THE PHONOLOGICAL CORRELATES OF SOCIAL STRATIFICATION IN THE SPANISH OF COSTA RICA

Susan BERK-SELIGSON
University of Arizona, Dept. of Linguistics, Tucson, U.S.A.
Mitchell A. SELIGSON
University of Arizona, Dept. of Political Science, Tucson, U.S.A.

This study examines Costa Rican Spanish phonology, relating linguistic variation to differences in socio-economic status, as well as to differences in speech formality levels. The study finds that as the level of formality increases the use of prestige phonological forms increases. Of the wide range of socio-economic variables measured, a particular subset, consisting of household artifacts, house construction materials, and education, as grouped together using factor analysis, are shown to be the strongest measures of linguistic performance. Finally, using multiple regression, it is found that the level of formality which most highly correlates with S.E.S. factors is the most formal level.

Introduction

In the growing field of sociolinguistics Latin America stands out as one area where little original research is being conducted, as Lavendera (1974) has recently noted. In an attempt to help fill the lacuna in sociolinguistic research on New World Spanish, the present study focuses on a series of Costa Rican phonological variables, following the efforts of others who have been empirically testing the hypothesis that "[...] the structure of language is related at various levels of abstraction to social structure" (Hasan 1973: 253), and that this relationship can be demonstrated by predictable correlations. A number of quantitative studies dealing with phonological and grammatical variables (Fishman et al. 1971; Labov 1972a; Levine and Crockett 1966; Sankoff and Cedergren 1971; Shuy et al. 1967; Shuy et al. 1968; Wolfram 1969) have attested to the existence of such a relationship. In particular, a variety of linguistic variables have been shown to correlate both with social stratification as well as with context, the latter viewed from the perspective of levels of formality, or attention paid to speech.

The present study[1] is set within the analytical framework of the ones mentioned above, although it diverges from them in its use of certain methodological innovations. Specifically, through the use of a large number of socio-economic status (S.E.S.) variables beyond merely income, occupation, and education, it was possible to discriminate finely among the informants of this study along three separate dimensions of S.E.S. and then correlate these status dimensions with linguistic output. What was found, among other things, is that material demonstrators of wealth (e.g. household artifacts, house construction materials, etc.) constituted the best predictor of phonological production, better even than occupation or income. Such a finding was made possible only because S.E.S. was determined *individually* for each informant, and only *after* the interviews were completed, rather than in an *a priori* manner through the use of neighborhood census data, as has been the practice in many previous studies of this nature. Interpretation of the S.E.S. data was greatly facilitated through the use of the data reduction technique of factor analysis.

Procedures

S.E.S. variables

As has been mentioned earlier, when it came to measuring socio-economic status, what was done here departs rather sharply from the approach used by other studies. Whereas most studies have limited their inclusion of socio-economic variables to education, occupation, and income, the present study gathers data on a series of variables (beyond the standard three major ones found to be most significant in most social science research), variables which we found to be especially important in discriminating finely between persons of varying socio-economic status in Costa Rica.[2] Thus, a description of the house was included, one which included such things as the materials the house was constructed of (e.g. the type of front door, windows, floor, walls and ceiling), facilities (e.g.

[2] Data of this nature has been found to be extremely useful in studies such as the one by Xavier Albó (1970), for example.

piped-in water, electricity, bathroom), and artifacts (e.g. furniture, vehicle, electrical appliances). Labov (1972a: 115) himself points out the utility of using a combined index of socio-economic indicators in correlating linguistic variables with socio-economic ones, explaining that when linguistic variables are "correlated with the individual social indicators of productive status – occupation, education and income [...] it appears that no single indicator is as closely correlated with linguistic behavior as the combined index". However, even a combined index of these three important variables is less adequate for accurately determining socio-economic status than is an index comprising a larger number of variables (see Hasan 1973: 257).

The question which arises, naturally, is what should the variables be beyond the three traditionally used ones. As Grimshaw (1966: 192) so aptly states it, "Sociologists are frustrated by the apparent necessity to choose between simple one-variable indicators (e.g. occupation) which explain much of the variance but leave unanswered questions of status inconsistency, and the more cumbersome, nongeneralizable, atheoretical indices [...]". But whether it is one variable or many that are used to measure socio-economic status, if data is gathered via census returns rather than by collection directly from the very persons who are being interviewed, then the researcher runs the risk of mistakenly attributing characteristics to individuals in his sample which in reality are not true of them. This is what has been referred to as "measurement by fiat" (Macaulay 1970: 767). In a review of Wolfram's *A Sociolinguistic Description of Detroit Negro Speech* (1969) Macaulay (1970) criticizes Wolfram for using this method of measuring social status. What Macaulay (1970: 767) says explains in effect our reasons for including into the index of socio-economic data the number and kinds of variables that were enumerated above.

Social stratification indices of the kind that can be assembled from census returns are no doubt useful for describing gross differences of behavior in large populations, but that is not what we're attempting in sociolinguistic work. Instead, we are interested in relatively small differences of behavior in groups the size of which we can hardly guess at [...].

In calling for the collection of sufficiently accurate and detailed information on the part of the fieldworker, Macaulay (1970: 768) notes that,

[...] any markedly deviant behavior on the part of even one informant may seriously distort the results [...]. In such circumstances, it is important that social status (or any other variable) should not be assigned on the basis of arbitrary numerical indices alone, but that

the actual social status of the informant should be estimated as carefully as possible by the interviewer in the course of his fieldwork. In this respect, it may be time for sociolinguists to update their sociology in the direction of those who are interested in the problem of SEEING THE SOCIETY FROM WITHIN (e.g. Garfinkel, Cicourel, Sudnow).

It is felt that the S.E.S. variables selected for measurement in this study have been chosen on the basis of 'seeing the society from within', and that, for this reason, they are highly accurate measures of socio-economic status.

Linguistic variables

The linguistic variables that were chosen for consideration in this study were phonological rather than grammatical or lexical, this for the reasons pointed out by Labov (1972a: 110–121): (1) they provide us with a large body of data from relatively small samples of speech; (2) their variation is largely outside of the conscious control of the speaker; and (3) of all linguistic systems, phonological ones "[...] show the highest degree of internal structure [...] and thus provide the investigator with an extensive series of parallel and convergent results". In an effort to discover which phonological variables were socially relevant we found that a large number had been identified in Costa Rica as being characteristic of lower class speech (Agüero 1962, 1964; Arroyo 1971; Azofeifa 1947; Salguero 1967; Wilson 1970).[3] From among the ones so identified, 25 variables summarized into three major categories (Accent Shift, Vowel Change, Consonantal Change) were selected for use in this study. For a complete listing of these variables, see the Appendix.

The occurrence of these variables in the speech of the informants interviewed was not analyzed in the light of any conditioning factors, whether phonological, grammatical, or otherwise. This certainly would be an object of study in future investigations of Costa Rican dialectal variation. Given the present large number of variables, however, sorting out the effect of linguistic environment upon them would have been an unfeasible task, since it would have entailed increasing by tenfold the number of items to be elicited.

