan.

Wolfram, W. A. 1969. *A sociolinguistic description of Detroit Negro speech.* Washington, D.C.: Center for Applied Linguistics.

Wood, Gordon. 1960. An atlas survey of the interior South. *Orbis,* 16 (1960): 7–11.

————. 1963. Dialect contours in the Southern States. *American Speech,* 38 (December 1963): 243–255.