What can be said tentatively at this time regarding the role of a conditioning environment in determining the distribution of the variables –

[3] It should be noted that the variables used in this study are by no means peculiar to the Spanish of Costa Rica. They are widespread phenomena found throughout Latin America, and, for the most part, have their origins in peninsular Spanish dialects (see, for example, Canfield 1962; Lapesa 1968; Menéndez Pidal 1966).

and this must be stated as an impressionistic observation rather than as a conclusion derived from rigorous analysis – is that grammatical factors appear not to condition the occurrence of the non-prestige forms. In particular, part of speech seems to play no role, as the following examples will indicate:

I.   Accent shift
     (verb) *reir*: [réir]
     (noun) *maíz*: [máis]
     (adj.) *increíble*: [inkréible]

Especially revealing in illustrating this point for the case of accent shift is the couplet *oído* (noun) and *oído* (past participle) which both occur as either [óido] or [óydo].

II.   Vowel change
      (noun) *poeta*: [pueta]
      (adj.) *feo*: [feu]

III.   Consonantal change
       A. [d, l, r]
       (verb) *admirar*: [almirar ~ armirar]
       (noun) *admiración*: [almirasyón ~ armirasyón]
       (adj.) *fácil*: [fasir]

       B. Nasals
       (verb) *ignorar*: [iŋnorar ~ innorar]
       (noun) *himno*: [iŋno ~ inno]
       (adj.) *ignorante*: [iŋnorante ~ innorante]

       C. Stops and fricatives
       (verb) *vomitar*: [gomitar]
       (noun) *buey*: [gwey]
       (adj.) *apsurdo*: [absurdo ~ aksurdo ~ awksurdo ~ asurdo]

The above examples are by no means meant to constitute a proof, for it might turn out, for example, that even though both verbs and nouns sometimes undergo accent shift, nouns may undergo it most of the time, while verbs may do so only rarely.

Phonological environment, on the other hand, seems to be largely

determinant in conditioning the occurrence of the non-prestige forms, many of the specific environments having been worked out by Wilson (1970). Although phonological environment was not taken into account in the process of selecting the lexical items for use in this study, it should be kept in mind that in the standardized portions of the interview (the ones which elicited variables for levels 3, 4, and 5), phonological environment was held constant, in that the conditioning elements surrounding the variables were identical, since it was a specific lexical item that was elicited for any given question in every case. Such control of environment was lost, of course, at levels 1 and 2, where variables were produced within a context of open-ended conversation.

That the stigmatized variants of the variables under consideration are indeed socially salient and do hold low social value for Costa Ricans is a thesis that could be supported by some type of subjective reaction test. Although no such test was conducted on this particular sample, one in fact was carried out on Costa Rican elementary and high school students concurrently with the present study. The specific type of subjective reaction test is the 'matched guise technique' developed by Wallace Lambert and his associates (Lambert 1967; Lambert 1965; Lambert, Frankel, and Tucker 1966; Markel 1967; Tucker and Lambert 1972), which was modified and adapted for use on Costa Rican Spanish. The results of that study have been analyzed (Berk-Seligson 1976) and show a clear consciousness of the social significance of the non-prestige forms.

The analysis of the linguistic data collected in this study proceeds by creating summary indices of all 25 variables at each of the formality levels (discussed below). The summary approach, as opposed to an analysis of each of the 25 variables taken separately, was taken in order to help reveal the overall pattern of interrelationships between the S.E.S. variables and the linguistic variables. In future studies we plan to study the linguistic variables individually. Such a presentation here, however, would unnecessarily burden the discussion.

Formality levels

As far as the formality levels (or 'contextual styles' (Fishman 1971)) are concerned, they were defined and elicited in the manner of Fishman et al. (1971). What must be made clear is that these formality levels are not meant to be viewed as existing styles, that is, replicas of real speech situations, but rather as analogs of them. Thus, the styles are actually

elicitation procedures which merit the term 'styles' only so long as "... the speech samples observed by these procedures *do* continuously and measurably (hence systematically) shift along a hypothetical style continuum" (Fishman et al. 1971: 368). Thus there are five levels of formality, each of which, according to Labov's conceptualization, is a reflection of the speaker's attention paid to his speech. The five levels, going from least formal to most formal, are the following: (1) casual style, (2) interview style, (3) word naming and leading questions, (4) passage reading, and (5) word list reading. With regard to the elicitation procedure used for each, the most formal level was elicited by having the informant read a list of words, each of which contained one of the phonological variables under study. The passage reading section of the questionnaire involved the reading of a short narrative in which were embedded the variables; however, in order to divert the informant's attention from the words containing those variables, a technique designed by Levine and Crockett (1966) was employed. Namely, into the passage were inserted blanks which the informant was requested to fill, yet the words that were meant to fill in the blanks were *not* the ones that contained the variables; rather, the variables were embedded in words that were part of the text. Thus, the informant's attention was focused not on the words containing the variables, but rather, on the 'dummy' words which the informant himself was supposed to provide. For example, one sentence read, 'His mother gave him --- juice', where 'juice' (*jugo*) was the word being elicited for study. The third level of formality, the one mid-way between the most and least formal, consists of variables elicited by two means: (1) leading questions (e.g. 'Your mother's mother is your ---') and word naming (e.g. 'What do you call the ferocious animal that lives in the jungle and roars a lot?'). The questions used in this study were basically those used by Wilson (1970), although many modifications had to be made on them once pretesting had determined that they did not easily elicit the variables. As for the fourth formality level, 'interview style', this is composed of any variables occurring during the informant's replies to the socio-economic questions, as well as any speech produced during the word-naming section other than the production of the particular variables being elicited for word-naming. The most informal level, 'casual', was elicited by asking the informant to recount an experience during which he had been in danger of losing his life, either through accident or illness; however, not all the speech produced during the narration was counted as 'casual'. The criteria used for inclusion of the variables as 'casual' were the presence

of one or more 'channel cues' specified by Labov (1964: 168), which were:
(1) changes in tempo, pitch or volume; (2) laughter; and (3) changes in
breathing. All other speech produced during this section of the interview
which did not satisfy these criteria were categorized as Level IV (interview
style).

Admittedly there is one major limitation to the methods of eliciting
speech which are discussed above, and that is the element of artificiality
in the speech situation. Since the five levels of formality are merely
analogs of real situations, the speech that is produced in them is not truly
natural speech. The observation of more natural situations should be
striven for, although the advantage of the method presently employed is
its efficiency and its ability to elicit the same variable in precisely the same
phonetic environment (i.e. the same lexical item), for purposes of com-
parison across formality levels.

Interviewer

Regarding the role of the interviewer in the data gathering process,
there was only one interviewer (Berk-Seligson) doing the fieldwork, which
was an advantage when considered from the point of view of insuring
interviewer consistency in coding (i.e. it eliminated one source of possible
error). However, the fact that the interviewer was a female was somewhat
problematical in light of the all-male sample of informants and the fact
that Costa Rican culture does not favor an unknown young woman's
entering a house and requesting to speak at length to a male. What Ma
and Herasimchuk (1971: 372–373) found to be true of the Puerto Rican
community which they studied applies equally as well to the Costa Rican
case. Specifically, when a Costa Rican man is interviewed by a female
whose educational and occupational achievements seem to him to be higher
than his own, he will feel threatened, and consequently will become
defensive with the interviewer. The higher his level of education and
occupational status, the greater will be his ability to relax during the
conversation. Fortunately, the three standardized sections of the question-
naire were not affected by this factor as the data analysis will show. Its
greatest impact was on the section designed to elicit the most casual speech.
However, the fact that the interview took place in the informant's home
rather than on the interviewer's private territory, helped to ameliorate
somewhat the negative influence of the status differential upon the eliciting
of casual speech.

Field instruments

As far as the field instruments are concerned, essentially, they consisted of a precoded, closed-ended (with the exception of the portions which elicited casual speech) questionnaire, and a tape recorder. The closed-ended sections of the questionnaire were the ones that were precoded; however, even here room was left to fill in any variant which had not been previously anticipated. Thus, if a word containing a given variable was known beforehand to have three variants (one prestige and two non-prestige), then these three were listed on the mimeographed questionnaire, and whichever one the informant pronounced was the one circled at the moment of his uttering it. However, if he responded with an unanticipated variant, then that word was jotted down phonetically on the spot. The tapes were played immediately after each interview, to check the coding.

Although the questionnaire was pretested in Costa Rica, certain unexpected difficulties arose in the word naming/leading questions section. Essentially the problems in eliciting the desired lexical items derived from the following types of differences, differences unknown to us prior to the formulation of the questionnaire: (1) differences in connotation between English and Spanish; (2) differences between English and Spanish arising from polysemy; (3) change in the usage of a lexical item over time; (4) differences in cognitive maps among informants; (5) differences in informants' ability to perform certain logical operations in a testing situation; and (6) differences among informants regarding factual knowledge. Thus, an example of differences in connotation would be the word *admiración*, which in Costa Rica does not mean precisely what the English word 'admiration' does, but rather, encompasses affection and good will, as well as esteem. An instance of differences between English words and their Costa Rican Spanish glosses differing because of polysemy is *toalla* 'towel', which, aside from meaning towel, refers to a sanitary napkin; furthermore, the most commonly used term for towel in Costa Rica is *paño* whose primary meaning is a piece of cloth. Similarly, the word *fusil* refers primarily to a rifle, and secondarily to the penis, so that men are wont to refrain from using the word *fusil* in conversation with women because of the nature of its secondary meaning. An instance of a change in the usage of a lexical item over time is that of *ritmo* 'rhythm' which, from what we gather, is not used by people of approximately 65 or older; instead, the term *compás* is utilized. As for problems arising from differences in cognitive maps, that is, in frames of reference, an example

would be the word *calle* 'street' as a fill-in for the frame, 'The main
avenue of San José is called Central Avenue; and what's the name of the
main street?' (the answer being 'Central Street'). Many people living
even in places such as Sabanilla, the district sampled, which is a 15-minute
bus-ride from the heart of the capital, do not know that the city is divided
up into streets and avenues along quadrants of north, south, east, and
west, the dividing lines being Central Avenue and Central Street. Essen-
tially, the problem is one of recognizing the concepts 'street' and 'avenue'
as artificial organizational constructs in city planning. In places like
Sabanilla, which until recently was still an entirely agricultural area, there
were no 'streets' and 'avenues', and even today the latter are scarce, the
few existing ones being unnamed, as is customary in rural villages. A
further source of difficulty in eliciting specific lexical items bearing the
variables in question, this one stemming from varying degrees of inability
to perform under a test situation certain logical operations such as provid-
ing a word which means the opposite of the stated word, is exemplified by
frames such as, 'One of the children fell asleep, but the other one stayed
---'. Here the desired lexical item was *awake*, but several informants
responded *asleep*. Similarly, the frame 'Joey isn't handsome. On the
contrary, he's ---' was sometimes filled in by such items as *friendly*,
beautiful, and *nice* instead of *ugly*. Lastly, there were unforeseen problems
on the order of lack of basic knowledge, so that some informants could
not produce the word *Atlantic* to complete the frame, 'One of our ports
is on the *Pacific* side of the country and the other port is on the --- side'.

The data

The sample of respondents interviewed, which totaled 48 adult males,
was drawn from the district of Sabanilla in 1973. Only male heads of
household were included for a pragmatic rather than a theoretical reason.
At the same time that this study was being conducted, another one
(Seligson 1974, 1975, 1977a, b, 1978, forthcoming; Seligson and Booth
forthcoming) was being carried out, the latter investigating peasant
political attitudes in relation to land tenure. Among the responses elicited
from the 531 peasants sampled, 13 were linguistic variables and were
among the ones employed in the present study (Seligson and Berk-
Seligson 1978). Since the respondents in the larger study were all males,
for the sake of potential comparison this study was restricted to males as
well, so that there would be a sufficient number of respondents in each

Table 1

S.E.S. characteristics of Sabanilla and the sample.

		Sabanilla[a]	The sample
Population	(men)	1,705	48
Education[b]	none	8%	2%
	primary	68%	52%
	secondary	19%	21%
	university	5%	25%
Illiteracy	(10 years and older)[c]	4.3%	19.1%
Monthly income	less than ₡100[d]	1%	6%
(economically active	100–399	27%	21%
population 12 years	400–699	36%	21%
and older)	700–999	17%	21%
	1,000–1,299	8%	6%
	1,300–1,599	3%	4%
	1,600–2,199	3%	4%
	2,200–2,799	1%	0%
	over 2,800	4%	17%
Occupation	professionals	9%	2%
(economically active	managers/administrators	2%	15%
population 12 years	office workers	11%	6%
and older)	businessmen and salesmen	10%	6%
	farmers	6%	10%
	transportation workers	6%	4%
	craftsmen and machinists	25%	15%
	laborers and farm day workers	8%	29%
	other	23%	13%

[a] Population data from 1973 national census (Dirección General 1974).

[b] The population education data includes all individuals 6 years and older, whereas the sample includes persons who are 18 and older.

[c] Note that the population illiteracy figures are based on 10 years and older, whereas the sample includes persons 18 and older.

[d] Income figures are in *colones* (₡). At the time of the study, 8.6 colones = 1 U.S. dollar. Note that the population income figures are based on 12 years and older, whereas the sample includes persons 18 and older.

cell to be able to make a statistically meaningful comment about the relation of one sample to the other.[4]

Sabanilla belongs to the county of Montes de Oca, and lies on the very

[4] Since the sample was composed of only male informants, nothing definitive can be said regarding the distribution of the selected variables along the lines of gender differentiation. Nevertheless, a hypothesis can be formulated (and it is planned to follow up this study with a comparable one, one including females) based on several factors. It is hypothesized that within each S.E.S. group, in careful speech, women will tend to use fewer stigmatized forms

border of the nation's capital, San José. It is today essentially a suburban area, although a few coffee farms still remain in production, reminders of what the area looked like less than a generation ago. Thanks to the recently published 1973 national census (Dirección General, 1974), we have available to us fresh data on Sabanilla. Table 1 presents comparative S.E.S. information on Sabanilla and the sample of this study. As can be seen, the informants are fairly representative of Sabanilla as a whole, although their level of education is higher, particularly at the upper end of the scale, as is the income (compare 17% of the sample earning more than ₡2,800 per month in contrast to 4% of the district earning that amount). At the opposite extreme, too, the sample included many more (6%) persons earning less than ₡100, whereas among the Sabanilla population only 1% earns such a low income. Furthermore, occupationally, whereas 9% of the district is professional, only 2% of the sample is, the major difference between the two being in the manager/administrator and laborer/farm day worker groups, both of which include a large segment of the sample, while the population of Sabanilla has a more limited percentage of persons working in such occupations. What is interesting is the much higher rate of illiteracy (19.1%) in the sample than in the overall population (4.3%), despite the previously mentioned generally higher levels of education among the former. The fact that the sample includes more persons at the extremes than does the census is probably due to our making sure that we did not ignore such informants in the sample. In other words, we made sure to go to shacks as well as to quite luxurious looking houses, for the former are sometimes overlooked by census takers because they do not even appear on census maps, and the latter, with

---

and will tend to be more sensitive to the prestige forms than men. This hypothesis is made both on the basis of findings to this effect in the studies of Sankoff and Cedergren (1971), Anshen (1969), Labov (1966), Levine and Crockett (1966), and Trudgill (1972), as well as on the basis of several observations that we have made in Costa Rica. First of all, very often the wife of the informant would correct her husband's responses to the standardized items, revealing a greater competence than his with regard to prestige forms. How often this would have happened remains an unknown, since the woman would have to be asked to refrain from further such interfering in order not to bias the informant's responses. Secondly, based on our participant observation in a peasant village, we find that women have a greater contact with prestige forms than do men through their frequent church attendance (for men, on the whole, do not regularly attend religious services) and through their listening to the radio for many hours a day (which men cannot do because of the nature of their work which keeps them out in the fields). Naturally, the hypothesized greater use of the prestige pattern on the part of women must be tested out in Costa Rica in order to be able to make any valid claims in this regard.

their uninviting gates, do not make themselves accessible to visits from unknown persons.

As far as the age of the sample is concerned, while the mean was 43, the informants ranged from 18 to 89 years old, 18 being the minimum age accepted for participation in the study. It should be pointed out that the minimum was set at 18, since, as Labov (1965) has found (at least with respect to American English) it is only at about the age of 14 or 15 that an adolescent's sensitivity to the social consequences of speech begins to resemble that of an adult. At this point (usually the first year of high school) the speaker "becomes sensitive to the social significance of his own form of speech, and other forms", whereas "complete familiarity with the norms of the community seems to be attained at the age of 17 or 18" (Labov 1964: 100). Thus, it is only at the age of 16 or 17 that the ability to use prestige forms is acquired (Labov 1964: 100). Although age-grading in relation to the acquisition of evaluative sociolinguistic norms may not be the same for Costa Rican adolescents, if there is a difference, it would be in the direction of an earlier rather than a later acquisitional stage since most Costa Rican youths do not attend high school, enter the adult world (in terms of financial and family responsibility) at an earlier age, in short, have a briefer period of adolescence than do Americans of an equivalent age group. Thus, cutting the minimum age at 18 seems to be appropriate in the Costa Rican context when it comes to the goal of interviewing persons who have adult sensitivity to language.

Apparently, there is still a lack of adequate knowledge as to what the most efficient size is for studying social dialects reliably. However, along with Labov, some researchers (Wolfram 1969) conclude that for the study of social dialects reliability can be obtained with a smaller sample than is used in other kinds of surveys because of the relatively more uniform nature of linguistic behavior in comparison to other types of social behavior. Labov (1972a: 204), pointing to his own studies (1966; et al. 1968) and to research done by others (Kučera 1961; Shuy et al. 1967), finds samples as small as 25 speakers to be sufficient for revealing basic patterns of class stratification. Specifically, Labov (1972a: 204) finds that,

[...] regular arrays of stylistic and social stratification emerge even when our individual cell contains as few as five speakers and we have no more than five or ten instances of the given variable for each speaker.

Thus, the sample of 48 interviewed in this study should be adequate to deal with the subject it sets out to.

Table 2
Mean scores of five formality levels.

Formality level	Standard deviation	Mean score in per cent	t	Sig.
1. Casual	0.268	47.0		
			1.34	n.s.
2. Interview	0.196	54.2		
			4.46	< 0.001
3. Word-naming	0.151	65.0		
			2.83	0.007
4. Text-reading	0.091	72.4		
			1.91	0.06
5. Word-list reading	0.082	73.9		

Findings

Formality levels and phonological performance

With regard to the hypothesis that as the context becomes more informal the use of stigmatized forms will increase, we find that it is borne out in the Costa Rican case. This is demonstrated by table 2, which presents a mean score for the entire sample at each formality level. The numbers, which range from 0% to 100%, 100% being a theoretically perfect score, that is, one in which there are no mentions of stigmatized forms, and 0% being the case where all the forms mentioned are stigmatized, represent an average 'correct' score (i.e. prestige form) per formality level, of all the items in which variables are embedded. The average prestige score was arrived at by first summing all the variable scores on each formality level for each informant, and then dividing that sum by the total number of variables on that level. This procedure yielded each informant's mean score on each formality level. These means were then averaged for each level so that a mean score of prestige responses for the entire sample on each formality level was obtained.

The actual method of arriving at a score for any given variable was the following. Prestige responses were scored 1, and non-prestige variants were given 0. What happened in those cases where there was more than one stigmatized variant per variable (and this was the case for over 90% of the variables)? Attempts had been made to rank the variants in order of distance from the prestige norm. 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 enough there was no consensus whatsoever 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 totally arbitrary from the point of view of native intuition, and, apparently, achieving a valid rank ordering would have involved a study in itself, with a sufficiently large sample to make the order an accurate one.

The coding system for formality levels 1 and 2 had to be different from that for 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 recordings all the instances in which the variables under study occurred, using 100% of the corpus 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 and 2 was arrived at by dividing the number of stigmatized mentions by the total number of variables mentioned.

Returning now to table 2, what it clearly demonstrates is that the informants use increasingly fewer of the stigmatized variants (the percentages approach 100 rather than 0) as their speech becomes more formal. Thus, the mean scores become higher (approach 100) as we approach the more formal contexts. It should be noted that the range of means illustrated in table 2 is quite narrow, that is to say, the mean scores are not arrayed over the entire range of 0% to 100% but rather are concentrated in the range of 47.0% to 73.9%. While such a narrow range is not unexpected, given the fact that these are averages across all the informants in the sample, the actual scores ranging from 0.08 to 1.0, we would still want to know if these small differences are statistically significant because if some of them are not, it might mean that we have been working with too many levels of formality and that two or more of the levels should have been collapsed. To determine this, a difference of means test (*t*-test) was run on each level of formality, the mean score for each level being paired[5] with the one beneath it (see table 2).

[5] A paired *t*-test should only be run on variables which are positively correlated. The five levels of formality were in fact positively correlated, their Pearson *r*s ranging from a low of 0.48 to a high of 0.85. The test, therefore, is an appropriate one in this case.

A look at table 2 reveals that the differences in means between formality levels 2 and 3 and between 3 and 4 are clearly significant (< 0.001 and 0.007, respectively, the generally accepted level of significance being 0.05), and that the difference in means between levels 4 and 5 barely misses significance. But most striking is the fact that the t-test results for levels 1 and 2 show no significant difference in means, which implies that perhaps levels 1 and 2 should have been collapsed into one category as, too, might levels 4 and 5 have been. A possible explanation for the lack of a significant difference in means between levels 1 and 2 is that the testing situations which elicited variables on these two levels may not have been sufficiently different, so that even though the mean score between levels 1 and 2 is higher, it is not high enough to be statistically significant.

Formality levels and S.E.S.

In this paper the wide range of S.E.S. variables mentioned previously are individually correlated with the formality levels. This offers the researcher a number of very important advantages. First, it does away with the problem of nominal classification of social status (that is, making distinctions in status in terms of gross types, as Labov does, rather than in terms of degree, so that persons are classified as 'lower class', or 'working class' or 'lower middle class', etc.). The difficulty with nominal classification of social status is that one is forced to classify persons on the basis of a very few variables (e.g. occupation, income) instead of the much wider range of variables which are known to comprise important elements of socio-economic status (see Seligson 1977b). Furthermore, by using a nominal classification one is frequently confronted with the difficulty of categorizing persons who, for instance, have had little schooling but earn a high income. In this study, instead of S.E.S. variables being lumped together for each informant, they are individually associated with the informant's responses to the linguistic variables, which obviates the necessity of grouping persons artificially into homogeneous appearing classes. An additional advantage of this approach is that it permits us to determine precisely which elements, or variables, comprising social status best predict linguistic behavior. And, finally, the use of multiple variables in determining social status permits us to employ a technique, namely factor analysis, for sorting out the variables into distinct categories, or dimensions of socio-economic status.

The S.E.S. variables gathered in this study thus were subjected to factor

Table 3
Varimax rotated principal components of S.E.S. variables[a]

	Material demonstration of wealth (Factor I)	Income/ occupation (Factor II)	Poverty markers (Factor III)
Care of grounds	0.85	0.37	0.10
Windows (absence/presence glass)	0.85	0.06	0.14
Walls (unpainted/painted)	0.79	0.14	0.14
Type of bed	0.76	0.22	0.26
Type of front door	0.75	0.37	0.00
Refrigerator	0.74	0.12	0.28
Record player	0.73	0.10	0.11
House building materials	0.73	0.55	0.16
Type of toilet facility	0.70	0.18	0.51
Type of wall decoration	0.69	0.10	0.16
Sewing machine	0.68	−0.09	−0.07
Storage for clothing	0.65	0.20	0.43
Years of schooling completed	0.63	0.37	0.16
Ceiling (absent/present)	0.63	0.17	0.24
Number of bedrooms	0.62	−0.27	0.16
House security measures (abs/pres)	0.63	0.55	0.04
Floor material	0.61	0.45	0.18
Car	0.53	0.35	0.14
Total weekly family income	0.45	0.80	0.08
Father's occupation	0.09	0.80	0.06
Respondent's occupation	−0.13	0.76	−0.12
Piped-in water	0.04	0.08	0.88
Electricity	0.25	−0.09	0.83
Eigenvalue	11.67	2.48	1.46
% of variance	48.6	10.3	6.1

[a] Total N = 48, but varies due to non-response.[b]

[b] Nearly every case of non-response was caused by the informant's illiteracy, and consequent inability to perform the reading tasks required for responses to the elicitation techniques used at Levels 4 and 5. Out of the sample of 48, nine informants (or 19.1%) were functionally illiterate.

We considered having illiterate informants complete the word-list reading and passage reading tasks by repeating the words after the interviewer. We rejected this possibility on the grounds that an element of mimicry would enter into the performance of the informant, an element whose influence could vary from one informant to another, and from one response to another by any given informant. We decided that rather than to deal with the unknown degree of distortion which mimicry might introduce, we preferred to have missing cases.

analysis. This technique is perhaps the most widely used data reduction procedure in use in the social sciences today. The procedure examines a correlation matrix computed on the raw data (in this case, the S.E.S. variables listed in table 3) and searches for common patterns. In so doing the procedure, if it is successful, uncovers the underlying pattern of relationships (i.e. dimensions) in the data such that a fairly large number of variables may be represented parsimoniously. More extended non-technical explanation and illustration may be found in Rummel (1967) and Boni and Seligson (1973). For our purposes, the major advantage of factor analysis is that it enables us to employ a wide range of S.E.S. variables (23 in all) in classifying respondents into their appropriate status position and thus obviates the necessity for using the single variable, nominal scheme which has seen such frequent use.

The results of the factor analysis are presented in table 3. The individual S.E.S. variables which were used to calculate the correlation matrix are listed in the rows. The numbers listed in the three columns are the 'loadings'. These numbers represent the relative weights of each variable in describing a given factor. The higher the loading the more closely related is a given variable to its underlying factor. As a general rule, variables which have a high (positive or negative) loading on one factor and low loadings on all other factors are considered as serving to define that factor. The researcher then examines the pattern of loadings on each factor and gives it a name. The factor analysis presented in table 3 has been given a 'varimax rotation', a commonly used procedure which helps the investigator more clearly visualize the pattern of loadings. After examining the patterns on the three factors we were able to label them as follows: (I) Material Demonstration of Wealth, (II) Income/Occupation, and (III) Poverty Markers.[7] What the factor analysis has accomplished, then, is to group the 23 S.E.S. variables into three clearly interpretable categories. The only anomaly in the solution presented in table 3 is that education ('number of years of schooling completed') loads on the factor denominated 'Material Demonstration of Wealth'. However, as is re-

[7] While this solution is quite clear, in most factor analyses which include a large number of variables a problem of distributed loadings often appears, that is to say, one variable might load highly on two or more factors. This occurred with a few variables in the study, as an examination of table 3 will show. In particular, several material demonstrators of wealth loaded highly on both Factors I and II, which is only to be expected since the presence of material demonstrators of wealth is largely a result of income. The variables in question were considered to be part of the factor on which they loaded most highly, a procedure which led to factors that were open to meaningful interpretation.

vealed below, the patterns of correlation between linguistic performance and the Material Demonstration of Wealth variables, on the one hand, and between linguistic performance and education, on the other, are very similar, hence justifying the grouping together of variables which are substantively distinct. We have called the third factor Poverty Markers because only the poorest individuals in our sample had no piped-in water or electricity.

Having organized the S.E.S. variables into three categories through factor analysis we then proceeded to examine the relationship between each of the formality levels and each of the S.E.S. variables as they are grouped into their individual factors. The resulting correlation matrix appears in table 4. The importance of the data on Material Demonstration of Wealth (Factor I) in predicting linguistic performance becomes apparent upon examining these correlations. Nearly all the variables on Factor I, across all formality levels, are solidly related to linguistic performance. In marked contrast to these strong relationships are the much less impressive relationships found on Factor II. Hence, Material Demonstration of Wealth is a better predictor of linguistic performance than is Income/Occupation.

An examination of table 4 makes it difficult for us to determine if any particular formality level or group of formality levels reveals stronger relationships between linguistic performance and our measures of S.E.S. than any other. Although we now have the linguistic variables summarized into five groups (by formality levels) and the S.E.S. variables grouped into three factors, we still are left with 115 correlations to examine. To further simplify the data and to make it easier to understand we can reduce table 4 by computing 'factor scores' on the three S.E.S. factors. A factor score is a composite scale computed for each individual in the study on each factor. Those informants who, for example, have a high score on income and occupation, would have a high factor score on Factor II. Since factor scores are weighted in proportion to the loading of each variable on a given factor, the individual variables that have higher loadings make a greater contribution to the factor score than do those which have lower loadings.[8] For instance, referring to Factor I, 'care of

---

[8] Factor scores, which have a mean of 0 and a standard deviation of 1, are calculated by multiplying the standardized value (i.e. 'Z-score') of each variable by the factor score coefficient, and summing the products for all the variables in the analysis. Thus, the complete estimation method is used here. Missing data (which occurred with one case on two variables) were replaced by the conservative method of substituting the mean of the variable.

Table 4

Statistically significant correlations of S.E.S. variables with formality levels.[a]

	Casual	Interview	Word naming	Passage reading	Word-list reading
Levels:	1	2	3	4	5
Factor I; Material demonstration of wealth					
Care of grounds	0.52	0.37	0.40	0.49	0.57
Windows (absence/presence glass)	0.29	0.35	0.33	0.34	0.24
Walls (unpainted/painted)	0.41	0.37	0.40	0.46	0.33
Type of bed	0.48	0.48	0.41	0.32	0.39
Type of front door	0.47				
Refrigerator		0.29	0.35	0.27	0.27
Record player	0.36	0.36	0.28		
House building materials	0.40		0.25	0.27	0.29
Type of toilet facility	0.29	0.29	0.45	0.28	
Type of wall decoration	0.36	0.30			
Sewing machine	0.40	0.25			
Storage for clothing	0.33	0.25	0.37	0.35	0.40
Years of schooling completed	0.42	0.41	0.48	0.42	0.45
Ceiling (absent/present)	0.38	0.41	0.36	0.36	0.34
Number of bedrooms		0.28	0.43		0.31
House security measures (abs/pres)	0.32		0.26	0.34	0.36
Floor material (indoors)				0.27	0.27
Car					
Factor II; Income/occupation					
Total weekly family income					0.31
Father's occupation	0.33	0.29			
Respondent's occupation					
Factor III; Poverty markers					
Piped-in water	0.33		0.43		
Electricity			0.32		

[a] Significant at the 0.05 level or better; all correlations are Pearson rs. Total N = 48, but varies due to non-response.

grounds' has a loading of 0.85, whereas 'car ownership' loads only at 0.53, indicating that car ownership makes a considerably smaller contribution to the factor score than does care of grounds.

What we are left with is an easy to interpret table (see table 5) which presents the correlations of the factor scores with the formality levels. Hence, in order to facilitate interpretation in our effort to uncover the central patterns in the data, we have reduced our massive correlation matrix to a small 3 × 5 table. We find there, first, of all, a consistent pattern of significant correlations across the five levels of formality on

Table 5

Statistically significant correlations of factor scores with formality levels.[a]

		Formality levels				
		1	2	3	4	5
Material demonstration of wealth	(Factor I)	0.39	0.43	0.37	0.42	0.45
Income/occupation	(Factor II)					
Poverty markers	(Factor III)			0.36		
Multiple *R*s		0.45	0.44	0.51	0.52	0.61
% Variance		20%	19%	26%	27%	37%

[a] Significant at the 0.05 level or better; all correlations are Pearson *r*s. Total N = 48, but varies due to non-response.

Factor I. This confirms Labov's general finding[9] that, regardless of the formality level, the relation between S.E.S. and level of formality is apparent. Factor I is consistently strongly correlated with all of the formality levels, whereas no significant relationships appear between the latter and Factor II. The Poverty Marker factor is significantly correlated only with formality level 3. The combined effect of all performance is revealed at the bottom of table 5, represented by the coefficient called the multiple *R*. This coefficient ranges from – 1 to + 1, the larger the number, be it positive or negative, the greater the strength of association. The square of the multiple *R* represents the per cent of variance explained in the linguistic data by all of the predicting variables (i.e. the three factors) acting together. The per cent of variance figures are shown in table 5, directly below the multiple *R*s. Also evident is a clear pattern of an increasing strength of relationship as the level of formality progressively shifts from the least formal (Levels 1 and 2) to the most formal (Level 5). Hence, in the least formal context the three dimensions of S.E.S. can explain only one-fifth of the variance in language performance, but at the most formal level, over one-third of the variance is explained. The significance of this finding is discussed below.

Discussion

The findings presented above have a number of interesting implications. First of all, we now see that by using more sensitive measures of S.E.S.

[9] The term 'general finding' is used because of Labov's finding of the hypercorrect (cross-over) pattern of the lower middle class.

(i.e. household artifacts, house construction materials, and other material demonstrators of wealth) we are able to predict more of the linguistic variance. This would imply that future social research on S.E.S. should include data on material demonstration of wealth, beyond the standardly used indicators (i.e. income, education and occupation).

Furthermore, all S.E.S. data that is gathered should be specific to the informants themselves, rather than generally characteristic of an ecological census unit. Gathering S.E.S. data on an informant directly, during the course of the interview, permits the researcher to later relate that informant's individual S.E.S. attributes to the variation in his linguistic production. This kind of approach undoubtedly will be more accurate in its findings than one which uses neighborhood census data as its source of S.E.S. information.

Corroboration presented here of the strong relationship between S.E.S. and language bears a particularly important message to social scientists whose primary interest is not linguistic or sociolinguistic. It tells them that sociolinguistic data is data about S.E.S., and that as such, rather than being the dependent variable which it is for sociolinguistics, it is an independent variable, constituting a *predictor* of social behavior.

That the relationship between S.E.S. and language is strong at all formality levels means that tapping linguistic data at any level would be fruitful. In other research (Seligson and Berk-Seligson 1978), for example, we have focused on Level 3 (word-naming), and have found language to be the best socio-economic predictor of three different modes of political participation (organizational participation, community improvement activism and interaction with local officials), better even than income, education, and occupation, which have typically been used in research of this nature. Although some may protest that the state of sociolinguistics is not sufficiently advanced, itself, for other disciplines to be adapting its methodology for their own particular purposes, others are openly encouraging its use as a tool of the other social sciences (Hymes 1970).

While research using linguistic variation at any formality level would be fruitful, our research in this study has singled out Level 5, the most formal level, as being most strongly related to standard measures of S.E.S. Thus, we found that as the formality of the speech situation increases the correlation between S.E.S. and language increases, reaching its peak at the most formal level (Level 5). Our explanation for this finding is that the more formal levels are least influenced by nonrandom biases introduced by the interview situation itself. As we pointed out earlier in

this paper, the data for this study was collected under an artificial situation: the respondent was aware that his speech was being recorded. Since the lower S.E.S. respondents in the sample, being interviewed by a considerably higher S.E.S. person, felt considerably more intimidated by the interview than did the upper S.E.S. respondents, whose status was more on par with that of the interviewer, we hypothesize that there was a tendency on the part of the former to be more conscious of their speech throughout the interview. The upper S.E.S. respondents, on the other hand, feeling less intimidated by the interview situation, on the whole were probably less careful with their speech. As a result, in the unstructured (i.e. less formal) part of the interview, the lower S.E.S. respondents consistently tried to use more of the prestige forms than they would have in more natural situations, whereas the upper S.E.S. respondents did not strive as much to do so. Hence, the correlation between S.E.S. and linguistic performance at the more informal levels was deflated by this nonrandom bias (i.e. poorer respondents performed better than expected, whereas richer respondents performed as expected). This parallels Labov's (1964a) finding regarding the hypercorrection of the lower middle class.

This bias became progressively less noticeable as the formality of the interview situation increased, for at the most formal levels all respondents recognized the situation as a test of their language ability, and hence strove to do as well as possible. Thus, at Level 5, in which respondents were asked to read a list of words, individuals were at their most careful; they all tried their best to produce prestige forms. Level 5, then, was free of most of the nonrandom biases that were found at the more informal levels.[10] If our findings can be generalized, then other investigators might seriously consider focusing on Level 5 in their research. We conclude then, that Level 5 makes the finest discriminations among respondents. Unfortunately, however, Level 5 requires that informants read, a requirement which is untenable for illiterate persons.

Conclusions

Using a sample of Costa Ricans, this study has found, first of all, that as the level of formality increases the use of prestige phonological forms

---

[10] The influence which the interviewer may have had upon the respondents in the course of the present investigation is similar to the impact which an experimental situation has upon any scientific study. The experimental situation always imparts a certain amount of 'energy' to what is being observed and hence alters its state. This phenomenon has come to be known as the 'Heisenberg Uncertainty Principle'.

increases. Secondly, a wide range of socio-economic variables beyond the common ones of education, income, and occupation were found to be associated with linguistic performance. Of these socio-economic variables a particular subset, consisting of household artifacts, house construction materials, and education, as grouped using factor analysis, were shown to be the strongest measures of linguistic performance. Finally, looking across formality levels using the statistical technique of multiple regression, it was found that the level of formality which was most highly correlated with the S.E.S. factors was the most formal level (operationalized as word-list reading).

Appendix

I. Accent shift (plus, optionally, loss of syllabicity)

aú ⋅ áu (→ àw)   *ataúd*: [ataùd] → [atáud] (⋅ [atáwd])
aó → áo (⋅ áw)   *ahogo*: [aógo] ⋅ [áogo] (→ [áwgo])
aí ⋅ ái (⋅ áy)   *maíz:* [maís] → [máis] (→ [mays])
ać → áe (→)   *maestro*: [maéstro] ⋅ ⋅ [máestro] (→ [máystro])
éa ⋅ eá (→ yá)   *océano*: [oséano] → [oseáno] (→ ⋅ [osyáno])
eí ⋅ ⋅ eí (→ éy)   *increíble*: [inkreíble] ⋅ [inkréible] (→ [inkréyble])
oí ⋅ ói (→ óy)   *oido*: [oído] → [óido] (⋅ [óydo])
Note: The arrow (⋅) simply represents 'alternates with', and makes no implications as to diachronic processes.

II. Vowel change

o → u   *poeta*: [poeta] ⋅ ⋅ [pueta]
e ⋅ i   *peor*: [peor] → [pior]
i → e   *policía*: [polisía] → [polesía]

III. Consonantal change

A. [d, l, r]
  1. [d] has the variants [l] or [r]
     admirar: [admirar] → [almirar] or [armirar]
  2. [l] has the variant [r]
     fácil: [fásil] ⋅ [fásir]
  3. [r] has the variant [l]
     comprarlo: [komprarlo] ⋅ [komprallo]

B. Nasals
  1. change of [n, m] to [ŋ] or [n] before a following [n]
     himno: [imno] ⋅⋅ [iŋno] or [inno]
  2. change of [g] to [ŋ] or [n] before a following [n]
     ignorante: [ignorante] ⋅ [iŋnorante] or [innorante]

C. Stops and fricatives

  1. consonant clusters

    (a) [p, b, k] followed by [t]

    (i) realization of the first consonant as some other stop

    aceptar: [aseptar]   · [asektar]

    or (ii) dropping the first consonant, and lengthening the second one

    [aseptar]   ·· [asettar]

    or (iii) simply dropping the first consonant

    [aseptar]   · [asetar]

    (b) [p, b, k] followed by [s]

    (i) realization of the first consonant as another stop

    apsurdo: [apsurdo] ·· [aksurdo] and optionally   · [awksurdo]

    (ii) voicing of the first consonant: [apsurdo]   · [absurdo]

    (iii) dropping of the first consonant

    [apsurdo]   ·· [asurdo]

    (especially frequent in the cluster [ks])

    exacto: [eksakto] ··· [esakto]

    (c) [t] followed by another consonant

    (i) realization of [t] as [k] *atlántico*:

    [atlantiko] → [aklantiko]

    (ii) voicing of [t]

    [atlantiko] →· [adlantiko]

    (iii) dropping of [t]

    [atlantiko] →· [alantiko]

  2. non-clusters

    realization of one stop or fricative as another stop or fricative

[b, b] · ·· [g, g]	*abuela*:	[abwela] ·· [agwela]
[d, d] ─· [g, g]	*padre*:	[padre] · · [pagre]
[p] · · [k]	*pupitre*:	[pupitre] · [kupitre]
[g] · [k]	*gangrena*:	[gangrena] · [kangrena]
[f] ─ · [h]*	*fuego*:	[fwego] · · [hwego]
[h] · [f]	*juego*:	[hwego] · · [fwego]

* Costa Rican Spanish regularly employs [h] where Peninsular Spanish and the Spanish of many Latin American countries regularly uses [x].

It should be recognized that the particular lexical items presented above are not isolated instances of the given variables, but rather, are simply examples of generalized phenomena which are found in the speech of all Costa Ricans to varying degrees. Thus, further examples of the categories mentioned above are:

I. Accent shift

    recaída, increíble, reir, océano, oído, periodo

II. Vowel change

    almohada, feo, toalla, calle, individuo

III. Consonantal change

    A. [d, l, r]

    advertir, ataúd, alcohol, sueldazo, verb infinitive .. direct/indirect object pronoun

B. Nasals
magnífico, alumno
C. Stops and fricatives
1. Consonant clusters
correcto, exacto, absoluto, concepción, cápsula, observar, aritmética, ritmo
2. Non-clusters
vomitar, huevo, fogón, jornalero, agonisar, fusil, madre, viruela

The preceding classification of variants into three major categories parallels Wilson's (1970: 39) categorization according to (1) accent shift, (2) changes in syllabic nuclei, and (3) nonstandard distributional changes of consonants, all three of which are shown by him to result from nonstandard syllable restructuring of the phonological word, tending toward open syllabicity. However, whereas Wilson's study examines additional processes (e.g. changes in the placement of syllable boundaries, epenthesis) that are related to syllable restructuring in Costa Rican Spanish, the sociolinguistic analysis presented here limits itself to the three categories outlined above.

References

Albó, X., 1970. Social constraints on Cochabamba Quechua. Ithaca, N.Y.: Cornell Univ. Ph.D. Dissertation Series.

Agüero, A., 1962. El español de América y Costa Rica. San José, C.R.: Editorial Antonio Lehmann.

Agüero, A., 1964. El español de Costa Rica y su atlas lingüístico. Separata del libro 'Presente y futuro de la lengua Española'. Madrid: Ediciones Cultura Hispánica.

Anshen, F., 1969. Speech variation among negroes in a small southern community. New York: New York Univ. (Ph.D. Diss.).

Arroyo Soto, V.M., 1971. El habla popular en la literatura costarricense. Costa Rica: Ciudad Univ. Rodrigo Facio (Tesis de grado, no. 18).

Azofeifa, I.F., 1947. Pronunciamos nuestra lengua y Conversaciones sobre literatura costarricense. San José: Librería Atenas.

Berk-Seligson, S., 1976. Costa Rican listener response to socially-based phonological variation: A subjective reaction test. Unpubl. MS.

Boni, F.D., and M.A. Seligson, 1973. Applying quantitative techniques to quantitative history: The case of Mexico. Latin American Research Review 8, 105–110.

Canfield, D.L., 1962. La pronunciación del espannol en América. Bogotá: Publicaciones del Instituto Caro y Cuervo, XVII.

Dirección General de Estadística y Censos, 1974. Censos Nacionales de 1973, población. San José.

Fishman, J.A., 1971. Sociolinguistics: A brief introduction. Rowley, Mass.: Newbury House Publishers.

Fishman, J.A., R.L. Cooper, R. Ma, et al., 1971. Bilingualism in the Barrio. Bloomington: Indiana Univ. Publications, Language Science Monographs.

Giglioli, P.P. (ed.), 1972. Language and social context: Selected readings. Harmondsworth: Penguin.

Grimshaw, A.D., 1966. Directions for research in sociolinguistics: Suggestions of a non-linguist sociologist. In: Stanley Lieberson (ed.), Explorations in sociolinguistics. The Hague: Mouton.

Gumperz, J., 1972. The speech community. In: P.P. Giglioli (ed.), Language and social context: Selected readings. Harmondsworth: Penguin.

Hasan, R., 1973. Code, register and social dialect. In: B. Bernstein (ed.), Class, codes and control. Vol. 2 of: Applied studies towards a sociology of language. London: Routledge and Kegan Paul.

Holt, R.T., and J.E. Turner (eds.), 1970. The methodology of comparative research. New York: The Free Press.

Hymes, D., 1970. Linguistic aspects of comparative political research. In: R.T. Holt and J.E. Turner (eds.), The methodology of comparative research. New York: The Free Press.

Kučera, H., 1961. The phonology of Czech. The Hague: Mouton.

Labov, W., 1964a. Hypercorrection by the lower middle class as a factor in linguistic change. In: W. Bright (ed.), Sociolinguistics. Proceedings of the UCLA Sociolinguistics Conference, 1964. Los Angeles: UCLA.

Labov, W., 1964b. The phonological correlates of social stratification. American Anthropologist 66 (no. 2, Part 2).

Labov, W., 1965. Stages in the acquisition of Standard English. In: R. Shuy (ed.), Social dialects and language learning. Champaign, Ill.: National Council of Teachers of English.

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

Labov, W., 1972a. Sociolinguistic patterns. Philadelphia: Univ. of Pennsylvania Press.

Labov, W., 1972b. The study of language in its social context. In: P.P. Giglioli (ed.), Language and social context: Selected readings. Harmondsworth: Penguin.

Labov, W., P. Cohen, C. Robins, and J. Lewis, 1968. A study of the non-standard English of Negro and Puerto Rican speakers in New York City. Final Report, Coop. Research project 3288. 2 vols. Philadelphia, Pa.: U.S. Regional Survey.

Lambert, W.E., 1967. A social psychology of bilingualism. The Journal of Social Issues 23, 91–109.

Lambert, W.E., M. Anisfeld, and G. Yeni-Komshian, 1965. Evaluational reactions of Jewish and Arab adolescents to dialect and language variation. Journal of Personality and Social Psychology 2, 84–90.

Lambert, W.E., H. Frankel, and G.R. Tucker, 1966. Judging personality through speech: A French-Canadian example. The Journal of Communications 4, 305–321.

Lapesa, R., 1968. Historia de la lengua española, séptima edición. Madrid: Escelicer.

Lavendera, B.R., 1974. On sociolinguistic research in New World Spanish: A review article. Language in Society 2, 247–292.

Levine, L., and H.J. Crockett, Jr., 1966. Speech variation in a Piedmont community: Post-vocalic r. In: S. Lieberson (ed.), Exploration in Sociolinguistics. Sociological Inquiry 36 (no. 2).

Ma, R., and E. Herasimchuk, 1971. The linguistic dimensions of a bilingual neighborhood. In: J.A. Fishman et al. (eds.), Bilingualism in the Barrio. Bloomington: Indiana Univer. Publications, Language Science Monographs.

Macaulay, R.K.S., 1970. Review of 'A Sociolinguistic description of Detroit Negro speech'. Language 46, 764–773.

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

Menéndez Pidal, R., 1966. Manual de gramática histórica española (6th ed.). Madrid: Espasa-Calpe.

Pahlsson, C., 1972. The Northumbrian burr: A sociolinguistic study. Lund: Berlingska Boktryckeriet.

Rummel, R.J., 1967. Understanding factor analysis. Journal of Conflict Resolution 11, 444–480.

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

Sankoff, G., and H. Cedergren, 1971. Some results of a sociolinguistic study of Montreal French. In: R. Darnell (ed.), Linguistic diversity in Canadian society. Edmonton, Canada: Linguistic Research Inc.

Seligson, M.A., 1974. The peasant and agrarian capitalism in Costa Rica. Pittsburgh: Univ. of Pittsburgh (Ph.D. diss.).

Seligson, M.A., 1975. Agrarian capitalism and the tranformation of peasant society: Coffee in Costa Rica. Buffalo: State Univ. of New York Special Studies Series.

Seligson, M.A., 1977a. Agrarian policy in dependent societies: Costa Rica. Journal of Inter-American Studies and World Affairs 19 (May), 201–232.

Seligson, M.A., 1977b. Prestige among peasants: A multidimensional analysis of preference data. American Journal of Sociology 83 (November), 632–652.

Seligson, M.A., 1978. Development and participation in Costa Rica: The impact of context. In: J.A. Booth and M.A. Seligson (eds.), Political participation in Latin America, Vol. 1: Citizen and State. New York London: Holmes and Meier Publishers, in press.

Seligson, M.A., forthcoming. Unconventional political participation: Cynicism, powerlessness, and the Latin American peasant. In: M.A. Seligson and J.A. Booth (eds.), Political participation in Latin America, Vol. 2: Politics and the poor. New York/London: Holmes and Meier Publishers.

Seligson, M.A., and S. Berk-Seligson, 1978. Language and political behavior: A methodology for utilizing the linguistic component of socio-economic status. American Journal of Political Science (August), forthcoming.

Seligson, M.A., and J.A. Booth, forthc. Structure and levels of political participation in Costa Rica: Comparing the countryside with the city. In: M.A. Seligson and J.A. Booth (eds.), Political participation in Latin America, Vol. 2: Politics and the poor. New York London: Holmes and Meier Publishers.

Shuy, R.W., W.A. Wolfram, and W.K. Riley, 1967. Linguistic correlates of Social stratification in Detroit speech. Washington, D.C.: Office of Education Cooperative Research, Project No. 6-1347.

Shuy, R.W., W.A. Wolfram, and W.K. Riley, 1967. 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, P., 1972. Sex, covert prestige and linguistic change in the urban British English of Norwich. Language in Society 1, 179–195.

Tucker, G.R., and W.E. Lambert, 1972. White and Negro listeners' reactions to various American-English dialects. In: J.A. Fishman (ed.), Advances in the society of language. Vol. II. The Hague: Mouton.

Wilson, J.L., 1970. A generative phonological study of Costa Rican Spanish. Ann Arbor: Univ. of Michigan (Ph.D. diss.).

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